Master's Thesis : Evaluating outcome following knee arthroplasty using inertial measurement units

Auteur : Cantamessa, Astrid
Promoteur(s) : Geris, Liesbet
Faculté : Faculté des Sciences appliquées
Diplôme : Master en ingénieur civil biomédical, à finalité spécialisée
Année académique : 2019-2020
URI/URL : http://hdl.handle.net/2268.2/10555

UNIVERSITY OF LIEGE
FACULTY OF APPLIED SCIENCES

UNIVERSITY OF AUCKLAND
AUCKLAND BIOENGINEERING INSTITUTE

Evaluating outcome following knee arthroplasty using inertial measurement units

Master thesis conducted by

ASTRID CANTAMESSA

with the aim of obtaining the degree of Master in Biomedical Engineering

Under the supervision of

LIESBET GERIS

THOR BESIER

PAUL MONK

ACADEMIC YEAR 2019-2020

Abstract

Osteoarthritis is a frequent and debilitating disease whose burden is set to increase given our ageing population. Arthroplasty is the only curative treatment for end-stage arthritis and can greatly improve joint function, control pain and enhance quality of life. However, surgery is only part of the picture, and ensuring successful outcomes requires both extensive tailored physiotherapy and close patient monitoring for complications. Currently, patient reported outcome measures (PROMs) are used with minimal clinical follow-up. Not only does this allow for limited opportunities to assess postoperative function, but PROMs are also inherently subjective. As such, the orthopaedic clinic lacks of quantitative information with which to actively monitor a patient's progress. In addition, due to resource limitations, it struggles to closely monitor patients during the first six weeks following surgery, a key period for ensuring adequate long-term joint function. Inertial measurement units (IMUs) provide an opportunity to objectively measure important biomechanical gait variables in both clinic and home settings. This allows clinicians and physiotherapists to remotely monitor patients through cloud-computing technologies.

The aim of this thesis, which is part of a larger research project at the Auckland Bioengineering Institute (Auckland, New Zealand), is to develop and assess a new workflow based on machine learning algorithm to quantitatively evaluate joint function during walking gait of patients following knee replacement surgery using only two ankle-worn IMUs. To evaluate this algorithm, predictions of joint kinematics were compared to 'ground truth' joint kinematics recorded from optical motion capture.

Twelve patients undergoing knee arthroplasty were recruited. They participated in two gait sessions before and around six weeks after their surgery during which optical marker trajectories and acceleration and angular velocity from IMUs were recorded. However, in view of the issues encountered with their quantity and quality, two other datasets, previously collected for other studies, were also exploited. One involved ten healthy volunteers performing treadmill walking and the second was composed of four overground walking healthy participants.

Two types of models were generated and evaluated: a personalised model, trained on a portion of a subject's data and predicting the remaining part, and a generalised model, trained on every individual of the cohort but one used for prediction. Moreover, a sensitivity analysis was performed to select the most optimal combination of parameters and data processing ways. Our method enables to predict knee kinematics with more than 95% accuracy for personalised models. This also holds for the treadmill generalised model. However, the poor performance of the overground generalised model was due to limited number of steps per person which could not capture the variability within the dataset.

The continuation of this study should increase the patient dataset and include other motions than walking. Moreover, information obtained about the outcome recorded in patients' environment will be contrasted with other metrics (PROMs, range of motion) collected during their clinical follow-up. Ultimately, this may help clinicians to identify potential complications during recovery and provide the opportunity for early intervention.

Acknowledgments

This master thesis is the fruit of long months of work punctuated by various challenges and an unpredictable worldwide pandemic crisis. I have learned a lot during this time both professionally and personally. I have been helped and supported by many people along this somewhat eventful path and I would like to express them all my gratitude.

First of all, I would like to thank Professor Liesbet Geris, my supervisor from the University of Liège, who gave me the opportunity to carry out my master thesis in New Zealand. She was very supportive regarding the unfortunate and unexpected situation.

Then, I would like to express all my gratitude to Professor Thor Besier and Doctor Paul Monk, my two co-supervisors from the Auckland Bioengineering Institute, for the opportunity to work for such an interesting project within an amazing institution. Their warm welcoming and ongoing assistance helped me integrate into an environment different from anything I had known before. They were always willing to help me when needed and I am very grateful to them for supporting my decision to come back in Belgium earlier due to the Covid-19 pandemic. A special thanks for the assistance that Professor Besier provided me after my return with our weekly meetings and his constant follow-up of my work, even with the difficulties associated with the time difference between our two countries.

I also want to thank the other members of my examination committee for taking the time to read this thesis and to listen to my defence.

I am also extremely grateful to Ted Yeung, the postdoc research engineer with whom I have worked all along my internship. He was always present to answer all my questions, which were numerous, and he taught me the basics of python programming language and of the use of Nexus software from Vicon. He was the one who developed the machine learning algorithm that has been used to obtain the results of this work.

Afterwards, I would like to thank every person of the team involved in the study: Shasha, Scott and Bruno. They all took time to update me on the study progress at my arrival and to explain to me the conduct of the gait lab and physiotherapy sessions. Furthermore, I am thankful to all my colleagues from the Auckland Bioengineering Institute and from the company *IMeasureU* for their kind welcoming and for making these two places very convivial to work in.

Moreover, I am deeply grateful to my family for their unconditional moral support throughout my studies and for their financial contribution to this amazing opportunity that I have been given to carry out my internship on the other side of the world.

Finally, I would like to express my thanks to all my friends with whom I have shared the ups and downs of these past five years of study, especially Sébastien, Ani, François and Arthur. Their help and encouragement were greatly valued all of this time.

Contents

Introduction

Osteoarthritis is a common and debilitating disease whose burden is expected to rise given our ageing population [1] [2]. Arthroplasty, also known as joint replacement surgery, is the only curative treatment for end-stage arthritis and can greatly improve joint function as well as reduce pain, enhancing quality of life [3]. However, surgery is only part of the picture, and ensuring successful outcomes requires both extensive tailored physiotherapy and close patient monitoring for complications. Encompassing more than just clinical outcomes, success is a term that has been complex to quantify in an orthopaedic setting. Realising the importance of a patient's view on their surgery and the result following it, there has also been a shift to collect patient reported outcome measures (PROMs) to monitor an individual's progress and gauge of success [4] [5] [6]. Currently, PROMs are used with minimal clinical follow-up occurring at two, six, twelve, and fifty-two weeks post-surgery [7]. Not only does this allow for limited opportunities to assess postoperative function, but PROMs are also inherently subjective. They have been shown to over-report the level of actual physical function and a ceiling effect has frequently been observed, whereby scores fail to discriminate patient outcomes. Moreover, there is reported to be poor correlation between PROMs and objectively assessed physical function [4] [8] [9]. As such, the orthopaedic clinic lacks of quantitative information with which to actively monitor a patient's progress. In addition, due to resource limitations, it struggles to closely monitor patients during the first six weeks following surgery, a key period for ensuring adequate long-term joint function [4]. Furthermore, the recent restrictions on patient contact related to the Covid-2019 pandemic has highlighted the need for more extensive adoption of remote monitoring technologies in orthopaedics.

Since the primary goal of lower limb joint arthroplasty is to restore ambulation, biomechanical variables can be useful clinical measures in the place of a personal surgical review or clinical function check by a physiotherapist. Traditionally, biomechanical measures are obtained from three-dimensional gait analysis, which, despite providing an accurate assessment of knee function, is time-consuming and an expensive process. Moreover, the field is limited to a laboratory setting, which does not reflect real-world mobility patterns and does not account for total load exposure [10]. Wearable sensors, especially inertial measurement units (IMUs), provide an opportunity to objectively measure important biomechanical gait variables in both clinic and home settings. This allows clinicians and physiotherapists to remotely monitor patients through cloud-computing technologies.

Obtaining accurate joint kinematics from IMUs currently requires using sensors located on each side of the joint, often with a complex kinematic model that necessitates computation and known person-specific parameters (i.e. IMU placements relative to the body). Furthermore, the resulting kinematics are prone to errors such as numerical drift coming from estimating orientations from angular velocity [11]. This holds true for many types of human movements from lower limb gait to upper limb reaching tasks. However, for motion with a cyclic nature such as walking gait, machine learning can be used to predict knee flexion angles from raw IMU data – acceleration and angular velocity. The aim of this thesis, which is part of a larger research project at the Auckland Bioengineering Institute (Auckland, New Zealand), is thus to develop and assess a new workflow based on machine learning algorithm to quantitatively evaluate joint function during walking gait of patient following knee replacement surgery using only two ankle-worn IMUs.

Data to test this algorithm were collected from twelve patients undergoing knee arthroplasty during gait sessions planned before and around six weeks after the surgery. Subject follow-up with IMUs tracking was also performed at their physiotherapy sessions and during the rest of the day, from their hospital release to the post-operative gait. During this patient monitoring, other metrics were assembled, including PROMs and range of motion (ROM). However, in view of the issues encountered with the data quality and quantity, two other datasets, previously collected for others studies, were also exploited. One involved ten healthy volunteers performing treadmill walking and another one composed of four overground walking healthy participants. Algorithm predictions of joint kinematics were compared to 'ground truth' joint kinematics recorded from optical motion capture.

This thesis is divided into four main chapters. The first one is dedicated to the introduction of the general context of the study. Chapter two describes the timeline routine followed by the patients enrolled in this study, the methodology pursued to collect and process the data and a description of the machine algorithm pipeline. The third chapter concerns the evaluation of this algorithm. The results obtained from the two generated model–personalised and generalised models–are illustrated and discussed. Moreover, a sensitivity analysis was performed to select the best parameters and data processing. Chapter four aims at presenting and contrasting the other data from different metrics collected during the patient's follow-up. The general tendency of the evolution of the recovery is reported followed by an analysis of individual results. Finally, the last chapter presents the limitations of the study and a critical view of the results. Ideas of improvement for future work are also discussed.

Chapter 1

Background

The aim of this chapter is to provide the general context of this project. First, a description of the various components that make up the knee is presented. Osteoarthritis, the condition that is responsible for more than 95% of knee replacement surgery is presented in the next section, followed by a focus on joint arthroplasty. After, inertial measurement units are introduced and their applications in the biomedical field are emphasised. Finally, the main focus of this master thesis is outlined in the last section. Two short notes are added about the impact of the Covid-19 on this work as well as on my personal contribution for this study.

1.1 Description of the Knee Anatomy

As the main focus of this thesis concerns patients' rehabilitation following knee replacement surgery, a description of this joint anatomy is required before exploring the operation in itself. The next section briefly investigates the key players of this peculiar component of the musculoskeletal system: joints, ligaments, muscles, cartilage, fibrous capsule and synovial membrane.

The knee is a complex modified synovial hinge joint, the largest of the human body, connecting the femur to the tibia and fibula [12] [13]. It allows movement in flexion-extension about a transverse axis passing through the centre of the knee, as well as internal-external rotation, to a lesser extent, about a longitudinal axis passing through the centre of the knee (see Fig. 1.1). These motions are enabled thanks to the three joints that make up the knee: the patellofemoral joint, linking the patella to the femur, and the lateral and medial tibiofemoral joints, between the femur and the tibia. It owes its great stability during a variety of loading situations to the association of the ligaments and muscles [12] [13] [14].

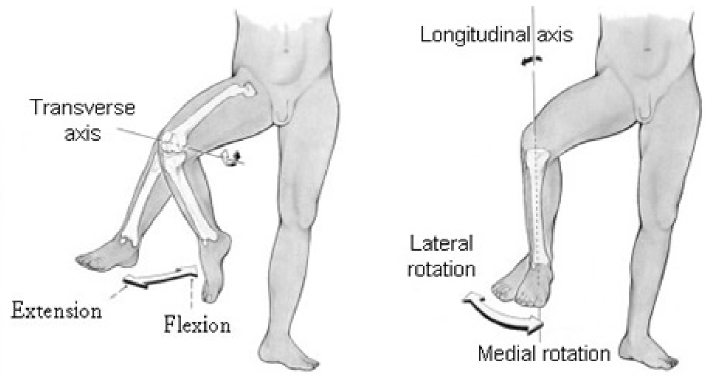

Figure 1.1: Movement of the knee joint (adapted from [10]).

1.1.1 Bones

The knee joint is made up of the distal end of the femur, the proximal ends of the tibia and fibula and the patella. These four bones are shown in Figure 1.2.

The **femur** is the longest bone in the body. The rounded protuberances located at the end of this bone are the *medial* and *lateral condyles* and are the part of the joint which articulates with the tibia. Moreover, the outermost projections on both condyles are called *medial* and *lateral epicondyles* and are the site of attachment of ligaments and tendons to the bone. A layer of *hyaline cartilage* covers the surface of the femoral end[12] [15].

The **tibia** is the bone running from the knee to the ankle. Its top part is also made up of medial and lateral condyles, relatively flat, topped by two shock-absorbing cartilaginous *menisci*. The tibial tubercle is a knuckle-like protuberance on the front of the tibial end where the patellar tendon attaches [15].

As mentioned previously, the *medial* and *lateral tibiofemoral joints* link the distal femur to the proximal tibia. Indeed, the femoral condyles articulate with the corresponding tibial ones. This joint is the weight-bearing component of the knee joint [12] [13].

The **patella** is a small, triangular, sesamoid bone sitting at the front of the knee and embedded with the tendon of quadriceps [12]. Its function is to increase the force generated by the quadriceps muscle by improving its leverage. It is also a bony shield for the knee and prevents excessive friction between the quadriceps tendon and the femoral condyles [12] [14]. The patella glides within the *patellofemoral groove*, formed by the two femoral condyles [16]. This peculiar articulation is the *patellofemoral joint*. It takes part to the knee extension by increasing the efficiency of the quadriceps muscle (this will be explained in more details in the section 1.1.6) [12].

The **fibula** is the long and thin bone next to the tibia. It is the site of attachment of the *biceps femoris* tendon, member of the *hamstring* group and of the *lateral collateral ligament*. It also helps to carry some weight but to a much lesser extent than the tibia [16].

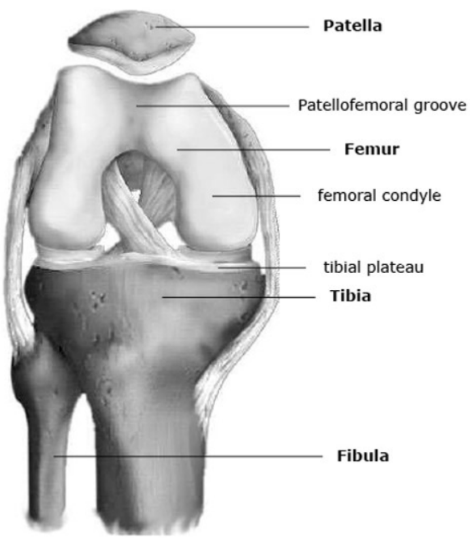

Figure 1.2: Schematic representation of the anterior view of the knee joint and emphasis on the bones (adapted from [16]).

1.1.2 Cartilage

There are two types of cartilage in the knee: *hyaline* or *articular cartilage* and *fibrocartilage* which constitutes the *lateral* and *medial menisci* [17].

The **articular cartilage** covers joint surfaces, i.e. the bottom of the femur, the top of the tibia and the back of the patella for the knee joint as represented in Figure 1.4. It is a tri-phasic material composed of a porous matrix made up of 10–20% in weight of type II collagen and 4–7% of proteoglycans, 70–90% of water and ions (to balance the negatively charged proteoglycans). It also contains other structural macromolecules such as non-collagenous proteins and glycoproteins. Its structure makes it slippery, strong and flexible which serves its two main functions. The first one is to allow the bones to glide over each other during motion by decreasing the friction forces and wear. Its second purpose is to act as shock absorber, preventing bones against impacting each other. There is no blood supply in cartilage, nutrients are brought by the flow of water through the diffusion process [17].

Cartilage is divided into four zones that differ in their composition, organisation, cells shape and mechanical properties. These areas are represented in Figure 1.3. The peripheral one, the *superficial zone*, is the thinnest one (10–20% of the cross section) and is in contact with the synovial fluid (see Section 1.1.3). It contains the lowest amount of proteoglycans and the highest of collagen arranged in thin fibres parallel to the articular surface as well as elongated chondrocytes. This zone is designed to reduce the friction and to enable cartilage to resist shear, tensile and compressive stresses generated during joint use. This area is followed by the *middle zone*, corresponding to 40–60% of the cartilage cross section. The collagen fibres are larger and randomly arranged and the chondrocytes shape rounder. The main role of this area is to absorb shocks. The *deep zone* (30% of the cross section) contains the highest proteoglycans content and the lowest water content. Collagen fibrils reach their largest diameter and are arranged perpendicular to the surface. These properties make this zone the strongest to resist to compressive forces and optimised to attach to the bone. Finally, the *calcified cartilage zone* is separated from the others by the *tidemark* and is designed to anchor the cartilage and its collagen fibrils to the underlying subchondral bone [17].

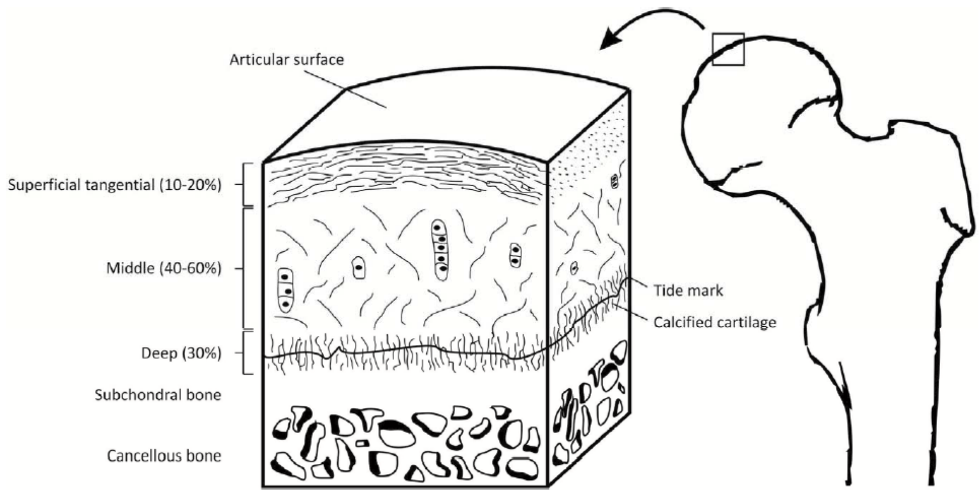

Figure 1.3: Schematic representation of the four zones that makes up the articular cartilage (adapted from [17]).

The **lateral** and **medial menisci**, shown in Figure 1.4 are thick layers of fibrocartilage located on top of the tibial condyles and articulating with the femoral condyles. Their superior surface is concave to accommodate for the femoral condyles and the anterior surface is flat. This shape enables to deepen the articular surface of the tibia, assisting and coordinating knee motion and therefore increases the stability. The menisci also act as shock absorbers and improve the weight distribution from the femur by increasing the area of contact on the tibia. Indeed, the arrangement of the fibres enables axial loads to be dispersed radially, which decreases the wear on articular cartilage [12] [14] [15].

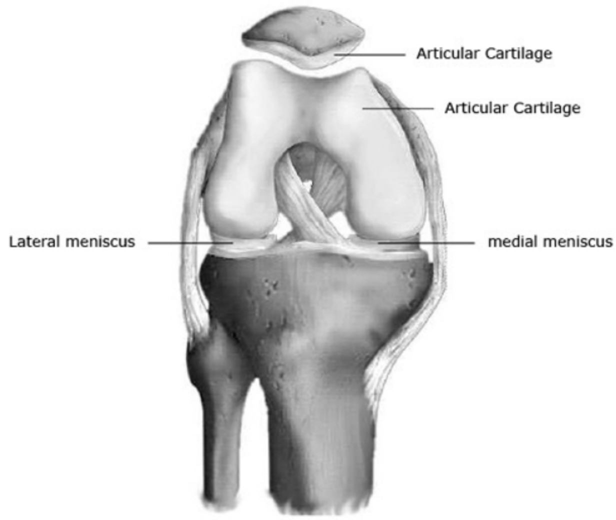

Figure 1.4: Schematic representation of the anterior view of the knee joint and emphasis on the cartilage (adapted from [16]).

1.1.3 Joint capsule, synovial membrane and synovial fluid

The **joint capsule** is a thick and fibrous structure that encapsulates the knee. Its outer layer is composed of fibrous connective tissue in continuity with the ligaments in order to hold the knee in place, therefore enhancing its stability [15]. The capsule contains **synovial fluid** surrounded by the thin **synovial membrane**. Its role is to lubricate the knee joint to reduce friction and wear and to provide nutrients and remove waste from the articular cartilage with nutrients [12]. These components as well as the articular cartilage are the key features that make the knee a *synovial joint* [18]. They are represented in Figure 1.5.

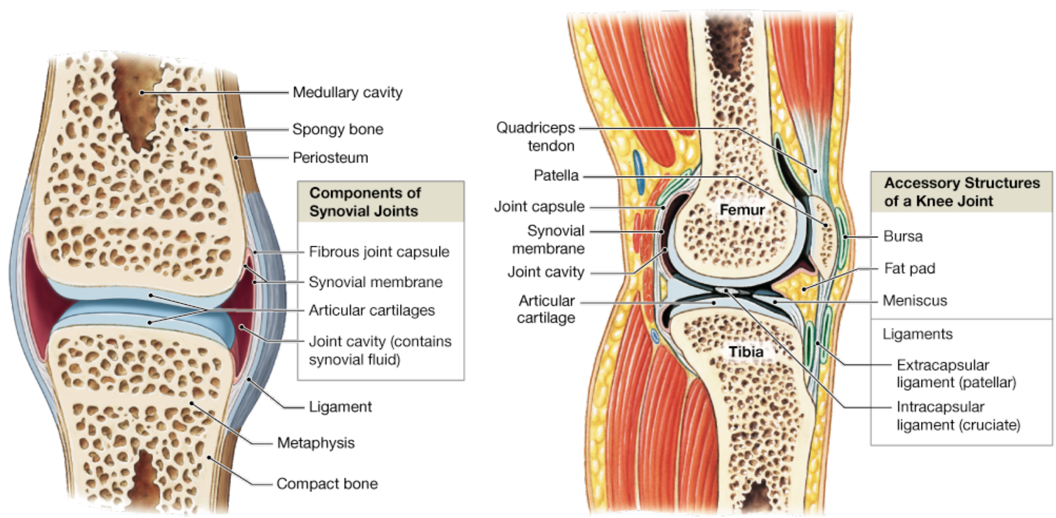

Figure 1.5: Schematic representation of a sagittal section of the knee joint and emphasis on the capsule, synovial membrane and synovial fluid (adapted from [18]).

1.1.4 Bursae and articular fat pads

The knee joint contains up to fourteen **bursae** which are pockets of synovial fluid of different sizes located between bones and soft tissues such as tendons, ligaments and muscles. Again, they help to cushion the joint and to reduce friction between adjacent moving structures [13]. For example, the *prepatellar bursa*, one of the largest, can be found just under the skin, in front of the patella and protects it [16]. Another one, the *infrapatellar bursa*, lies between the patellar tendon and the tibia, therefore preventing friction between these two [13]. Some of these bursae are shown in Figure 1.6.

The last component helping in the knee protection from external stresses are the **articular fat pads**. They are pockets of adipose tissue located around the knee and are represented in Figure 1.6 [15].

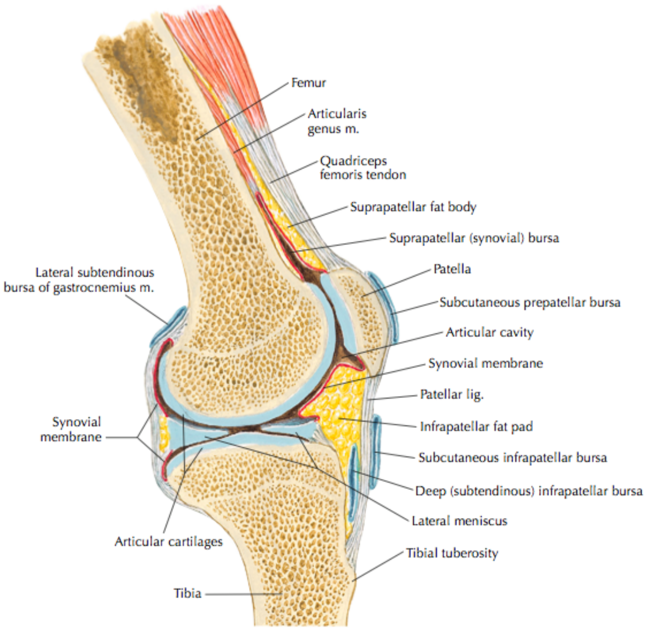

Figure 1.6: Schematic representation of a sagittal section of the knee joint and emphasis on the bursae and articular fat pads (adapted from [19]).

1.1.5 Ligaments

The ligaments in the knee connect bone to each other and act as joint stabilisers. They are composed of 70–80% in dry weight of collagen, 1–15% in dry weight of elastin and 55–65% in wet weight of water [17]. The four main ligaments of the knee are the *medial* and *lateral collateral ligaments* and the *anterior* and *posterior cruciate ligament* [12]. They are shown in Figure 1.7.

The **medial collateral ligament (MCL)** is located on the medial side of the joint and attaches the medial epicondyle of the femur to the medial condyle of the tibia. Its role is to stabilise the hinge motion of the knee by limiting any medial movement (such as valgus forces[1]) [12] [13] [15].

Conversely, the **lateral collateral ligament (LCL)** can be found on the lateral border of the knee and runs from the lateral femoral epicondyle to the lateral side of the fibula. It prevents forces applied to the medial side of the knee (such as varus forces[2]) from moving it laterally [12] [13] [15].

The **anterior cruciate ligament (ACL)** runs from the anterior intercondylar region of the tibia to the posteromedial aspect of the lateral femoral condyle. Its purpose is to resist anterior movement of the tibia relative to the femur and therefore preventing hyperextension. Moreover, it helps to limit some rotation and sideways motion (varus and valgus forces) of the knee [12] [13] [15].

[1]Forces that causes excessive outward angulation (away from the midline of the body) of the distal segment of an anatomical part. [20]

[2]Forces that causes excessive inward angulation (toward the midline of the body) of the distal segment of an anatomical part. [21]

Conversely, the **posterior cruciate ligament (PCL)** attaches the posterior intercondylar space of the tibia to the inner surface of the medial femoral condyle. It prevents posterior sheering forces of the tibia relative to the femur and helps to resist rotation and valgus and varus forces [12] [13] [15].

Figure 1.7: Schematic representation of the anterior view of the knee joint and emphasis on the main ligaments (adapted from [16]).

1.1.6 Muscles around the knee

The role of the muscles around the knee is obviously to produce motion but also to act second knee stabilisers, in addition to the ligaments. The two predominant groups of muscles involved in the knee motion are the *quadriceps* and the *hamstring* groups [10]. Other muscle also take part in the joint movement, albeit to a lesser extent. Some of them are shown in Figure 1.8.

The **quadriceps** muscles are located on the anterior aspect of the thigh and are responsible for the knee extension as well as the hip flexion. They are composed of the *rectus femoris*, *vastus lateralis*, *vastus intermedius* and *vastus medialis* [10]. Each originates from the top of the femur, and they all converge near the knee into the quadriceps tendon. This last one covers the patella and connects the quadriceps to the tibial tuberosity. The extension of this tendon from the patella to the tibia is known as the *patellar tendon* (or ligament since it is a bone to bone connection). It is one of the most important in the knee joint due to its important role in extension motion [12] [16] [13].

Conversely, the **hamstrings** group can be found on the posterior aspect of the thigh. This group consists of three muscles: the *biceps femoris*, *semimembranosus* and *semitendinosus* [10]. Each originates from the bottom of the pelvis and attaches to different sites to the back of the knee (to the head of the fibula for the biceps femoris and to the medial part of the tibia for the other two). They are the main knee flexors and play a secondary role in the hip extension. Moreover, they enable knee rotation in addition to other muscles [13] [15].

Another group of muscles involved in the knee motion to a lesser extent are the **calf muscles** composed of the *gastrocnemius* and *soleus*. They are located on the back of the lower leg and merge into the Achilles tendon. Their primary role concerns the ankle movement and more accurately the plantar flexion. However, they also help to bend the knee [10] [15].

The **gracilis** and **sartorius** muscles are two muscles running from the anterior pelvis to the medial tibia where they converge with the *semitendinosus* muscle into a tendinous attachment called *pes anserinus*. They take part to the knee flexion to a lesser extent, as well as the internal (medial) rotation [13] [15].

Other muscles such as the **popliteus**, the **gluteal muscles** and the **iliotibial band** are also involved in the knee motion but not as much as the previous ones [13] [15].

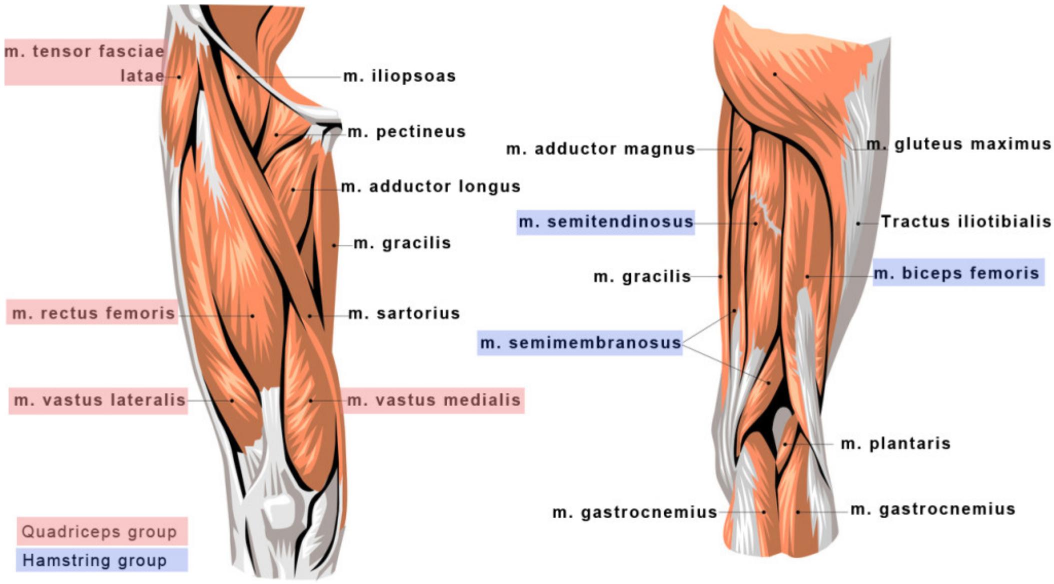

Figure 1.8: Schematic representation of the anterior (left) and posterior (right) view of the knee joint and emphasis on the muscles involved in its motion (adapted from [22]).

1.1.7 Neurovasculature

Blood is supplied to the knee through the genicular branches of the femoral and poplitaeal arteries, the circumflex fibular arteries, and the recurrent branches of the anterior tibial artery. As for the nerve supply, one can mention the femoral, tibial and common fibular nerves [13].

1.2 Osteoarthritis

According to the New Zealand Joint Registry covering data over a period of twenty years from January 1999 to December 2018, the main reason leading to knee arthroplasty is osteoarthritis in almost 95% of the cases [23]. In the 2018 annual report of the Belgian Hip and Knee Arthroplasty Registry, this percentage is similar [24]. This condition is the most common joint disease worldwide and among the most frequent health problem for middle-aged and older people with the prevalence of 30% in those over 75 years old [3] [25]. This is why this section examines the epidemiology, pathogenesis, diagnosis and treatment of osteoarthritis.

1.2.1 Epidemiology

Osteoarthritis affects people all around the world, ethnically diverse, with a higher prevalence in women than in men. The most common site of this disabling condition is the knee, the joint of interest in this study, followed by the hand and the hip [1] [2] [26]. Many risk factors have been identified with more or less likelihood to be related to the development or progress of the disease. The most important one is the age, which strongly correlates with the incidence of osteoarthritis [1] [2] [26]. Indeed, since the proportion of the population over 40 year old increases, the total number of people suffering from this health problem is rising too. Moreover, this is linked to the accumulation of biological age-related changes in the joint structure as well as the loss of mechanical properties. The higher incidence related to the three risk factors age, gender and site is illustrated in Figure. 1.9 It represents the age-specific and gender-specific incidence of the three main sites of osteoarthritis for the general population of Catalonia in Spain [1].

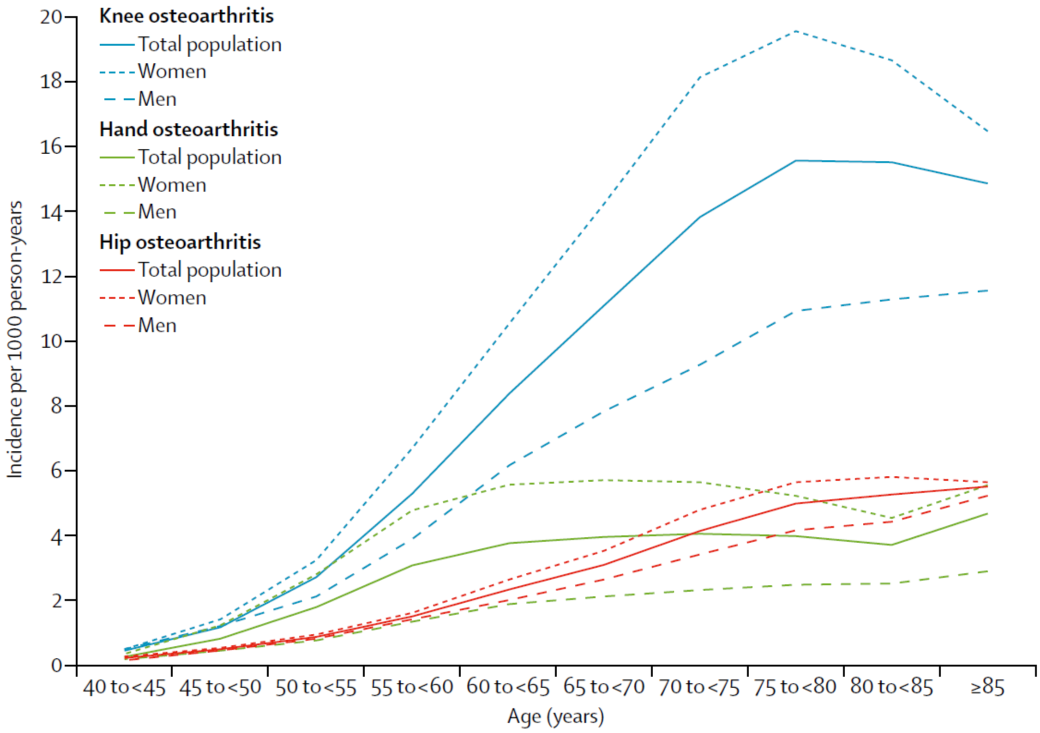

Figure 1.9: Representation of the age-specific and gender-specific incidence [per 1000 person years] of knee, hand and hip osteoarthritis for the general population of Spain (adapted from [1]).

A bad joint mechanics, both functionally and anatomically, is also considered as a risk factor but to a lesser extent. The tibial and femoral bone morphology, the limb alignment (varus or valgus alignment) or the leg length inequality are anatomical factors that could influence the development and the progression of knee osteoarthritis [2] [26]. On the functional side, a weakness in the knee extensor muscles, the quadriceps, is also considered as a weak risk [1] [2].

Another danger for the development and progression of osteoarthritis concerns heavy work activities. Indeed, repetitive joint overuse such as frequent kneeling, and excessive mechanical loading could exceed the ability of a joint to repair or to maintain itself. This also applies to obesity since it implies extra load on weight-bearing joints such as the knee [1] [26].

Previous knee injury can cause damage to any joint components (bone, cartilage, ligaments, menisci. . .) which could affect its biomechanics or make it more susceptible for further deterioration. Therefore, it would lead to an increased risk of progressive joint degeneration. These factors also comprise some high-impact sports including football or handball, since they are strongly related to knee injuries by exposing the joints to high levels of shock or torsion [1] [26].

Finally, genetic predisposition is also one of the risk factors that have been detected many years ago. Eleven loci have been identified to be associated with osteoarthritis by different genome-wide association studies [27]. Moreover, single-nucleotide polymorphisms have also been correlated with various known risk factors such as body mass index, bone mineral density,. . . However, the effect size is relatively small [1] [2].

1.2.2 Pathogenesis

Osteoarthritis is the resultant of biological and mechanical phenomena which destabilise equilibrium between the synthesis and degradation of cartilage [25]. This disease does not only affect the articular cartilage but the entire joint including alterations of the underlying subchondral bone, inflammation in the synovial tissue, weakness of the periarticular muscles,. . . [1] [2] [28] These changes are responsible for the symptoms (pain, restriction of motion, crepitus,. . .).

As previously mentioned, the role of chondrocytes is to develop, maintain and repair the extracellular matrix that makes up the cartilage. They are responsible for synthesising its components but also the proteolytic enzymes which degrades it. Therefore, osteoarthritis is a consequence of the failed attempt of chondrocytes to maintain the homoeostasis between the synthesis and the degradation of the matrix components [28]. As stated above, age is the number one risk factor for osteoarthritis. With age, the capacity of chondrocytes to synthesise proteoglycans, their proliferative capacity and their response to anabolic stimuli, like growth factors, decreases. Consequently, the maintenance and restoration of articular cartilage by chondrocytes are tainted, increasing the risk of development and progression of joint degradation [29]. Initial cartilage degradation could also be the results of a trauma leading to microfractures or inflammation which increases its susceptibility for disruption by physical forces [2] [28].

The pathophysiology of osteoarthritis can be divided into three overlapping stages. Firstly, damage or alteration of the extracellular matrix at the molecular level. Then, chondrocytes react to this damage. Finally, since their synthetic response decline, cartilage is not restored and there is a progressive loss of tissue [29].

The changes enforced by early stage osteoarthritis come in the form of erosion of the articular cartilage in the superficial zone; followed later by deeper fissures throughout the zones until reaching the subchondral bone [1] [29]. Free particles of extracellular matrix from cartilage tearing are released into the joint space and diffuse into the synovial fluid. This activates synoviocytes to release pro-inflammatory mediators and degrading enzymes which activate in turn chondrocytes to produce cytokines and matrix metalloproteinases (MMPs) which are some matrix-degrading enzymes. This results in further cartilage degradation and more inflammation in a positive feedback cycle [2] [28] [29]. Moreover, synoviocytes phagocytose the free particles which stimulates their proliferation, tissue hypertrophy and increased vascularity [2] [29].

The innate, i.e. non-specific, immune system has a key role in the inflammatory cycle of osteoarthritis. By recognising distinct molecular patterns such as damage-associated molecular patterns (DAMPs), it can take part in the host defence against microbial agents and modulate tissue homoeostasis. In osteoarthritis, DAMPs may be released due to the presence of free extracellular matrix fragments accumulated over time. They are recognised by Toll-like receptors of chondrocytes and activates the innate immune system reaction. This leads to the up regulation of matrix-degrading enzymes and inflammatory cytokines. These last ones are responsible for cartilage breakdown but also for decreasing the synthesis of key components such as aggrecans, part of proteoglycans, and collagen type II. Therefore, under pro-inflammary conditions, the balance between catabolic and anabolic events is disrupted and ultimately leads to the failure to compensate matrix cartilage damage [29] [30].

Moreover, the changes observed in the cartilage composition, i.e. decreases of proteoglycans and collagen content and increase of water content, make the matrix more permeable and reduce its compressive stiffness. Thus, it is more vulnerable to further mechanical damage [29].

In an attempt to respond to this degradative inflammatory cascade, chondrocytes are stimulated to boost their anabolic and proliferative activity. This stimulates the synthesis of extracellular matrix macromolecules leading to an increased deposition of collagen type II in the deepest cartilage zones. Moreover, during osteoarthritis, cartilage composition is modified with an increased expression of collagen type I which is a main component of fibrous cartilage. However, at a later stage of degeneration, these actions lose out to the increased degradative activity because of the malfunctioning and progressive loss of chondrocytes [29].

In addition to the cartilage degradation and synovial inflammation, alterations of the subchondral bone occur. Indeed, due to the close interactions between cartilage, bone and synovium, modifications in one of these tissues do not seem to occur independently from the others. However, it is unclear whether these changes take place before or after those in the cartilage [28] [29]. Endochondral ossification is reactivated leading to the formation of osteophytes, fibrous, cartilaginous and bony protrusions at the joint margins [1] [2]. The tidemark, which divides the superficial, uncalcified cartilage from the deeper, calcified cartilage, advances due to the progressive calcification of the cartilage. Vascular invasion takes places from the subchondral bone into the non-calcified cartilage which enables molecular interactions between bone and cartilage leading to further degradation of the last one [1] [2] [29]. The process of bone remodelling is increased and in later stages of osteoarthritis, bone volume and trabecular thickness increases leading to stiffer bone. Therefore, it is less able to absorb and dissipate impact loads, which increases stresses in the cartilage [28]. In addition, extensive bone sclerosis is the consequence of the severe remodelling process which takes places in areas of advanced cartilage degradation [28] [29]. When the synovial fluid gets access to the bone marrow in areas of total cartilage destruction, subchondral bone cysts, intraosseous lesions composed of fibroconnective tissue which initially contains fluid but ossify with time, appear [28]. Osteophytes and subchondral bone cysts can restrict joint motion and contribute to the associated pain [29]. Finally, bone marrow lesions can also be identified on magnetic resonance imaging (MRI) at the areas of most severe cartilage damage. They result from microfractures caused by the altered mechanical loading consequential from the changes in subchondral remodelling [2] [28]. Figure 1.10 represents a scheme of the structural changes and signalling pathways in the development of osteoarthritis by comparing it to healthy cartilage [2].

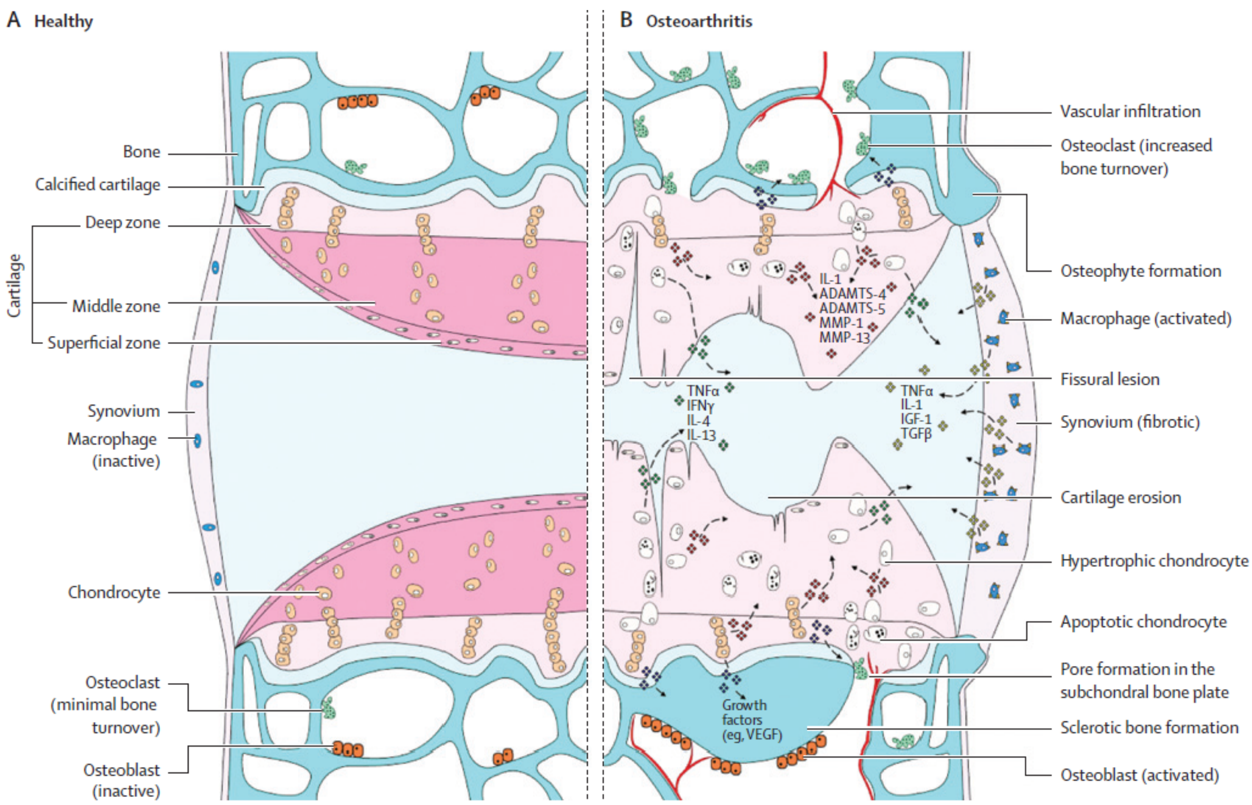

Figure 1.10: Scheme of the signalling pathways and structural changes in the development of osteoarthritis - ADAMTS: a disintegrin and metalloproteinase with thrombospondin-like motifs, IL: interleukin, MMP=matrix metalloproteinase, TNF: tumour necrosis factor, IFN: interferon, IGF: insulin-like growth factor, TGF: transforming growth factor, VEGF: vascular endothelial growth factor (adapted from [2]).

1.2.3 Symptoms and Diagnosis

Osteoarthritis can be diagnosed clinically based on the symptoms and a brief physical examination. The most common and disabling symptom, which leads patients to seek medical attention, is the pain. It is usually intermittent, depending on the activity, and mechanical (weight-bearing) and, generally, it increases with the progression of the disease. Among the other symptoms, one can mention joint stiffness, joint swelling, joint instability, muscle weakness, cracking, joint enlargement and functional discomfort with the limitation of activities due to the reduction of the range of motion [1] [25] [26]. These symptoms can vary among people and also depend on the stage of the osteoarthritis. The clinical diagnosis being mainly based on the symptoms, the disease is already well advanced and probably irreversible when patients start worrying about them enough to see a doctor [2].

Changes on plain radiographs can help to confirm the diagnosis even if they are not always needed. A narrowing of the cartilage space, an increased density of the subchondral bone and the presence of osteophytes and subchondral bone sclerosis and cysts are included among those radiographic changes [2] [26]. Figure 1.11 compares a normal knee with another one suffering from advanced osteoarthritis where several features that have been mentioned previously can be identified [31].

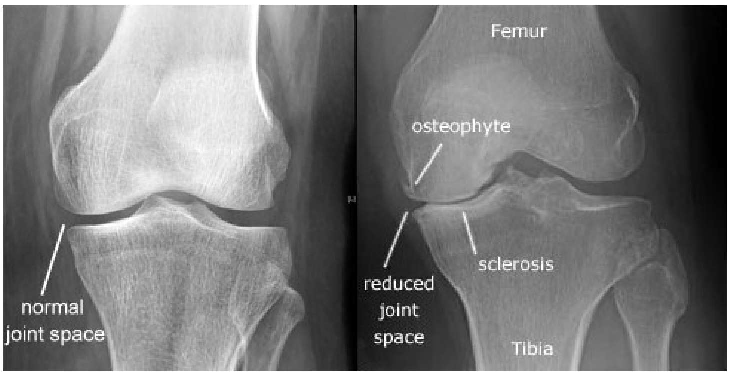

Figure 1.11: Radiograph of a normal knee and knee suffering from advanced osteoarthritis (adapted from [31]).

Other imaging techniques such as CT scans or MRI may be useful in the evaluation of the early stages of the disease but they are rarely necessary to establish the diagnosis [26].

1.2.4 Treatments

There are different ways to treat osteoarthritis depending on its state of progress. As a first-line treatment, several recommendations are made to the patients. First of all, they are required to lose weight if needed in order to prevent mechanical overload on an already weakened cartilage. Exercises can also be helpful to improve joint lubrication and nutrition, to maintain joint amplitude of motion, to relieve the stiffness and to increase muscles strength and aerobic capacity. A registration to a physiotherapy class is a good way to encourage to actually do this under supervision of a specialised staff. They should also reduce high-impact sports activities [1] [2] [25]. As pain medication, paracetamol and nonsteroidal anti-inflammatory drugs are the most often recommended [1] [2].

If all the appropriate conservative options have been unsuccessful, surgery should be considered. The three main surgical options are: osteotomy, unicompartmental knee prosthesis and total knee prosthesis [1] [25]. Other alternatives exist but they are not common so they will not be explored in this work. The choice between these surgeries is shared between the surgeon and the patients based on their analysis, their background, their functional complaints, their motivation as well as their clinical and radiological examination. Age, weight, autonomy, professional activity and leisure, general state and medical background, surgical background on the lower limbs, stage of osteoarthritis, joint amplitude,... are part of the anatomical and clinical factors guiding the therapeutic choice [25].

Osteotomy consists in correcting an axis defect and transferring the load from the diseased compartment to the healthy one. The mechanical axis of the lower limb is moved towards a healthy area which can tolerate an increased load so that the osteoarthritic area benefits from relaxation and can reshape itself [25]. This operation seeks to relieve the pain and improve the joint function [1]. In addition to being restricted for unicompartmental knee osteoarthritis only, it is also mainly considered for young (between 40 and 60 years old) and active subjects without end-stage joint disease [1] [25] [29]. This surgery enables them to postpone the prosthetic insertion for up to some ten years in a great majority of patients. This allows them to return to a sustained sporting and/or professional activity [25]. Two methods exist to realign the knee joint: the closing wedge osteotomy and the opening wedge osteotomy. The first one consists in removing a wedge of bone and pinning the open edges together using a plate. By contrast, in the opening wedge osteotomy, the surgeon opens a wedge and adds a bone graft to hold it open [25] [32]. Figure 1.12 shows these two approaches as well as the realignment observed following the closing edge osteotomy [25].

Figure 1.12: Representation of the closing wedge osteotomy (left) and opening wedge osteotomy (right) and realignment observed following the closing wedge osteotomy (centre) (adapted from [32]).

A surgical alternative of osteotomy is the **unicompartmental knee replacement** since it is also a therapeutic option when the disease is limited to only one compartment (see Fig. 1.13 (right)). This operation is more suitable for older patients because of the shorter rehabilitation time and faster recovery period. It is subject to a lower blood loss and a lower risk of infection [33]. Compared to the total knee prosthesis, it is much less tiring with less morbidity leading to a better functional outcomes for patients but this advantage is offset by a higher rate of revision [3] [25]. In general, patients undergoing joint replacement surgery (total or partial) have end-stage knee osteoarthritis and persistent severe pain [3].

Total knee arthroplasty is the gold standard operation for end-stage osteoarthritis when it concerns more than one compartment (see Fig. 1.13 (left)). As the partial knee replacement, the goal is to relieve pain and restore knee stability and knee function [1] [25]. Joint replacement surgery will be discussed in more details in the following section.

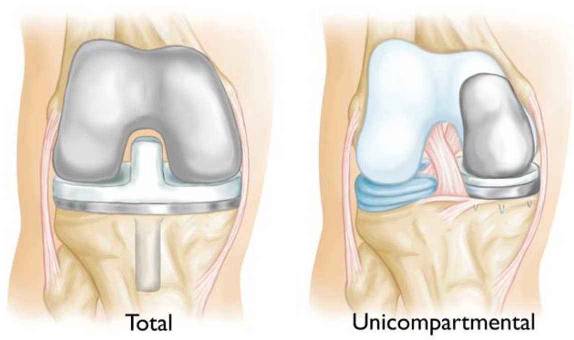

Figure 1.13: Representation of a total knee prosthesis (left) and unicompartmental knee prosthesis (right) (adapted from [34]).

1.3 Joint Replacement Surgery

In this section, joint replacement surgery is examined in more detail. Starting from statistics concerning this medical procedure, description is provided of the most common types of prostheses available. Then, the surgical procedure for knee replacement is detailed. In order to be as general as possible, the text focuses on a typical total knee arthroplasty, the principle being similar for a unicompartmental one.

1.3.1 Statistics

According to the New Zealand Joint Registry, the number of new operations in 2017 and 2018 is of 8,298 and 8,392 respectively, which represents 11.3% of increase [23]. Compared to Belgium, this rise is slightly higher with 12.77% going from 22,981 to 25,915 new surgeries during the same years [24]. As mentioned in the epidemiology, women are more concerned than men with and this difference is highlighted in Belgium with 61.3% of women affected against 51.64% in New Zealand. The mean age of people undergoing this surgery is relatively similar in both countries: 67.6 ± 10.2 in Belgium and 68.2 ± 9.49 in New Zealand [23] [24].

Osteoarthritis is the main cause of knee replacement, reaching almost 95% of cases in both countries. To a lesser extent, one can cite rheumatoid arthritis, avascular necrosis, post fractures, inflammatory arthropathy,... [23] [24]

1.3.2 Types of prosthesis

Prostheses used for total knee replacement are composed of four distinct parts: a femoral component, a tibial component, a tibial insert and a patellar component (see Fig. 1.14). The first two are usually made of cobalt-chromium or titanium alloys whereas ultra-high molecular weight polyethylene is used for the last two. Indeed, the criteria to choose these materials are that it must be biocompatible, able to fulfil the function of the knee to be replaced and keep its strength and shape during the prosthesis lifetime [35] [36] [37]. The femoral component comprises an anterior concave side to accommodate for the patella, and its most distal part is curved just like the real knee anatomy. The tibial component is a flat metal platform stabilised by a short stem inserted into the tibial medullar canal. It is covered by the tibial plastic insert which has a superior surface congruent with the outer surface of the femoral component. Finally, the patellar component, which is not always replaced, is just a polyethylene dome that replicates the shape of the patella [35] [37].

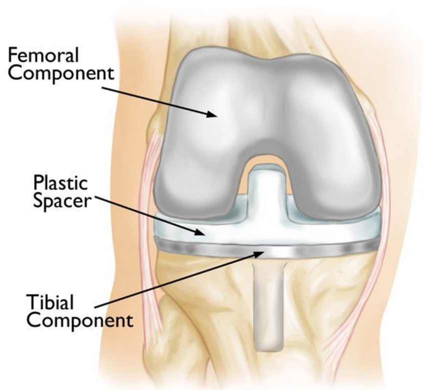

Figure 1.14: Representation of the front view of a total knee prosthesis and its different components (adapted from [35]).

Total knee prostheses can be classified by mechanical stability into three categories: unconstrained, semi-constrained and constrained. Usually, in total knee replacement surgery, the anterior cruciate ligament is removed whereas the posterior cruciate ligament can be either retained or removed depending on its condition. This difference is what distinguishes unconstrained and semi-constrained prostheses [35] [36] [37] [38].

Cruciate-retaining prosthesis is used when the posterior cruciate ligament is healthy enough to keep stabilising the knee joint. It is part of the unconstrained category because the femoral and tibial components are independent of each other. Figure 1.15 represents this design (left) [35] [36].

The posterior-stabilised design requires the removal of both cruciate ligaments. It is composed of a tibial component with a central spine fitting into the transverse cam of the femoral component which classifies it into the semi-constrained prostheses. This mechanism enables to replace the function of the posterior cruciate ligament and to replicates the rolling and sliding movement of the end of the femur over the proximal tibia. It is represented in Figure 1.15 (right) [35] [36] [37] [38].

These two types of implants are the most common and, despite some advantages and drawbacks controversial in several studies for each category, present excellent long-term results [39] [40]. Choosing one over the other is up to the surgeon (in collaboration with the patient) based on his experience with the design and brand, the patient's needs, and the cost and performance of the prosthesis [35] [37]. In New Zealand, 73% of all knee prostheses are cruciate-retaining and posterior-stabilised implants account for 23% [23]. The tendency is reversed in Belgium with 59% of the posterior-stabilised prosthesis against 19% cruciate-retaining designs [24].

Figure 1.15: Front view of a cruciate-retaining prosthesis (left) and a posterior-stabilised prosthesis (right) (adapted from [41]).

Among the constrained prostheses, the nonhinged one uses a deeper femoral cam and a larger tibial post, as represented in Figure 1.16 (left). This enables to reach a better stability in general. The drawback of this design is the increased femoral bone resection needed to accommodate the larger component but also the increased risk of aseptic loosening due to the larger constraint [42] [43].

Finally, the constrained hinged prosthesis comprises femoral and tibial components linked with a rotating hinged mechanism as its name implies it. It allows to decrease the risk of aseptic loosening at the expense of the increased level of constraint. This type also presents a longer femoral and tibial stems (see Fig. 1.16 (right)). This type is restrained to more severe damage or used in revision surgery. A constrained implant is illustrated in Figure 1.16 [37] [42] [43].

Figure 1.16: Left: Example of constrained nonhinged prosthesis (adapted from [42]). Middle: Front view of a radiograph of a constrained hinged prosthesis (adapted from [38]). Right: Example of constrained hinged prosthesis (adapted from [42]).

Another choice to make when choosing a prosthesis concerns the tibial insert: it can either be mobile or fixed (both possible for the two most common designs defined previously). The mobile one enables a slight rotation over the tibial component during the flexion-extension motion, allowing for a more accurate replication of the knee kinematics. Moreover, the load being distributed over a larger area, there should be less polyethylene wear in theory. By contrast, the main drawback is that the implant is more subject to dislocation (spin-out) since it requires more support from soft tissues. A fixed insert is attached strongly to the metal of the tibial component [35] [36] [37] [42]. The proportion of mobile insert is around 40% in New Zealand, this design being more recommended for younger and more active patients [23] [36].

There exist also different types of fixation to attach the components of the prosthesis to the bone. Polymethylmethacrylate is usually used as bone cement and ensures a good fixation of the implant to the bone as shown in Figure 1.17. However, unlike the cemented fixation, the prosthesis can also be press-fit onto the bone, allowing new bones to grow on the surfaces. This fixation is said cementless. A hybrid fixation can also be considered when the tibial and patellar component are locked thanks to bone cement whereas the femoral component is only ‘press-fitted’ [35] [37]. Despite these three possibilities of implant fixation, the cemented one is chosen in most cases whether in Belgium or in New Zealand, probably due to the poorer outcomes of the other two [23] [24] [37].

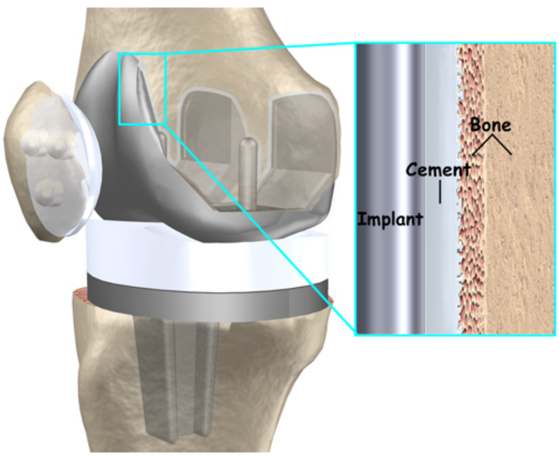

Figure 1.17: Representation of the cemented fixation for a knee prosthesis (adapted from [44]).

1.3.3 Surgical Procedure

In this subsection, the surgical procedure for knee replacement is summarised. First, the preoperative examination and planning are explained. Then, the main steps of the general surgical procedure are described. Finally, the rehabilitation process and the tools available to measure the outcome are exposed.

Preoperative Evaluation

Every surgery starts with a preoperative evaluation. This medical assessment is performed one to four weeks ahead of the surgery and consists of a complete patient's medical history and examination. Pain being the main indication of total knee arthroplasty, its source must be identified and all alternative conservative treatments must be exhausted. A history of any infections, allergies, previous treatments (surgical or not) is executed. Relevant comorbidities must be determined in order to reduce potential complications. Current medications, treatments or other conditions are also established. Furthermore, the doctor has to talk with the patient about his expectation after the surgery which is linked to his occupations, hobbies and activities [43] [45] [46].

This process is accompanied by a full examination of the patient. His gait is observed, any skin changes, swelling or deformities are identified. Crepitus, effusion and areas of joint tenderness can be determined by palpation. Finally, range of motion and ligament stability are also assessed [43] [45] [46].

As mentioned previously, the main indications for this surgery are osteoarthritis, rheumatoid arthritis, inflammatory arthritis,... By contrast, a chronic knee infection, severe vascular disease, presence of a functional and painless knee arthrodesis are some of the absolute contraindications [25] [43] [45] [46].

A full set of plain radiographs in many different views are required for both the diagnosis and the surgical planning. It enables the doctor to identify bone defects and deformities, to evaluate the mechanical alignment, to plan the major bone cuts and to determine the correct size of the various components [45] [46] [43]. Figure 1.18 (left) represents an anteroposterior X-ray view of an affected knee joint in which the femoral and tibial cuts planned are highlighted [45].

Surgery

In the first place, the patient is anaesthetised and prepared for the surgery to come. He is shaved and prepped with antiseptics, before being draped and placed in supine position with his affected knee flexed at $\pm 90°$, as represented in Figure 1.18 (right). The leg is stabilised with some supports but the knee and the ipsilateral hip remain freely mobile [33] [47].

The incision must enable a clear exposure and access to each side of the joint for the surgeon. It consists of an anterior midline cut of about 20 to 30 centimetres, the length varying depending on the patient's anatomy and the surgeon's experience. Starting, a few centimetres above the superior side of the patella, it goes down to just below the tibial tubercle [33] [47].

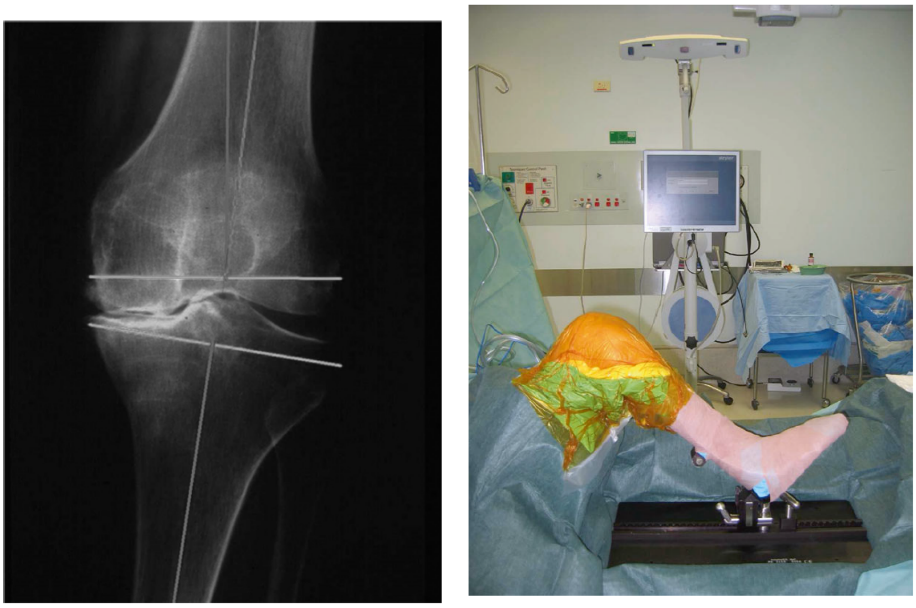

Figure 1.18: Left: Anteroposterior X-ray view of an affected knee joint undergoing total knee replacement in which femoral and tibial cuts planned are drawn (adapted from [45]). Right: Patient positioning in the operating room for a knee replacement surgery (adapted from [47]).

There are different surgical approaches for the knee arthrotomy. The *medial parapatellar* technique is the most commonly used because it provides an excellent exposure for proper component placement and alignment. However, the extensor mechanism is impaired due to the incision of the quadriceps tendon. Two possible alternatives are the *midvastus* and *subvastus* approaches which spare the quadriceps mechanism and therefore may lead to a faster recovery. The first one dissects the *vastus medialis obliquus* muscle belly but the insertion of this muscle on the quadriceps tendon is spared. On the other one, the muscle belly of *vastus medialis* is lifted off the intermuscular septum. The drawback of these alternatives is the limited exposure they provide, especially for obese patients or when this is not the first knee surgery. These three approaches are illustrated in Figure 1.19 [33] [43] [47].

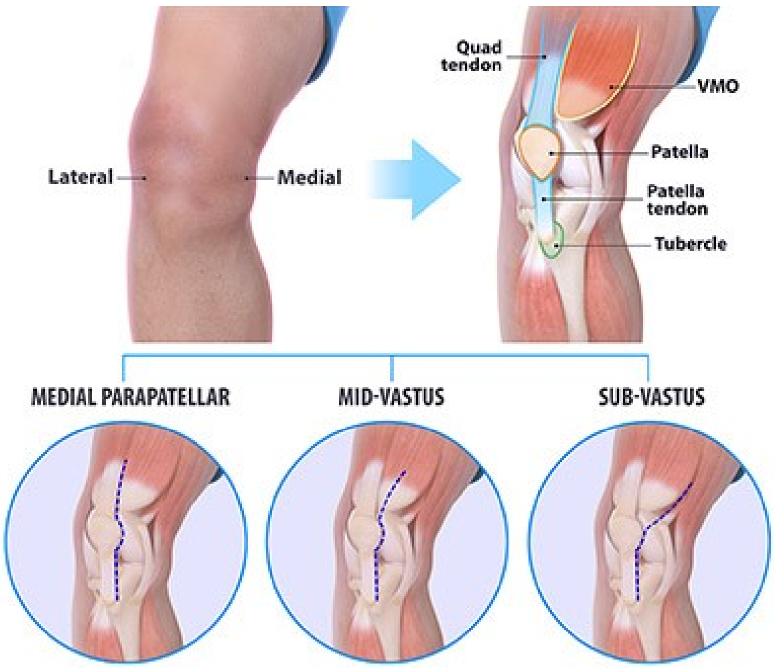

Figure 1.19: Illustration of the three main approaches for a knee replacement surgery (adapted from [48]).

Once the joint is fully open, the patella is everted [43]. The areas of damaged cartilage are removed and osteophytes are excised on the distal femur and proximal tibia [47]. Then, the different bone cuts can start. Their order can vary depending on the surgeon preference [43]. With the help of specific instruments (adjustable resection instruments, intramedullary guide, intramedullary drill, saw blades...), bones are resected and resurfaced [43] [49]. The distal femoral cut is usually based on an intramedullary referencing system guided by the medullary canal. The proximal tibial resection, as for it, is typically performed perpendicular to the tibial axis. The anterior and, depending on the prosthesis type chosen, posterior cruciate ligament attachments are also excised from the two bones [43] [49] [50].

Before going any further, the extension gap is assessed by inserting a spacer block with the knee fully extended. If it fits, i.e. the cuts are parallel and form a rectangle, it means that the extension gap is balanced. If it is not the case, adjustments must be made [43] [47] [51]. Next, the femoral component size is determined thanks to the assessment of the flexion gap. As this last one should match the extension gap to achieve the balance (see Fig. 1.20 (left)), the same spacer is inserted and thanks to anterioposterior sizing guide, the next femoral resections can be visualised. In total, four more cuts are made on the femur to accommodate for the femoral component depending on the size that has been determined: anterior, posterior, anterior chamfer and posterior chamfer cuts [43] [47] [51]. All the cuts needed for knee replacement surgery are illustrated in Figure 1.20 (right) [51].

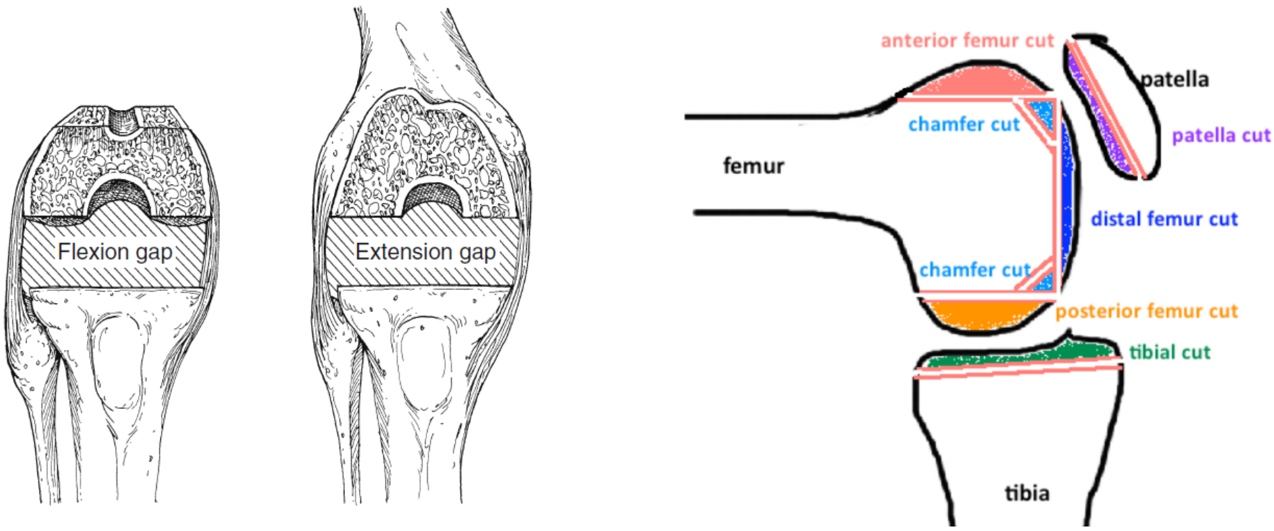

Figure 1.20: Left: Representation of the flexion and extension gap that must be equal (adapted from [52]). Right: Illustration of the bone cuts made for a knee replacement surgery (adapted from [51]).

The size of the tibial component is determined thanks to specific sizing tools again by drilling the centre of the tibia. If necessary, the posterior side of the patella can be resurfaced with a saw blade and based on another specific cutting guide. This resection dimension is based on a sizing chart [49] [50].

Finally, the implant components are attached to their corresponding bone with bone cement in most cases. The knee is then extended and flexed to ensure the correct fitting, alignment, sizing, positioning and stability of the components. Before closing the incision with stitches, the surgeon repairs any deep tissue that had to be cut during surgery [43] [50].

1.3.4 Other Approaches

Minimally invasive surgery is a more recent alternative of the typical procedure with the objective of reducing the rehabilitation time and postoperative pain. The incision made by the surgeon is much smaller. It also involves as little soft tissue dissection as possible with the help of specific instruments that can be manoeuvred around them. However, the resulting reduced visibility is probably the main drawback of this technique. Moreover, its clinical advantage over the traditional approach has not yet been established. Indeed, published studies have presented contradictory results about its use, notably in terms of complications [37] [47].

Another rising procedure is the *computer-assisted surgery*, used for surgical planning and for guiding during surgical interventions. It improves the accuracy of the bone resections and knee balance and therefore leads to the prosthetic fitting and alignment. This technique also allows a smaller incision which contributes to the smaller blood loss during the surgery and therefore a faster recovery and a reduced postoperative pain. However, such systems are expensive both in terms of purchase and maintenance. Furthermore, there is again no strong evidence for their long-term benefits for the patient and implant survival over the conventional arthroplasty [37] [53].

In a similar way, the widespread use of robots in orthopaedic surgery is relatively recent and is expected to further expand over the next years. There are currently two types of systems: semi-active and active one. The first one comprises a shared control mode, i.e. the surgeon is responsible for executing the operations but the robot prevents any deviations from the virtual operative plan. *MAKO* (MAKO Surgical Corporation, Fort Lauderdale, FL, USA), one of the most common systems, is classified into this category. By contrast, active systems, such as *ROBODOC* (Integrated Surgical Systems Davis, Sacramento, CA), are autonomous for steps or even the whole procedure based on the supervisor's orders [53] [54] [55]. Despite their excessive cost and the lack of long-term studies to prove the clinical advantage of such systems, they enable an accurate resection, fit, filling and alignment of the prosthesis [54] [55]. The Auckland City Hospital has recently acquired a semiautonomous robot: *ROSA Knee Robot*® (Zimmer Biomet, Warsaw, IN), represented in Figure 1.21. It provides real-time intraoperative data and their analysis derived from integration of the three-dimensional model, bone surface mapping and landmark registration and soft tissue laxity measurements. Rosa supports surgeon in performing total knee replacement by haptically positioning the resection guide and assessing the balance of soft tissue, which improves the accuracy of bone cuts and implant positioning [55].

Figure 1.21: Rosa Knee Robot® system (adapted from [56]).

1.3.5 Rehabilitation

After the surgery, the patient is led to his hospital room where he typically stays from one to four days before being discharged [43]. This length depends on the patient's recovery, pain and his ability to take care of himself by performing simple tasks such as eating, drinking, using the bathroom, getting in and out of bed,. . . Medical staff informs him about possible complications and warning signs of infections. A family member, friend or caregiver provides him support at home for the few weeks following the hospital discharge. Exercises are prescribed by their physiotherapist in order to improve their knee mobility, range of motion and muscle strength. They can either be done at home or during physiotherapy sessions scheduled at the clinic in some cases. Medical appointment with their surgeon is typically planned after two, six and twelve weeks post-surgery to assess postoperative knee function. Improvement in the knee mobility should be noticeable over the weeks, and less pain medication should be required. The patient is able to resume most of his activities within the first three months and full recovery is expected to be reached after six months to one year [57] [58].

1.3.6 Outcomes

Knee replacement is a successful, durable and cost-effective procedure. One possible way to evaluate its outcome is by performing a survival analysis[3]. In 2018, 635 more knee revision procedures have been registered in New Zealand and 2,246 in Belgium [23] [24], which accounts for approximately 8% of all knee surgery. Revision is defined as ‘a new operation in a previously replaced knee joint, during which one or more of the components is exchanged, removed, manipulated or added’ (New Zealand Orthopaedic Association, 2019) [23]. This amount of revision is increasing mainly due to the rise of knee arthroplasties, but they do not present a similar outcome since the risk of undergoing a second revision is five to six times higher than for the first revision. Aseptic loosening is the most common reason for revision accounting for almost 30% in Belgium. It is the consequence of an inflammatory response that induces bone loss and therefore components loosening. The next most frequent indications for revision are instability, infection and pain, each accounting for 20% of all revisions. Among the other less frequent causes, one can mention periprosthetic fractures, wear of the polyethylene component(s), stiffness and malalignment. Figure 1.22 represents all these reasons and the percentage associated with them in Belgium for the year 2018 and Figure 1.23 shows the Kaplan-Meier curve for age at primary knee replacement, i.e. the survival analysis by age group [24]. The goal of revision procedure is similar to the one of the primary surgery with the added complexity of its cause [3] [23] [24] [59].

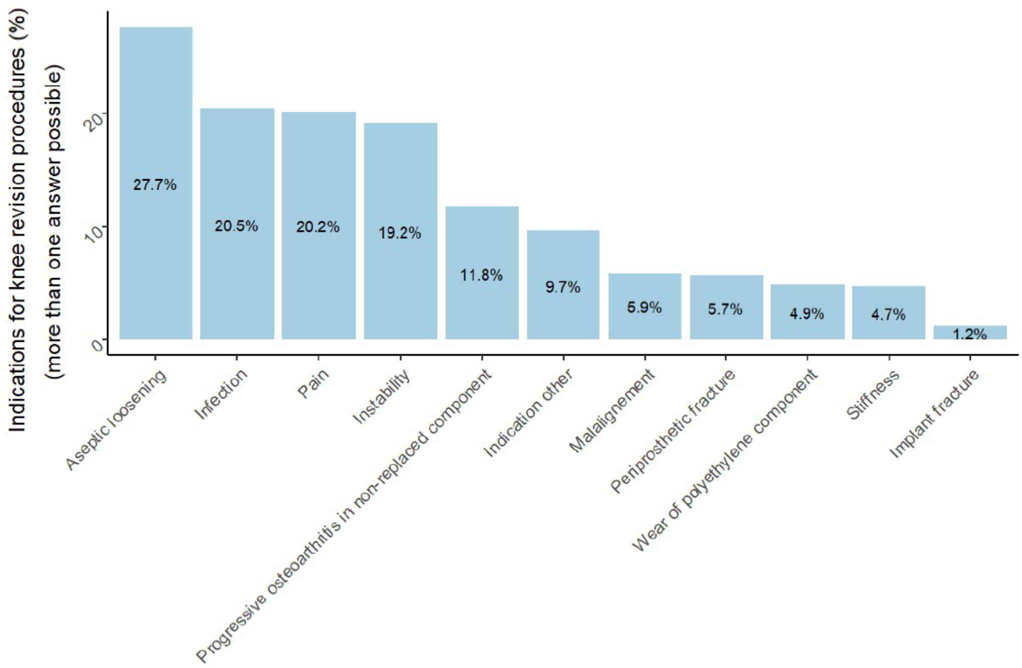

Figure 1.22: Indications for knee revision procedures in Belgium in 2018 (adapted from [24]).

[3] A set of methods for analysing data where the outcome variable is the time until the occurrence of an event of interest, i.e. the failure of the implant in this case

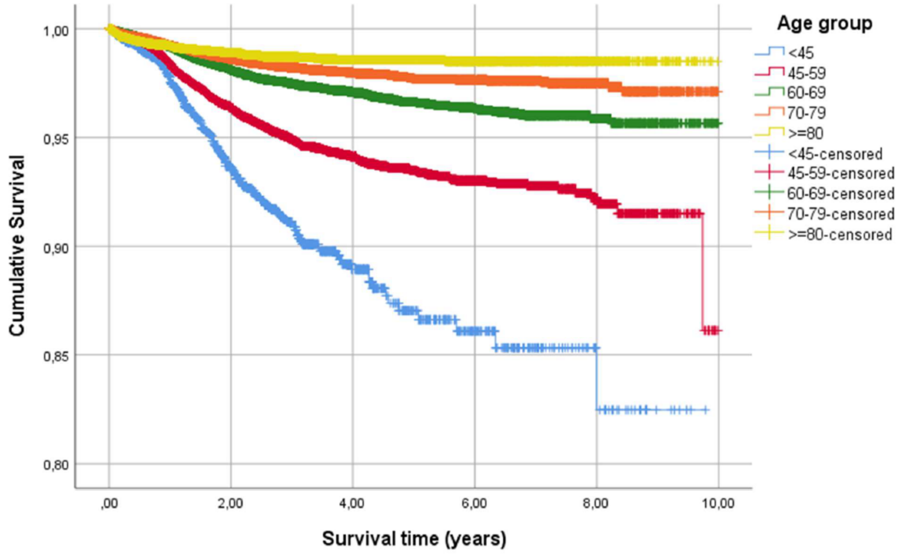

Figure 1.23: Kaplan-Meier curve for age at primary knee replacement in Belgium in 2018 (adapted from [24]).

Despite these excellent results, the number of patients dissatisfied with the outcome is higher than the number requiring revision surgery. Indeed, between 10 and 20% of patients continue to experience postoperative pain and functional limitations [3] [4]. The patients view on their surgery and their relative postoperative outcome is an important factor to take into account to quantify the success. This is the reason for the wide use of **Patient-reported outcome measures (PROMs)**. These tools reflect the patient's perception of function, symptoms and activity. PROMs can be divided into two categories: generic outcome measures and specific outcome measures. The purpose of generic health outcome measures is to assess the patient's overall quality of life without restriction to a peculiar disorder [5] [6]. The Short-form 36 (SF-36) is the most common example of generic tool and consists in a 36-item questionnaire measuring general health and quality of life. Another example of this kind of measure is the EQ-5D developed by the EuroQol group. It assesses the general health status based on five domains: mobility, self-care, usual activities, pain/discomfort, anxiety/depression. The use of generic tools allows comparisons to be performed across conditions and treatments. However, they might not reflect clinically important changes [5] [6]. By contrast, specific measures depends on a particular pathology, condition or anatomic location. Therefore, they are more responsive to changes in the phenomenon of interest and patients find them more relevant [5] [6]. For example, the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) score is a frequently used health questionnaire designed for patients with osteoarthritis of the knee or hip. It assesses three domains: pain, stiffness and physical function. The Oxford-12 Knee Score (OKS) is another widely employed questionnaire dedicated for patients undergoing knee arthroplasty evaluating knee pain and function. Finally, the last example is the Forgotten Joint Score (FJS), again designed for patients with artificial joints. It assesses their ability to forget it in everyday life [3] [5] [6]. To date, it is not possible to recommend one single best test although there is agreement that these outcome tools have been proven to be valid, reliable and reproducible [5] [6].

Nevertheless, commonly used PROMs have been shown to over-report the level of actual physical function and a ceiling effect[4] has frequently been observed. Therefore, the test score might fail to discriminate patient outcomes, i.e. a patient can be dissatisfied with the outcome despite presenting a good score [4] [8] [9]. Moreover, it is only a subjective measure based on the patient's perception. This is why the need for a more objective assessment is required in combination with PROMs. **Performance-based outcome measures (PBOMs)** assess objectively the ability of a patient to perform directly observed tasks. Associated to PROMs, they give a clearer view of the functional outcome [5]. Among the exercises used for the functional evaluation of patients undergoing knee replacement, one can mention the

[4]The ceiling effect is said to occur when participants' scores cluster toward the high end (or best possible score) of the measure/instrument. The opposite is the floor effect. [60]

6-min walk test, stair climbing tests, chair to stand test and self-paced walk test. For example, the 6 min walk test measures the distance covered by a patient on a flat surface over six minutes [3] [5] [4]. Despite addressing the current limitation of PROMs, PBOMs require an assessor, equipment and an area dedicated to perform the tasks. They are time-consuming and incorporating them into a routine clinical follow-up might be difficult [4] [8].

Direct monitoring of performance is another objective alternative or additional test to evaluate the recovery. For example, range of motion can be easily evaluated with a goniometer and is considered as a reliable outcome measures [3] [4] [5]. Devices such as accelerometers can be used to quantify the patient's physical activity [4] [8]. However, it involves once again some specific equipment and medical staff to evaluate these measures and it can only be performed during doctor follow-up, which may not be frequent enough.

Several studies have reported some discrepancies between objective and subjective measures. Therefore, they recommend using a combination of both instead of only relying on self-reported measures [4] [8] [61] [62].

1.4 Inertial Measurement Units

As explained previously, one of the ways of assessing patient outcome after joint replacement objectively is to monitor its activity directly. Nonetheless, the low frequency of clinician follow-up (typically at 2, 6 and 12-week post-surgery) and the scarcity of physiotherapy within a group setting after this surgery, limit the opportunities to assess postoperative function [7]. The use of wearable sensors, such as inertial measurement units (IMUs) can overcome this issue by providing a method to monitor patients' physical function objectively and remotely. This chapter presents the basic functioning of IMU components and their application in the biomedical field.

1.4.1 Principle

An IMU is an electronic device that integrates multiaxes combinations of accelerometers, gyroscopes and magnetometers to measure linear acceleration and angular velocity as well as magnetic fields surrounding the device. It uses microcontrollers to process its collected measurements and Bluetooth modules for system communication [63] [64]. A brief description of each inertial sensor is provided below.

Accelerometers

A single axis accelerometer is a type of inertial motion sensors measuring linear acceleration along a peculiar axis. Its principle can be explained with a simple system composed of a mass, suspended by a spring in a housing as represented in Figure 1.24. According to Hooke's law, a spring (within its linear region) will exhibit a restoring force F proportional to the distance x it has been expanded or compressed, or in mathematical terms: $F = kx$ where k is the spring constant. Moreover, as stated by Newton's second law of motion, the acceleration a of the mass is directly proportional to the force F applied on this mass: $F = ma$. This force causes the mass to either compress or expand the spring as: $F = ma = kx$. Therefore, the acceleration can be determined by measuring the induced displacement of the mass connected to the spring $a = \frac{kx}{m}$. By duplicating this system along other axes enables to provide multiple axes of acceleration and thus, multiple degrees of freedom of translation. The three most common types of accelerometers are piezoelectric, piezoresistive, and capacitive. They use capacitive sensing and the piezoelectric effect to sense the displacement of the proof mass which is proportional to the applied acceleration [63] [64] [65] [66].

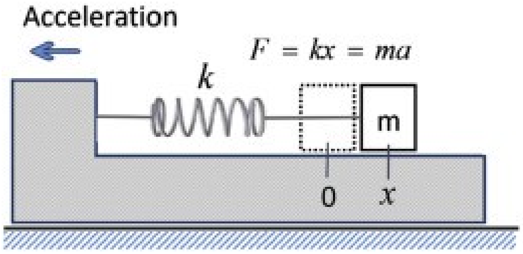

Figure 1.24: Mass-spring system used for measuring acceleration (adapted from [65]).

Gyroscopes

Gyroscopes measure angular velocity about three axes: roll (x-axis), pitch (y-axis) and yaw (z-axis) as represented in Figure 1.25 (left) [64]. The type used in human motion analysis are vibrating mass gyroscope. They are based on the concept of using the Coriolis force, a fictional force acting on a moving object when observed on a rotation reference frame, to induce a response to a movable mass held by two orthogonal sets of springs (see Fig. 1.25 (right)). This mass is forced to vibrate at its resonant frequency along the direction of one set of springs (e.g. x-axis). Therefore, the vibration provides the structure with a linear velocity component. When the structure is rotated (e.g. about the z-direction), the Coriolis force drives the mass in the direction perpendicular to the original vibrating direction, causing a second oscillation in this direction (e.g. y-axis). The magnitude of this secondary oscillation is proportional to the angular rate of rotation. The Coriolis force is given by: $\mathbf{F}_c = -2m(\omega \times \mathbf{v})$, where m is the mass, $\mathbf{v}$ the drive velocity of the mass relative to the moving system to which it is attached and ω the angular velocity of this system. Gyroscopes based on other operating principles also exists but are not as used for applications in human motion analysis (optical, mechanical,...) [63] [66] [67].

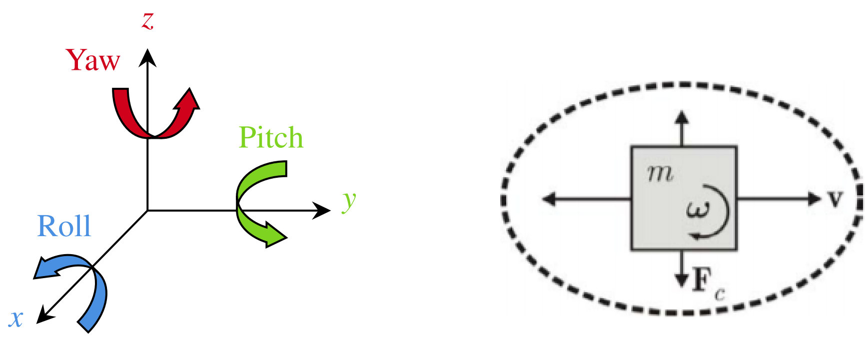

Figure 1.25: Left: Illustration of the roll, pitch and yaw angles. Right: Representation of the structure composed of a movable mass held by two orthogonal sets of springs (adapted from [66]).

Sensor fusion enables to combine accelerometers and gyroscopes to provide the complete position and orientation of an object in a three-dimensional space. Usually, 3-axis accelerometers and 3-axis gyroscopes are used to provide the six degrees of freedom required for this purpose: three degrees of translation movement along each axis (front/back, up/down and right/left) and three degrees of rotation about each axis (pitch, roll and yaw) [63] [64].

Magnetometers

A magnetometer measures the magnetic fields and can therefore detect fluctuations in Earth's magnetic field. They are usually classified into two basic types: scalar and vector magnetometers. The first one measures the total strength of the magnetic field to which it is subjected, but not its direction, whereas the second one can measure a component of the magnetic field in a particular direction. Again in each category, there are various models. The most common used magnetometers relies on the Lorentz force acting on the current-carrying wire in the magnetic field. Indeed, in absence of magnetic field, electrons run parallel to the direction of the current. However, in presence of a magnetic field, they gather on one side of the conductive material. The mechanical motion induced can be sensed either electronically or optically to determine the direction of this magnetic field. Moreover, since the Lorentz force applied is proportional to the magnetic flux density ($\mathbf{F} = q\mathbf{E} + q\mathbf{v} \times \mathbf{B}$ where $\mathbf{F}$ is the Lorentz force, q the particle charge, $\mathbf{v}$ the particle velocity, $\mathbf{E}$ the electric field and $\mathbf{B}$ the magnetic field), its strength can also be determined. Therefore, these sensors can estimate changes in the orientation of a body segment in relation to the Earth's magnetic North or the vertical axis in gait analysis. Combined to accelerometers and gyroscopes, the motion, orientation and absolute heading can be determined [63] [64] [66] [68] .

1.4.2 Applications in Human Motion Analysis

During the recent years, this technology has seen increasing popularity in the field of human movement analysis. Indeed, it can estimate motion in space without the limitations encountered by traditional gait analysis tools such as force platforms, instrumented mats, and optoelectronic system. Despite being the gold standard to provide spatio-temporal parameters as well as accurate joint kinematics and kinetics, only a limited number of steps can be captured with these conventional instruments due to the restricted laboratory environment. Moreover, the recording process is labour-intensive and time-consuming and their prices are high. IMU, on the other hand, is a wearable, compact and low-cost solution to measure functional parameters similar to those captured by three-dimensional gait analysis. Their placement is easy and fast and there is no limitation by the captured volume. They can provide unsupervised, long-term recording outside peculiar environment, increasing the quantity of data collected and making possible to detect potential rare events such as falls. These advantages make them suitable to monitor activities of patients anywhere for clinical applications such as rehabilitation or clinical diagnosis [10] [11] [63]. For example, assessing ambulatory gait analysis is important for people with neurological conditions such as Parkinson's disease and stroke. Indeed, these conditions are characterised by motor dysfunctions including resting tremors, limb rigidity, slowing of movement. Therefore, gait evaluation can be used as a reliable diagnosis [63]. As mentioned previously, wearable sensors have also made their way into the field of joint arthroplasty, notably to assess patient's progress [63] [69] [70].

Many authors (Trojaniello et al., 2014, 2015 ; Esser et al., 2011 ; Storm et al., 2016 ; Change et al., 2016 ; Houdijk et al. 2008 ; Hundza et al., 2013 ; Boutaayamou et al., 2015) have suggested methods to estimate spatio-temporal parameters based on IMUs measurements [71] [72] [73] [74] [75] [76] [77] [78]. For example, Trojaniello et al. have proposed and validated an approach to determine both temporal and spatial parameters which could be applied to normal and to various pathological gait patterns. By attaching an IMU on each subject's ankle, gait events were detected thanks to the cyclic nature of gait and by exploiting lower limb invariant kinematics. The stride length was estimated using a combination of an IMU axes realignment along the direction of progression and of an optimally filtered direct and reverse integration. This method has been validated by comparing the results with the gold standard instrumented mat measurements [71].

Others (Weygers et al., 2020 ; Cooper et al., 2009 ; Tadano et al., 2013) have presented approaches to estimate joint kinematics from inertial measurement units data [79] [80] [81]. Theoretically, the position and orientation of an inertial sensor can be trivially determined by integrating the angular velocity or double integrating acceleration data. However, one of the drawbacks of IMU is the inherent bias. Therefore, when integrated, any small amount of noise will accumulate over time. Alternative methods have thus been developed. Picerno has reviewed the literature by classifying and describing the one used so far to estimate lower limb joint kinematics thanks to wearable inertial sensors [82]. Three main approaches exist to achieve this purpose: orientation-based methods, kinematics or musculoskeletal model-based approaches and data-driven methods [11]. The first one relies on the computation of the relative orientation of two adjacent segments based on the orientation of IMUs attached on the distal and proximal segments. However, among the limitations, one can mention the need of mounting IMUs on each segment for lower limb kinematics estimation which is bulky [81]. Moreover, errors relative to sensor-to-segment alignment require a calibration step to obtain the relative orientation of the IMU and the anatomical reference frame [79] [82]. Kinematic or musculoskeletal model-based methods take into account kinematics constraints of the human body to overcome the limitation of inertial sensors [83] [84]. Finally, the data-driven approaches depend on machine learning algorithms which compute joint angles based on the raw IMU data [85] [86] [87]. The main limitation encountered with this method is the dependency on the supplied dataset and the generalisation to other individuals [11].

1.5 Main aim of This Thesis

The previous sections have highlighted the need for a contemporary outcome assessment after knee replacement surgery which combines subjective PROMs, objective PBOMs, and objective direct measurement of the patient's physical activity. Using inertial measurement units allows clinicians and physiotherapists to remotely monitor their performance in both clinic and home settings and to replace conventional three-dimensional gait analysis. While IMUs have been applied extensively in younger orthopaedic patients following sports injury [88] [89] [90] [91], few studies have used them to evaluate elderly patients after knee arthroplasty [92] [93] [94]. The purpose of this study is to develop and validate a new workflow to quantitatively assess joint function following knee replacement surgery using only two ankle-worn IMUs. To this end, a machine learning algorithm has been implemented to predict knee flexion time series kinematics during walking gait. This approach provides a solution to obtain objective quantitative data over a long capture duration.

1.6 Covid-19

In view of the exceptional circumstances related to the Covid-19 pandemic, some difficulties have been encountered in the course of carrying out this work. New Zealand has been in lockdown (Alert Level 4) from the 25th of March to the 8th of June. During this period, all routine surgeries, including knee arthroplasty, have been postponed. Therefore, the number of patients initially planned throughout the period of my internship has been relatively reduced. This lack of data has had a subsequent impact for the training of the machine learning algorithm and therefore on the final results.

1.7 Work Performed

This master thesis is part of a larger research project at the University of Auckland (Auckland, New Zealand) for which several people from different backgrounds have collaborated. My personal contribution consisted in following the patients from their enrolment in the study until the end of their six weeks of rehabilitation process to collect the data. I was in charged of performing the preoperative and postoperative gait lab analyses as well as the tracking of the patients at their physiotherapy sessions. Given the reduced time I spent in the institute, I have only followed the last three patients (ID 10, 12, and 13). I have also processed the data obtained during the gait-based laboratory sessions, not just the ones I had personally performed but also the ones that had been recorded before my arrival, with the first patients. Finally, I have taken part in the evaluation of the machine learning algorithm that was developed by my colleague, notably by performing a sensitivity analysis on the parameters.

Chapter 2

Methodology

At this stage, twelve patients participated in the study (6 males and 6 females ; age: 68.17 ± 6.35 ; height: 1.67 ± 0.08 m ; body mass: 86 ± 20.28 kg ; body mass index: 30.65 ± 6.13 kg/m^2). Seven underwent a total knee arthroplasty and five a unicompartmental one for end-stage osteoarthritis. They all provided written informed consent prior to entering the study and local ethics committee approval was obtained from The Auckland Health Research Ethics Committee. Participants general information and information relative to their surgery are provided in Appendices A.1 and A.2 respectively.

The study design is outlined in Figure 2.1, consisting of a pre-screening and information session, preoperative gait analysis with IMU sensors, knee replacement surgery, a few weeks of IMUs tracking during physiotherapy and at home and a postoperative gait analysis after this period.

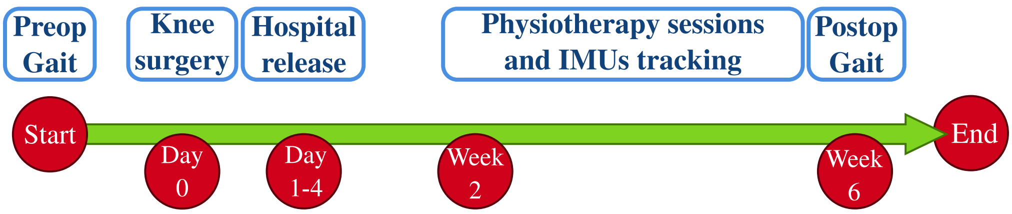

Figure 2.1: Timeline with routine measurement cycle using wearable IMUs following knee arthroplasty.

The purpose of this chapter is to present each step of the study methodology followed to collect the data. Then, a description of the machine learning pipeline is given.

2.1 Ethics approval

Local ethics committee approval was obtained from The Auckland Health Research Ethics Committee (AHREC). The study is planned to enlist two hundred consenting participants and it is taking place at Greenlane hospital in the Orthopaedic Department under the supervision of orthopaedic surgeons Mr Jacob Munro and A/Prof Paul Monk, as well as at North Shore Hospital in the Orthopaedic Department under the guidance of Mr Simon Young. The expiry date for it is the 4$^{\text{th}}$ of March 2022. A pilot study was conducted at Greenlane Hospital as part of a PhD project. In light of the results, the intention of expanding to further District Health Boards was made known. Ethical approval for this pilot study was obtained from the University of Auckland (UoA).

Eligible participants are adults undergoing primary knee arthroplasty at Auckland District Health Board (ADHB) or Waitematā District Health Board (WDHB). They are screened for gout and/or diabetes since these conditions can cause complications in their muscle structure and recovery period.

Moreover, if potential participants cannot attend routine follow-up appointments during their rehabilitation, they are also excluded. The recruitment process is based on flyers and posters regarding the research conducted as well as on clinicians' notifications to potentially eligible participants. A consent form has to be signed by the patients before anything and they have the right to withdraw from the study whenever. A copy of this consent form is provided in Appendix B.

Regarding the data confidentiality, de-identification of patients data is observed to ensure the strict secrecy. All information is entered into an encrypted database by the researcher and, in addition to images and other documents, they are saved onto a secure university hard drive for a maximum of six years. After this time, they will be permanently deleted. Only the investigators have access to the study files.

2.2 Data Collection

In this section, protocol applied to collect the data is explained. First, a review of the patient-reported outcome measures used in this study is provided. After that, the conduct of the gait-based laboratory sessions are described. Finally, the physiotherapy sessions and their following home monitoring are presented.

2.2.1 Patient-Reported Outcome Measures

Patients completed an online version of the Oxford Knee Score and EQ-5D with the EuroQol Analogue Scale at every stage of the study: during the pre-operative gait assessment, at each physiotherapy session, at the post-operative gait evaluation and throughout the clinical follow-up a few months later. The Forgotten Joint Score was obtained only at week 2 and 6 after surgery. These three questionnaires and their manner of scoring are presented in Appendices C, D and E respectively.

2.2.2 Gait-Based Laboratory

A gait analysis was performed before the planned knee replacement and after the approximate six weeks of recovery to enable the comparison pre- and post-surgery. The gait laboratory is located at the Department of Exercise Sciences, University of Auckland. A 10-camera motion capture system (Vicon MX Cameras, Oxford Metrics Group, Oxford, UK) combined with three force plates (Berec Force Plates, Bertec, Ohio, USA) was used to collect motion data (100 Hz) and ground reaction force data (2000 Hz) respectively. For this study, synchronous IMU data were captured at 1149 Hz from *VICON Blue Trident* (VICON, NZ) sensors using the application *Capture.U*. It is a wireless inertial measurement unit, which combines a 3-axis accelerometer, a 3-axis gyroscope and a 3-axis magnetometer. Its specifications are listed in Table 2.1. The protocol followed to record optical marker trajectories is explained below.

In the first place, the system, cameras and force plates had to be calibrated. Once the patients had arrived, the study design was explained to them as well as the conduct of the gait assessment and physiotherapy sessions they were going to attend. The signed consent form was returned to the supervisor. Patients also had to complete the online version of the questionnaire (OKS and EQ-5D).

General	
Dimensions	42 x 27 x 11 mm
Weight	9.5 g
IP68	
Bluetooth 5	
Battery	up to 12 hr life
Charge time	1.5 hr

Sensor	
Accelerometer sensor	Low-g 16 bit High-g 13 bit
Accelerometer axes	3 axis
Accelerometer range	Low-g ± 16g High-g ± 200g
Accelerometer frequency	Low-g 1125Hz High-g 1600Hz
Gyroscope sensor	16 bit
Gyroscope axes	3 axis
Gyroscope range	+/- 2000°/s
Gyroscope frequency	1125Hz
Magnetometer sensor	16 bit
Magnetometer axes	3 axis
Magnetometer range	+/- 4900µT
Magnetometer frequency	100Hz

Table 2.1: Specification of *Vicon Blue Trident* sensor [95].

For motion capture, participants were barefoot and were wearing tight shorts to facilitate the markers placement. They were fitted with a 'University of Western Australia (UWA)' marker set (see Fig. 2.2 (left)) [96]. It consists of the following anatomical markers: the most dorsal point on the acromioclavicular joint, the Processus Spinosus of the 7^{th} cervical vertebra, the deepest point of Incisura Jugularis (suprasternal notch), the most caudal point on lateral and medial humeral epicondyle, the most caudal–lateral and caudal–medial point on the radial styloid, the anterior superior iliac spine (ASIS), the superior aspect of the 5^{th} lumbar vertebra–sacral, the lateral and medial femoral epicondyles, the lateral and medial malleolus and the posterior aspect of the calcaneus [10] [97] [98]. Moreover, technical markers were placed on the lateral sides on the thighs and shanks (triad marker sets) and the head of the first and fifth metatarsal [10]. Four IMU sensors were used, two on each lower limb: the proximal sensor was placed on the distal anterior femur, approximately five centimetres cranial to the knee joint line and the distal sensor was placed two centimetres cranial to the ankle medial malleolus on the anteromedial aspect of the distal tibia, as represented in Figure 2.2 (right). IMUs on the ankle were attached thanks to a bracelet with a velcro strap in order to minimise any discomfort whereas the two on the thigh were just attached with tape.

The subject's age, gender, weight and anthropometric measurements were also collected. The leg length (ASIS to medial malleolus), femur length (greater trochanter to lateral condyle), tibial length (medial condyle to medial malleolus), thigh girth (mid-point femur), shank girth (upper-middle third tibia) and ASIS distance were measured for both sides.

During every session, participants performed five different tasks. First, a static trial was recorded where they were standing one foot on each force plates and raising their arms horizontally in the frontal plane. This static calibration was used to locate anatomical landmarks and estimate the position of the hip joints centres, knee joint centres and functional flexion/extension axes of each knee. Then, five walking trials at self-selected speeds with successful force-plate strikes were captured. In addition, five trials of squats, sit-to-stand and stair climb (with two steps only) exercises were performed by the subjects. Note that, in view of the unexpected lockdown, no pictures to illustrate this gait-based sessions are provided.

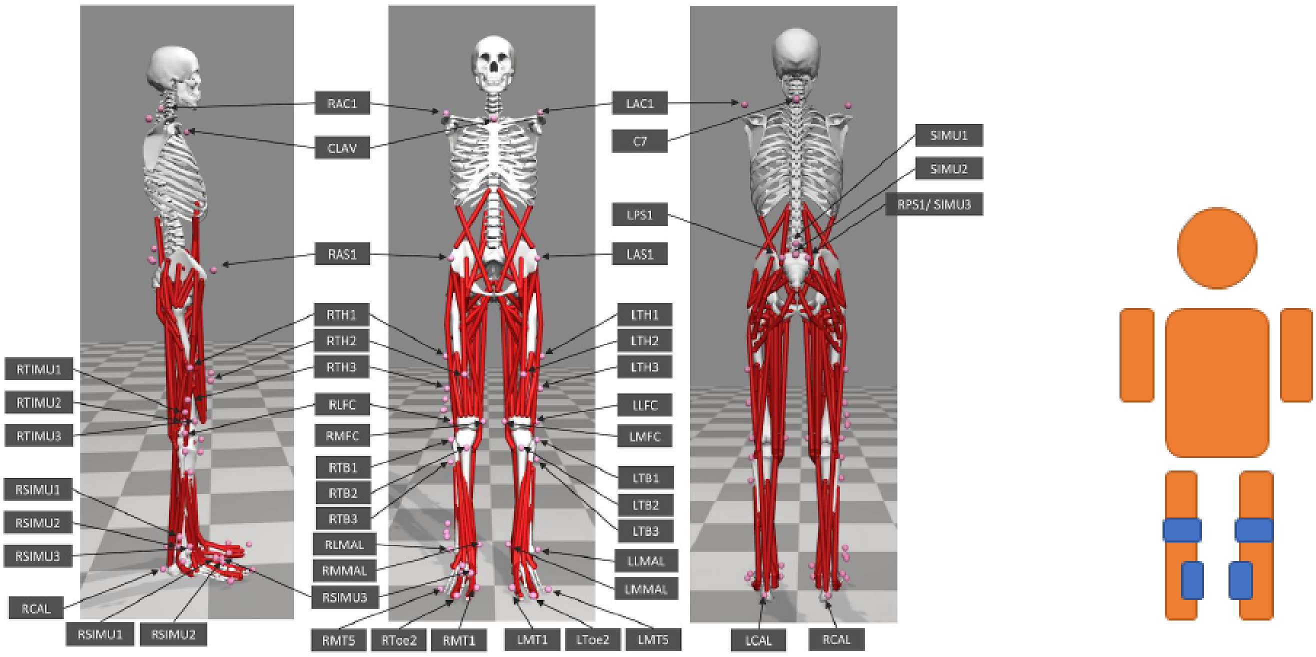

Figure 2.2: Left: Marker set with the *OpenSim gait2392* model that was used to track kinematics. Right: Placement and orientation of the inertial measurement units.

Since hardware synchronisation was not possible at the time of data collection, a light-emitting diode (LED) based method was used to synchronise IMU and motion capture data together. An IMU was modified to include an infrared LED visible to the optical motion capture cameras. This LED turned on 2.5 seconds after the IMU started recording. In the recorded motion capture data, this appeared as a marker in the capture volume. It enabled a synchronisation frame to be identified (with the accuracy of 10 msec at a capture rate of 100 Hz). The synchronisation protocol used for each trial was the following:

1. Motion capture cameras start recording;
2. IMU trial capture starts and the infrared LED on the IMU appears in the capture volume;
3. Participant performs the tasks;
4. IMU recording is stopped manually;
5. Motion capture recording is stopped manually.

Participants performed two other activities from the PBOMs category. First, the 6-min walk test to measure the maximal walking distance covered in six minutes. Patients were instructed to walk as far as possible in a safe manner in a fifty-meter loop track and the number of laps was recorded. Assistive devices were allowed if needed. This test has been proven to be reliable and is commonly used to evaluate functional performance in many different patient groups, including patients who have undergone knee replacement surgery [99]. Finally, a downhill treadmill test was performed on a *SportsArt Fitness T650ME* treadmill (SportsArt, Washington, USA). Patients were instructed to walk downhill with a slope of 7° decline (-12%). After a six-minute acclimatisation period, the speed is increased incrementally until they got uncomfortable [100]. This downhill top walking speed reached was documented as the fastest they could walk without running [101]. Every information collected during these appointments are gathered in Appendix A.3.

2.2.3 Physiotherapy Sessions

During the weeks following the surgery, patients attended an outpatient rehabilitation program led by physiotherapists twice a week. Physical therapy is an important part of the recovery process. The prescribed exercises restore strength and mobility to the operated knee and ensure a gradual return to everyday activities. One should notice that while most patients completed the common six weeks, some left early or took longer due to individual recovery timelines.

It is advantageous for the patients to go to these rehabilitation clinics because they are helped and encouraged by a specialised medical staff. This is a motivation to actually do their exercises. Moreover, they rub elbows with other patients who had undergone the same surgery, which can be a support for them. During these sessions, they had to attend to different activities:

- 50 metres at speed walk;
- Stairs climb (only from week 4 to week 6);
- Cycling for 3 minutes;
- Lunges: patients placed their operated leg onto the second step and leaned forwards bending their operated knee, reaching for the wall and holding the position for 10 seconds, 10 times;
- Steps ups and downs: patients stepped up onto the first step leading with their operated leg and then stepped back down leading with their non-operated leg, 20 times;
- Knee extensions: patients straightened their operated leg while keeping their thigh on the seat, held for 3 seconds and repeated the exercise 20 times;
- Sit to stands: patients stood up from the chair without using their hands 20 times;
- leg press: patients placed their operated leg onto the front plate, pushed back slowly (not fully straighten) and held for 3 seconds before returning to starting position with a weight of 30 pounds (for week 2) to 40 pounds (for week 3 to 6);
- Hamstring curl: patients was seated in the chair with the strap around the ankle of their operated leg and they slowly pulled backwards, held for 3 seconds and slowly released forwards 20 times with a weight of 2.5 kg (week 1) to 5 kg (week 2 to 6);
- Wobble board: patients stepped onto the wobble board, with their feet hip distance apart and remained in the position for 2 minutes (he could hold the bar for support when needed);
- Calf rises: patients faced the bar, rose onto tiptoes and held for 3 seconds before lowering down, 20 times;
- Kneeling: patients stepped their non-operated leg forward to help support themselves, knelt down onto their operated knee and hold for 10 seconds before standing up, 5 times (only from week 3 to 6);
- Stretches: a staff member helped patients to stretch the calf, hamstrings and quadriceps for a period of 20 seconds, 3 times each.

The person in charge of the study met with patients in at least one of their two weekly sessions. He placed two IMU sensors on each lower limb at the same locations as for the gait analysis (see Fig. 2.2 (right)) and started the recording with the application at the beginning of the sessions. The usual questionnaire is completed during the meeting. The scores as well as the range of motion measured during the session are gathered in Appendix A.3.

2.2.4 Home Monitoring

At the end of the physiotherapy session, the two IMUs over the distal thigh were removed, whereas the two devices over the ankle remained on. Patients were instructed to wear them for the rest of the day, which accounts for approximately six to ten hours, and to take them off prior to sleeping. They were asked to perform all routine normal activities, including washing and showering since the devices are watertight. These two sensors were returned to the responsible at the following physiotherapy clinic.

2.3 Algorithm Development

The purpose of the machine learning algorithm is to predict knee flexion during walking gait-based on data from only two ankle-worn IMUs. Two different walking models were generated: person-specific walking models, trained on a portion of a subject's data and predicting the remaining part, and a generalised knee model, trained on every individual of the cohort but one used for prediction. They will be explained in more detail in Chapter 3 Section 3.1. Predictions obtained from IMUs data were compared to knee flexion angle from optical motion capture recorded at the same time. However, during the development, it was determined that there may not be enough information in the patient dataset to build a robust generalised overground model. Indeed, a problem of IMU and optical motion capture synchronisation for the preoperative patient data has been encountered during the processing. To address this issue, the LED IMU mentioned previously has been created and introduced in the protocol for the gait-based laboratory sessions. Therefore, this model was trained on two other datasets already collected (at 200Hz) and processed from previous studies.

The machine learning algorithm is based on a feature engineering workflow. It automatically extracts time-series features based on the *FeatuRE Extraction based on Scalable Hypothesis tests* (tsFresh) to identify statistically significant features from the IMUs. The general idea behind this method is to characterise each time series by applying a library of curated algorithms to quantify them with respect to their distribution of values, correlation properties, stationarity, entropy, and nonlinear time-series analysis. This feature extraction, computationally expensive, is followed by a feature selection algorithm to prevent overfitting. Statistical significance of each feature is tested to predict the target and to control the false discovery rate. Finally, only statistically significant time-series features are selected [102] [103]. The steps of this algorithm are presented in Figure 2.3. Therefore, the machine learning pipeline comprises three components: data pre-processing, feature extraction, and feature selection and model training. One subject of the treadmill datasets (P08) was used to develop this pipeline and then the rest of the healthy cohort was used to create person-specific and generalised models. Finally, the algorithm was tested on the patient data.

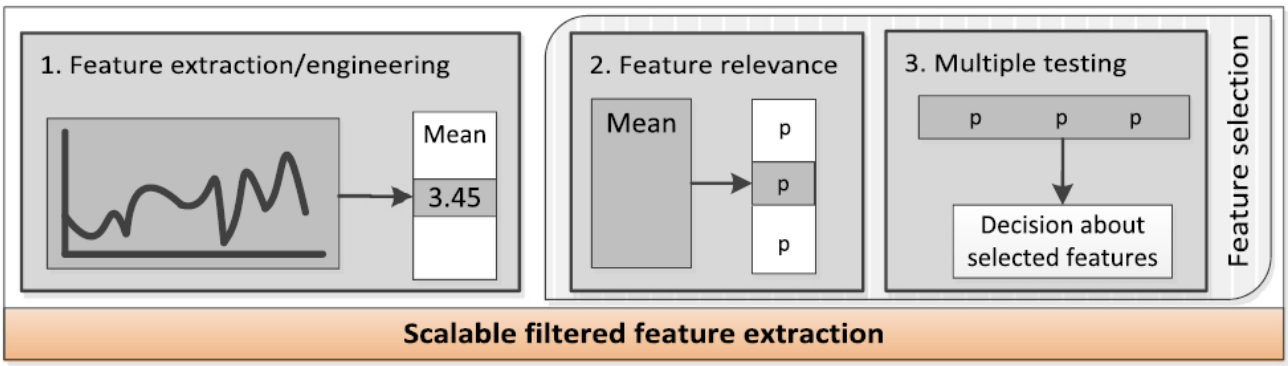

Figure 2.3: Steps of the `tsfresh` algorithm (adapted from [102]).

2.3.1 Data Pre-Processing

Following each participant's gait trial, markers were labelled using *VICON Nexus 2.9.3*, and inverse kinematics was performed using a scaled *OpenSim* model to track the motion of the patients using Optical Motion Capture (OMC). The last step of this pre-processing concerns the filtering and resampling of IMU and OMC data.

VICON Nexus

Starting from the trial files recorded during the gait-based laboratory sessions, *VICON Nexus* software was used to obtain the `.trc` files required by *OpenSim*. The first part consisted in creating the different segments composing the model: upper part of the body, thighs, shanks and feet (shown in orange in Fig. 2.4 (right)). This step was usually done in the static trial. Each segment is composed of several markers (at least three) represented by the white dots in Figure 2.4 (left). Sometimes, not all the markers were seen by the cameras during the whole trial so an adaptation was made with fewer markers. Once they have been created, adjacent segments were linked to each other.

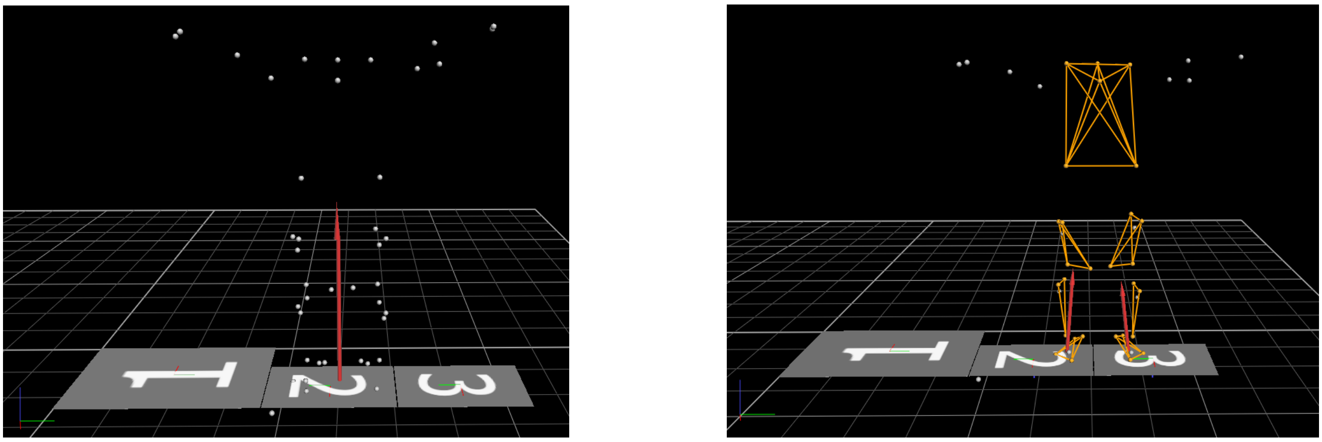

Figure 2.4: Left: Frame taken before markers labelling in the static trial of patient 13. Right: Frame taken after markers labelling in the static trial of patient 13.

Then, this specific marker labelling was applied for all the frames composing the different trials. When a frame was not recognised as similar enough to the one initially defined, labelling was performed manually by assigning each marker with the right name. The other common issue concerned markers swapping and was also corrected manually.

After that, the *gap-filling* tool was used to fill the gaps in the trajectory when certain markers were not seen by the cameras. Different types of filling were available but the most used were the *rigid body filling* based on the trajectory of the three other markers of the same segment and the *pattern filling* based on the trajectory of another marker of the same segment.

One should notice that the smoothness of each marker trajectory could be checked using the graph tool and representing the x, y and z components of each marker (see Fig. 2.5). Thus, any remaining marker swapping or bad gap filling could be detected and then corrected.

Finally, the corresponding `.trc` file was exported.

OpenSim

Once the `.trc` file had been obtained, *OpenSim* was used to scale the model and perform the inverse kinematics.

Scaling is performed by comparing experimental marker data to virtual markers placed on a unscaled model in the same anatomical locations as the experimental ones. The dimensions of each body segment in the model are scaled so that the distances between the virtual markers match the ones between the experimental markers [10] [104].

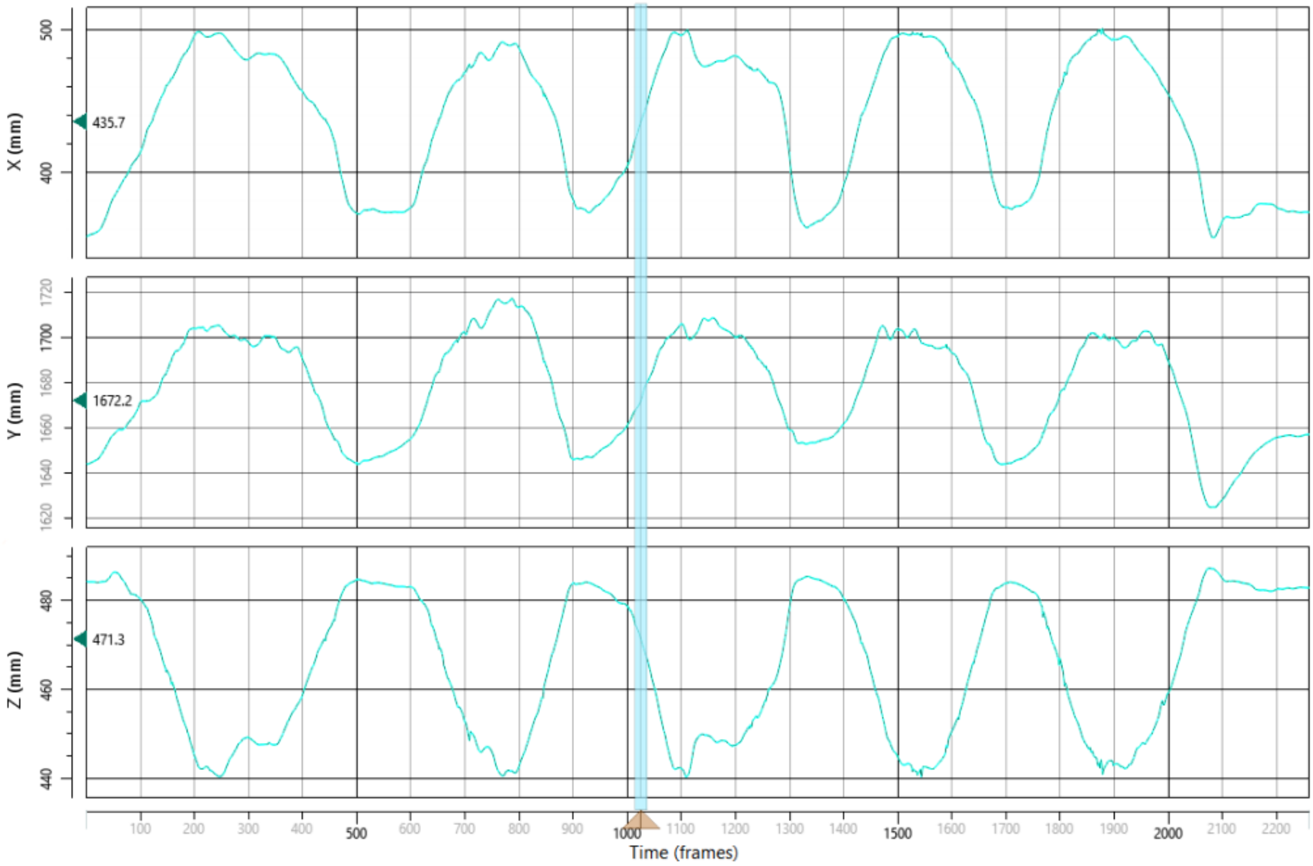

Figure 2.5: Example of graph obtained for the x, y and z component of the left lateral femoral epicondyle (*L.Knee.Lat)* for the squat trial of patient 13.

Inverse kinematics computes generalised coordinate values that position the model in a position and orientation which matches experimental markers at each time step. In mathematical terms, it is expressed as a weighted least squares problem whose solution intends to minimise both marker and coordinate errors. The latter refers to the difference between the experimental coordinate value and the one computed by the inverse kinematics. As for marker error, it is the distance between an experimental marker and the corresponding virtual one on the model when it is positioned using the generalised coordinates computed by the inverse kinematics solver. As an output `.mot` files containing these generalised coordinate trajectories is obtained [10] [104].

Filtering and downsampling

The output kinematics obtained from *OpenSim* were filtered using a Butterworth low pass filter with a cutoff frequency of 6 Hz. As for the IMU data, since they were collected at a higher frequency, they were resampled to match the OMC data at 100Hz for the patient dataset and 200Hz for the treadmill dataset. This step will be explained in more detail in Chapter 3 Section 3.2.1, since the data processing ways were investigated in the sensitivity analysis.

2.3.2 Features extraction

Features were extracted from the IMU signal using an open-source Python package called `tsFresh`. It automatically calculates up to 794 time series characteristics, called features, per channel of data. For the walking model, there are twelve channels for the two IMUs: three accelerations and three angular velocities for each sensor.

A moving cut-out window was used in the data extraction process. It shifted over the data to create smaller time series cut-outs ending one time step later than the one before and extract features only on these. This process, called rolling, is represented in Figure 2.6 [105]. This window size has an impact on the resulting predictions. The full machine learning algorithm was run on two reference treadmill individuals (P08 and P07) with different window sizes to determine the optimal one. The window sizes tested were 10, 20, 40, 60, 80, 100 and 150 frames and the root-mean-square errors, mean absolute errors, and coefficient of correlation were computed and compared. For the personalised models, a size of 1 second of data was selected, as well as for the generalised overground models. However, the generalised treadmill models comprised a larger sample of data, which could not be handled with a window size as large as 1 second. Therefore, it has been reduced at 0.25 seconds. All these results are presented in Chapters 3.

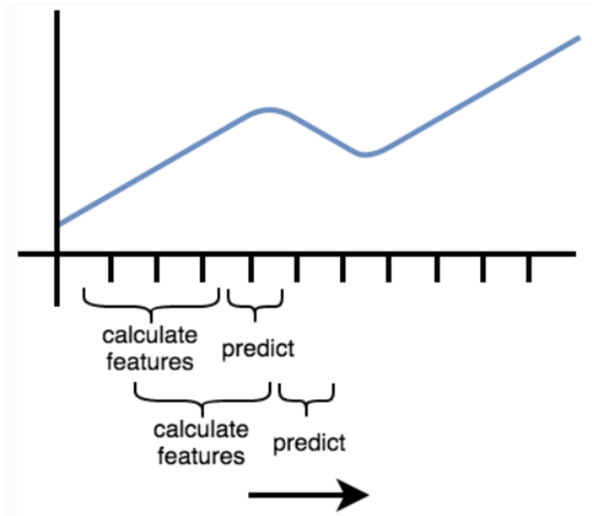

Figure 2.6: Principle of rolling (adapted from [105]).

2.3.3 Feature selection and model training

To select the best set of features for the walking, it was first trained with the full 794 features. Then, to reduce this number of features to only those that are relevant and significant for the target, i.e. the knee angle, a multivariate nonlinear regression was performed with the *Sklearn*-compatible estimator `RandomForestRegressor` [103] [105]. The correlation between each feature and the latter is estimated by computing the coefficient of determination R^2. After their sorting, only the subset of most important features was selected. Their number depends on several factors such as the computation time and the accuracy of prediction. For this model, the top 100 features were selected to estimate the knee angle. Finally, this set of features was used to train the multivariate nonlinear regression.

Chapter 3

Evaluation of the Machine Learning Algorithm

This chapter concerns the evaluation of the machine learning algorithm. The methodology followed to do so is described. Two types of model were generated and analysed: personalised and generalised models. In addition, a sensitivity analysis was performed to tune the parameters and data processing the best possible way. Finally, the results are discussed assessing the usability of this workflow for clinical applications.

3.1 Methodology for Assessment

First of all, as mentioned in Chapter 2 Section 2.3, preoperative data from patients 1 to 10 were not usable due to the synchronisation issue between the IMU and OMC data encountered during the processing. Moreover, for the postoperative gait of patients 4, 9, 10, and 12 another matching issue was encountered between the OMC and IMU data. A resynchronisation could have been possible for this problem (unlike for the other one which made these data useless) but have not been performed due to the lack of time to meet the deadline for this project. As for the most recent preoperative gait data obtained after addressing the first synchronisation issue, a choice was made to separate preoperative and postoperative models to keep them consistent. It would have allow us to compare gait before and after as well as possible, giving the preoperative model a diagnostic capability. Moreover, it was expected that the preoperative data would be more heterogeneous than the postoperative data, given the pain and difficulty to walk from their injured knee. Hence, the heterogeneity coupled to the extremely small sample (only Patients 12 and 13) would have meant that the model will have least chance of working. This is why only postoperative data were considered at this stage and no comparison can be made between preoperative and postoperative gait from this viewpoint. Finally, patient 13 did not have the chance to perform his postoperative gait due to unexpected confinement in New Zealand caused by the Covid-19 pandemic. Nevertheless, since the study extends over three years and that it is still its beginning, more patients will be enrolled which will expand the data samples, enabling to create the preoperative model and its comparison with the postoperative one. In addition, the resynchronisation, not performed by lack of time, will also be executed and patient 13 will perform its postoperative gait now that the lockdown is over.

Table 3.1 represents the data usability and the issues encountered. Consequently, only seven postoperative gait was used to train and assess the machine learning algorithm developed (3 Female, 4 Male ; 70.24 ± 7.2775 years old ; 1.69 ± 0.09 m ; 82.47 ± 18.38 kg).

ID	Preop Gait	Postop Gait	IMU Clinical follow-up
1	Sync issue	OK	OK
2	Marker visibility issue	OK	OK
3	Sync issue	OK	OK
4	Sync issue	Sync issue[1]	OK
5	Sync issue	OK	OK
6	Sync issue	OK	OK
7	Sync issue	OK	OK
8	Sync issue	OK	OK
9	Sync issue	Sync issue[1]	OK
10	Sync issue	Sync issue[1]	Not performed
12	Not used	Sync issue[1]	OK
13	Not used	Covid-19	OK

Table 3.1: Data usability and issues encountered.

In view of the lack of experimental patients' data obtained at this stage of the study, the model was also trained on two other walking gait datasets already collected and processed from previous studies: one set walking overground and another on a treadmill. The first one was collected at *VICON IMeasureU* headquarters. The cohort consisted of four healthy adult participants with no recent history of knee injury (2 Female, 2 Male ; 28.8 $\pm$ 3.43 years old ; 1.74 $\pm$ 0.08 m [body mass not recorded]). As for the treadmill dataset, it was collected at *AUT Millenium*'s motion capture facility (Mairangi Bay, NZ) and consisted of ten healthy adult participants with no recent history of knee injury (4 Female, 6 Male ; 27.4 $\pm$ 4.7 years old ; 1.75 $\pm$ 0.08 m ; 72.32 $\pm$ 10.33 kg). For these two datasets, optical marker trajectories were captured at 200 Hz using again a *VICON* motion capture system. Moreover, synchronous IMU data were captured at 250 Hz from *IMeasureU Blue Thunder* sensors–the older version of the *VICON Blue Trident*. It should be acknowledged that the new versions of IMeasureU's sensors, *VICON Blue Trident*, has an onboard 6-axis sensor fusion algorithm providing quaternion data at a frame rate that is acceptable for such applications. Therefore, the inherent drift, due to the numerical integration of IMU data to obtain orientation and position information, is corrected by the sensor itself instead of the principal component analysis-based algorithm that the company developed to overcome this limitation. Optical markers were placed on body segments in accordance with the UWA marker set [96], just as for the patients (see Fig. 2.2 (left)). IMU sensors were also placed on each thigh and shank of the participants (four sensors in total) (see Fig. 2.2 (right)). Several tasks were performed for both overground and treadmill data. Table 3.2 represents the parameters of the three experimental datasets used to evaluate the machine learning pipeline.

Two types of models were generated to test the algorithm. Firstly, a person-specific model generated directly from that person's data. One can trivially imagine collecting training data for this model in a clinic setting, prior to tracking and predicting knee flexion data outside the clinic. The second model tested was a generalised model in which data from a whole cohort are used to create a model that can be used to predict knee flexion angles from individuals who are not part of the training cohort. This would require a lot of data across a population to reach an acceptable accuracy. To simulate the effect of sample size on the predicted knee angle values, the machine learning pipeline was trained on the treadmill data using two scenarios: one which used 10 seconds of walking data, and another using 50 seconds of walking data.

[1]Second synchronisation issue encountered which could be fixed with more time (also linked to OMC file).

	Overground		Treadmill
	Patients	**Healthy**	**Healthy**
Number of participants	7 (3F ; 4M)	4 (2F ; 2M)	10 (6F ; 4M)
VICON Cameras	MX13	Vantage 16	MX T40-S
Marker sampling rate	100 Hz	200 Hz	200 Hz
IMeasureU sensors	Trident	Blue Thunder	Blue Thunder
IMeasureU capture rate	Collected: 1149 Hz Downsampled: 100 Hz	250 Hz	250 Hz
Static standing	√	√	√
Sit to stand	√	√	–
Squats	√	–	√
Seated knee flexion	–	√	–
Stair climb	√	–	–
Walking	√	√	√

Table 3.2: Experimental datasets summary.

Two analyses were performed to evaluate both models. For the person-specific model, a temporal cross-validation was carried out: the model was trained on the first 70% of the person's data and the remaining 30% were predicted. For the generalised model, a Leave-One-Out cross validation was executed to evaluate the potential generalisability of it: the model was trained on every individual of the cohort but one, which was used for the prediction. To assess each model, the root mean square error (RMSE), mean absolute error (MAE), and coefficient of determination (R^2) values were calculated.

As a reminder, the RMSE is the sample standard deviation of the differences between predicted values and observed values. Mathematically, the formula is:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{3.1}$$

where y_i represents the actual values, $\hat{y}_i$ is the prediction, and N the sample size. Since the errors are squared before being averaged, RMSE assigns a higher weight to larger errors. This metric is thus more useful to detect the presence of large errors which drastically affect the performance of the model [106] [107].

MAE is the average of the absolute difference between the observed and predicted values. This metric can be expressed as:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{3.2}$$

where y_i represents the actual values, $\hat{y}_i$ is the prediction, and N the sample size. Unlike the RMSE, the MAE is a linear score: all the individual differences are weighted equally in the average. Therefore, MAE is more interesting from an interpretation viewpoint [106] [107].

Finally, R^2 represents the coefficient of how well the model fits a given dataset. Mathematically, its formula is given by:

$$R^2(y,\hat{y}) = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{3.3}$$

where y_i represents the actual values, $\hat{y}_i$ is the prediction, $\bar{y}$ is the mean of actual values, and N the sample size. This metric indicates how close the regression line, representing the predicted values plotted, is to the given data values. Its value lies between 0 and 1, 0 indicating that the model does not fit the actual data whereas 1 represents the perfect fit [106] [107].

Figure 3.1 summarises the different steps to obtain the knee flexion prediction from the raw data.

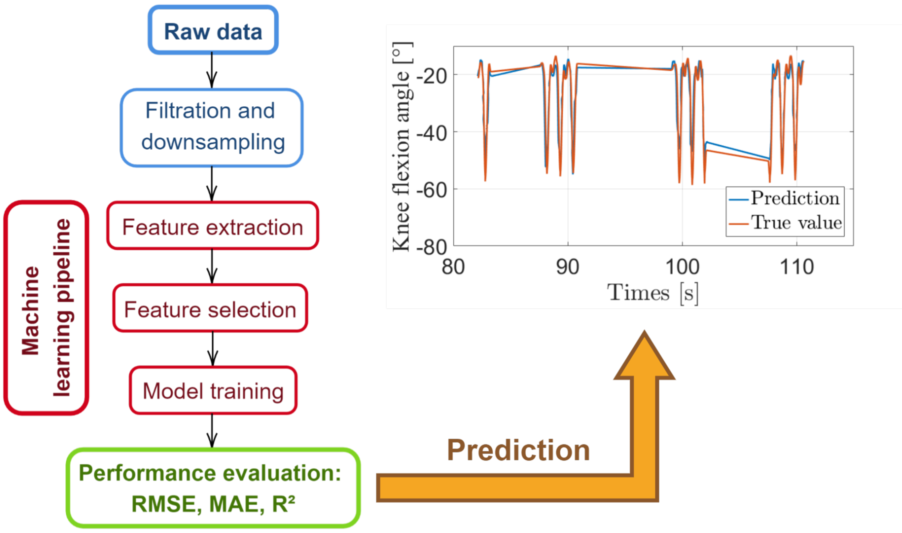

Figure 3.1: Resume of the different steps of the workflow presented to obtain the knee kinematics from raw IMU data.

3.2 Personalised Models

As previously explained, personalised models were trained only on data collected from that peculiar individual. Predictions were performed for the three datasets by training the models on a subset of people's data and validating, i.e. predicting, them on the remaining data (more or less 70% training and 30% validation). One should notice that, for the treadmill dataset, this split is not respected. Indeed, by running the algorithm, we found out that using 10 seconds (17%) of data was enough to train the model. Hence, the last 50 seconds (83%) were used for validation. Indeed, while overground datasets includes many variations in walking speed and is composed of five walking trials back and forth, the treadmill dataset is much more consistent and longer. Therefore, it does not require the 70-30 split and using a smaller sample size for training sped up the algorithm (this will be demonstrated in the sensitivity analysis).

3.2.1 Sensitivity Analysis

A sensitivity analysis was performed on the parameters used in the model: the window size, the sample size used for training and validation, and the number of features used in the regression model for two healthy individuals from the treadmill dataset (P08 and P07). Moreover, the effect of filtering and downsampling were also investigated. Therefore, the following combinations were generated and evaluated (see Tab. 3.3):

A. IMU filtered and downsampled at 100 Hz, OMC filtered and downsampled at 50 Hz, training on $\pm$ 10 seconds and validation on $\pm$ 50 seconds;

B. IMU filtered and downsampled at 100 Hz, OMC filtered and downsampled at 50 Hz, training on $\pm$ 50 seconds and validation on $\pm$ 20 seconds;

C. IMU filtered, downsampled at 100 Hz and using the resultant, OMC filtered and downsampled at 50 Hz, training on ± 10 seconds and validation on ± 50 seconds;

D. IMU filtered, downsampled at 100 Hz and using the resultant, OMC filtered and downsampled at 50 Hz, training on ± 50 seconds and validation on ± 20 seconds;

E. IMU filtered and downsampled at 200 Hz, OMC filtered (no downsampling: 200Hz), training on ± 10 seconds and validation on ± 50 seconds;

F. IMU not filtered but downsampled at 200 Hz, OMC not filtered (no downsampling: 200 Hz), training on ± 10 seconds and validation on ± 50 seconds;

G. IMU not filtered but downsampled at 200 Hz and using the resultant, OMC not filtered (no downsampling: 200 Hz), training on ± 10 seconds and validation on ± 50 seconds;

H. IMU filtered and downsampled at 200 Hz and using the resultant, OMC filtered (no downsampling: 200 Hz), training on ± 10 seconds and validation on ± 50 seconds.

	IMU			**OMC**		**Training**	**Validation**
	Filtered	**Downsampled**	**Resultant**	**Filtered**	**Downsampled**	**[s]**	**[s]**
A	√	√ (100 Hz)	–	√	√ (50 Hz)	10	50
B	√	√ (100 Hz)	-	√	√ (50 Hz)	50	20
C	√	√ (100 Hz)	√	√	√ (50 Hz)	10	50
D	√	√ (100 Hz)	√	√	√ (50 Hz)	50	20
E	√	√ (200 Hz)	–	√	–(200 Hz)	10	50
F	-	√ (200 Hz)	√	-	- (200 Hz)	10	50
G	–	√ (200 Hz)	√	–	–(200 Hz)	10	50
H	√	√ (200 Hz)	–	√	–(200 Hz)	10	50

Table 3.3: Summary of each combination of parameters.

The IMU and OMC data filtering were performed using a 4$^{\text{th}}$-order low-pass Butterworth filter at 30 Hz and 6 Hz, respectively. Indeed, such filter is commonly used in biomechanics to reduce high-frequency noise due to inaccuracies in the marker reconstruction or soft tissue artefact, for example. Different methods exist to optimally select the appropriate cut-off frequency for these signals. However, we just decided to test several frequencies by observing the difference between the filtered and the raw data and to select the best since our focus was not on this part. Using a too high frequency will remove only very little noise, whereas a low one will reject some of the signals.

For OMC data, a commonly used cut-off frequency is between 4 and 8 Hz since most human movement are typically low frequencies [108]. Therefore, four different cut-off frequencies have been investigated: 3, 6, 9, and 12 Hz. This test was performed on the right knee angle obtained after the inverse kinematics on `OpenSim` from P08 (treadmill dataset). Figure 3.2 (left) represents a zoom on the filtered signals obtained with these four cut-off frequencies, as well as the raw data. As can be observed, a cut-off frequency of 3 Hz does not maintain the signal of interest, whereas the others seem to work well. Therefore, 6 Hz was chosen since it was the one who most filtered noise while keeping the signal. The comparison between raw, filtered and downsampled signals that are going to be used for the sensitivity analysis is made in Figure 3.2 (right).

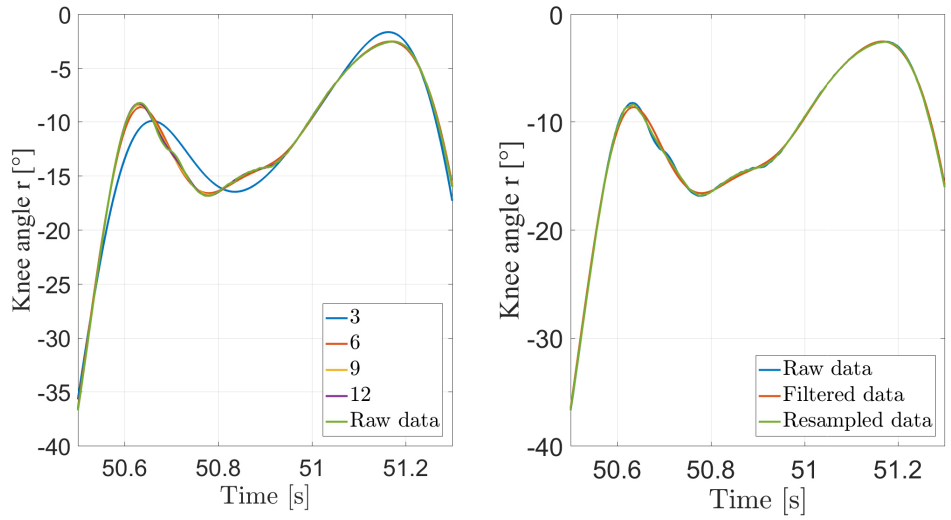

Figure 3.2: Left: Zoom on the right knee angle [°] of one participant (P08) filtered using different cut-off frequencies compared to the raw data as a function of time [s]. Right: Comparison of the zoom on the right knee angle [°] of one participant (P08) obtained initially (raw data), after filtration with a cut-off frequency of 6 Hz and after downsampling from 200 to 50 Hz as a function of time [s].

For IMU data, a similar test was made on the resultant of the acceleration and angular velocity obtained from the IMU positioned on the right knee angle for a range of cut-off frequencies of 15, 30, 45, and 60. Figure 3.3 represents a zoom on the filtered signals (acceleration and angular velocity) obtained with these four cut-off frequencies, as well as the raw data. By the same reasoning as previously, 15 Hz was rejected since it was too low to maintain the signal of interest. Hence, 30 Hz was selected for IMU data. The comparison between raw, filtered and downsampled signals that are going to be used for the sensitivity analysis is represented in Figure 3.4.

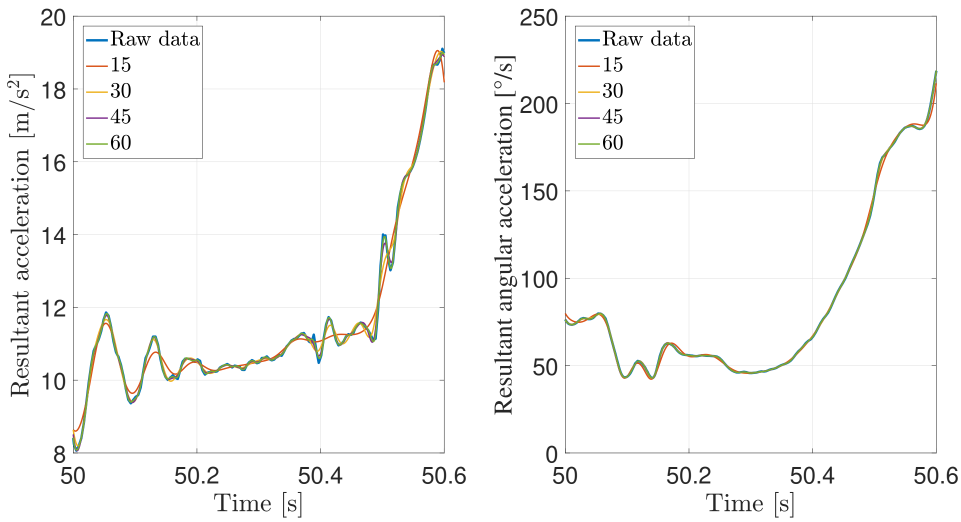

Figure 3.3: Left: Zoom on the right knee resultant acceleration [m/s^2] of one participant (P08) filtered using different cut-off frequencies compared to the raw data as a function of time [s]. Right: Zoom on the right knee resultant angular velocity [°/s] of one participant (P08) filtered using different cut-off frequencies compared to the raw data as a function of time [s].

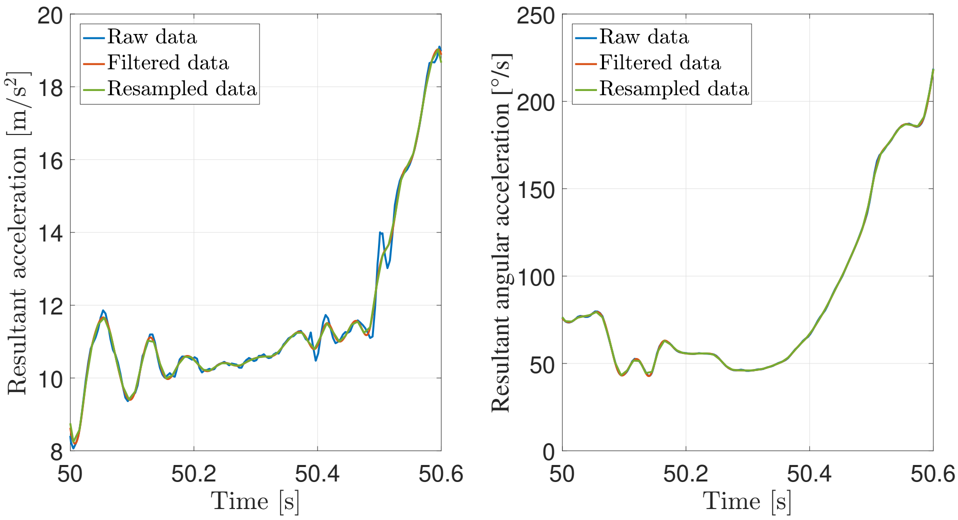

Figure 3.4: Left: Comparison of the zoom on the right knee resultant acceleration [m/s^2] of one participant (P08) obtained initially (raw data, after filtration with a cut-off frequency of 30 Hz and after downsampling from 1000 Hz to 100 Hz) as a function of time [s]. Right: Comparison of the zoom on the right knee resultant angular velocity [°/s] of one participant (P08) obtained initially (raw data, after filtration with a cut-off frequency of 30 Hz and after downsampling from 1000 Hz to 100 Hz) as a function of time [s].

Figures 3.5, 3.6, and 3.7 present the RMSE, MAE, and Correlation coefficients from the sensitivity analysis, where eight combinations of data processing and parameters used in feature extraction were investigated (identified as A through D above, with 100–solid lines–or 50 features–dashed lines).

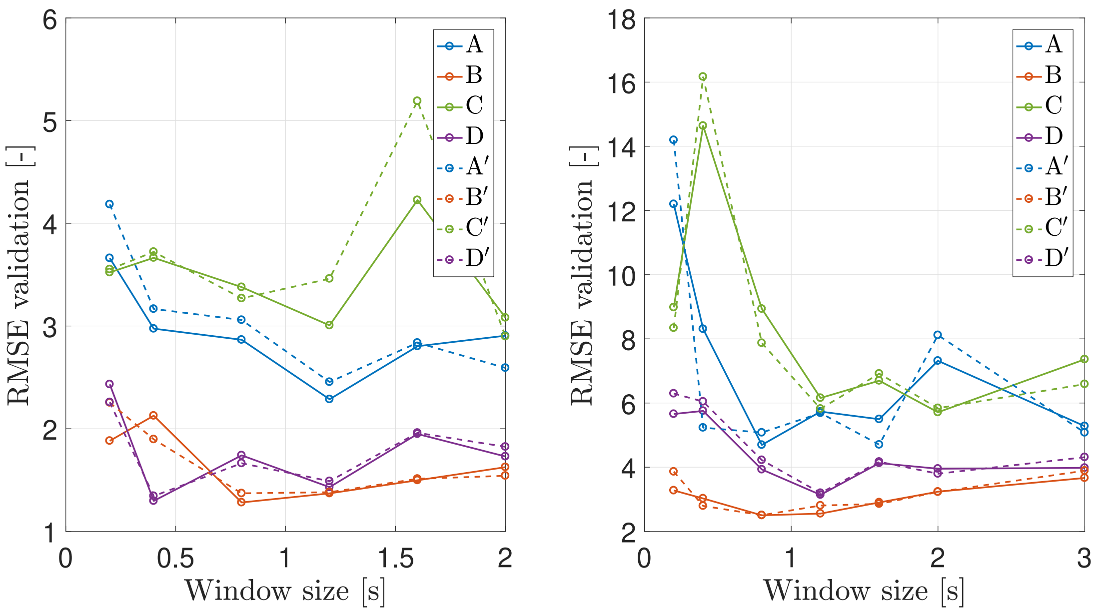

Figure 3.5: RMSE for the prediction of the knee angle from a person-specific machine learning model of two participants (P08 on the left and P07 on the right) using treadmill data, illustrating prediction error as a function of window size (x-axis). The dotted line shows the reconstruction using the top 50 features and the solid line used the top 100 features.

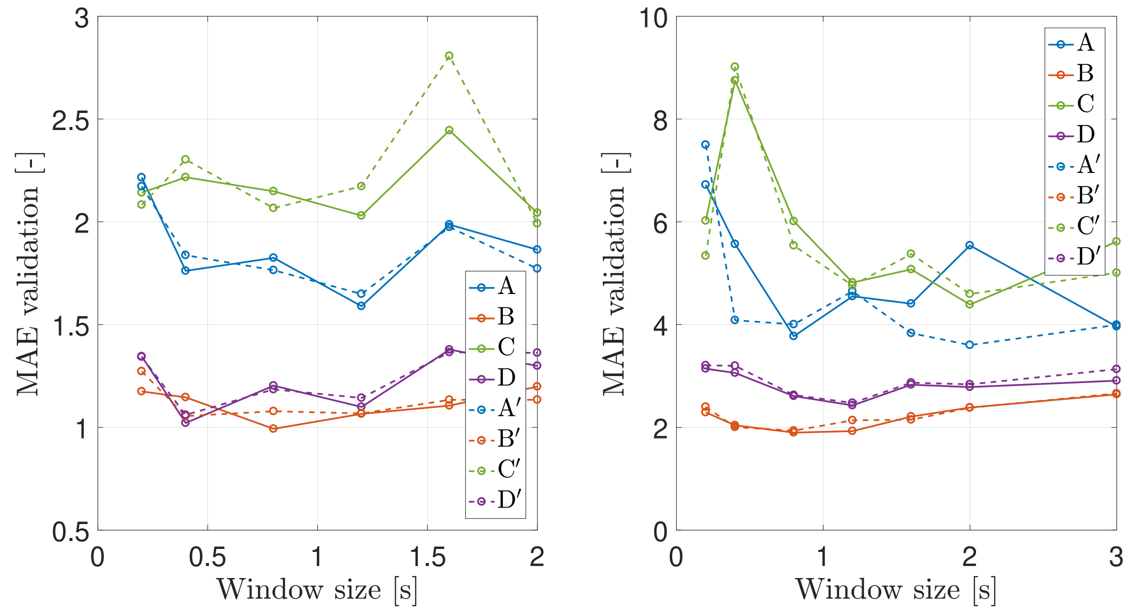

Figure 3.6: MAE for the prediction of the knee angle from a person-specific machine learning model of two participants (P08 on the left and P07 on the right) on a treadmill for different window size and for different cases. Dotted line used for the top 50 features and solid line used for the top 100 features.

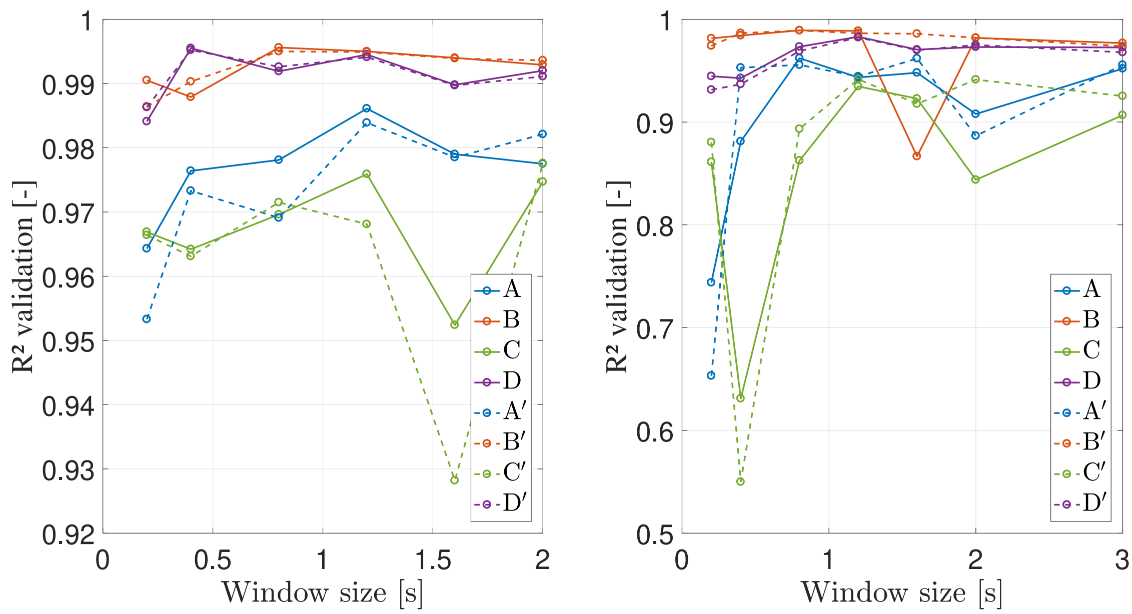

Figure 3.7: Correlation coefficients for the prediction of the knee angle from a person-specific machine learning model of two participants (P08 on the left and P07 on the right) on a treadmill for different window size and for different cases. Dotted lines used for the top 50 features and solid line used for the top 100 features.

These figures show that the window size matters (as error changes along the *x*-axis), but not as much as the length of the sequence of data given to the model to train (50 seconds worked much better than 10 seconds). Indeed, there is a distinguishable gap between Models A and C and Models B and D, whereas the effect of the window size is not repeatable on the two reference individuals (bad results obtained around 1.5 seconds for P08 and around 0.4 seconds for P07). The cases based on the raw data seem slightly better than the ones using the resultant data, but not as robust. Indeed, using the resultant means not taking into account the orientation of data and thus, the orientation of the sensors. It also appears that using 50 features was adequate to reconstruct the data when training on the longest sequence of data–dashed vs solid lines. However, the overground dataset collected has a very short sequence of data with much larger variations meaning that using at least 100 features was necessary to get similar results. Therefore, Models B and D outperformed the other models consistently across the cohort.

The combination of paramters in Model B (IMU filtered and downsampled at 100 Hz, OMC filtered and downsampled at 50 Hz, training on ± 50 seconds and validation on ± 20 seconds) performed the best. Figure 3.8 (top) illustrates the ability of Model B to predict knee flexion angle across a period of sixteen strides with a window size of 40 frames (0.8 seconds). For comparison, the prediction of Model C (IMU filtered and downsampled at 100 Hz, OMC filtered and downsampled at 50 Hz, training on ± 10 seconds and validation on ± 50 seconds) with the same window size of 40 frames (0.8 seconds of data) is also shown in Figure 3.8 (bottom). Note that a zoom has been made on sixteen strides for the sake of visibility. It illustrates the difference in model predictions with less training data. One should observe how the raw data predictions are presented. It would be logical to filter them using a low pass filter at ± 6 Hz. However, for the purpose of comparison, only the raw data are presented here. A comparison between the errors and correlation coefficients can also be observed in Table 3.4. Both types of errors are more than twice higher in Model C than in Model B.

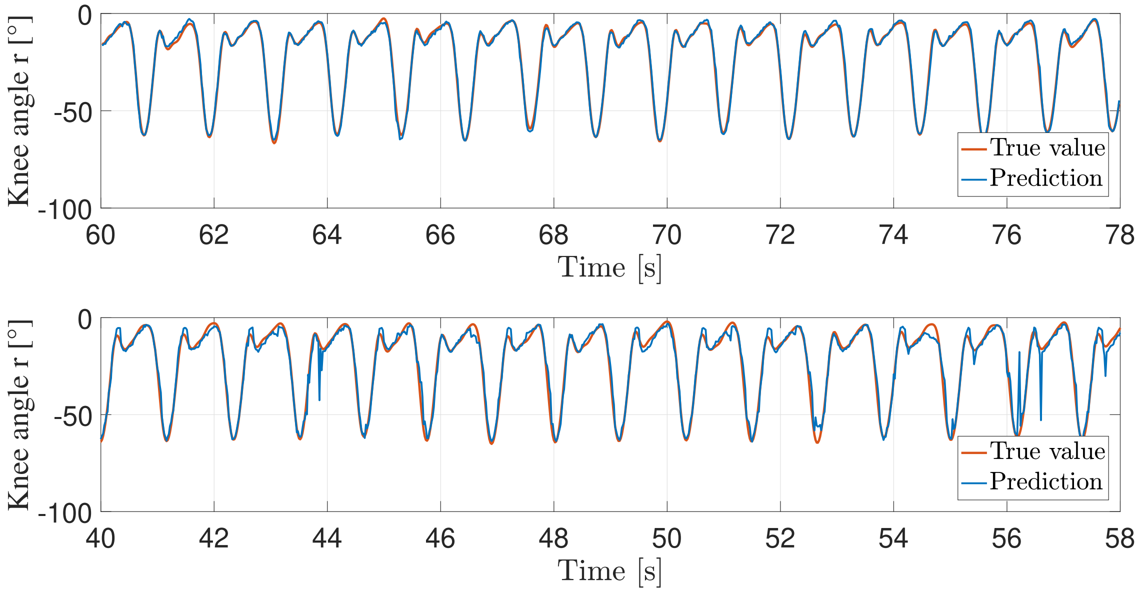

Figure 3.8: Top: Right knee flexion angle obtained for individual P08 for the combination B (filter and downsampling without using the resultant and trained during ± 50s) for a window size of 40 frames (corresponding to 0.8s) and extracting the top 100 features. Bottom: Zoom on the right knee flexion angle obtained for individual P08 for the combination C (filter and downsampling using the resultant and trained during ± 10s) for a window size of 40 frames (corresponding to 0.8s) and extracting the top 100 features.

Model	**RMSE [°]**	**MAE [°]**	**R^2**
B	1.28	0.99	1
C	3.38	2.15	0.97

Table 3.4: Comparison between errors and correlation coefficients obtained with Model B and C.

Figures 3.9, 3.10, and 3.11 show the RMSE, MAE, and Correlation coefficients from the sensitivity analysis, where the last four combinations of data processing and parameters used in feature extraction are investigated (identified as Model combinations E through H above). One should notice that for these, only the training of ± 10 seconds was considered. Indeed, the process time to run the machine learning algorithm was around 2 minutes only for Models A and C with only $\pm$ 10 seconds of training. It increased to around 8 minutes for Models B and D with $\pm$50 seconds of training and to more than 10 minutes for Models E to H with again only $\pm$ 10 seconds of training. Therefore, since the effect of the training data size has already been demonstrated, the longer sequence of data to train the model was not investigated for these last models.

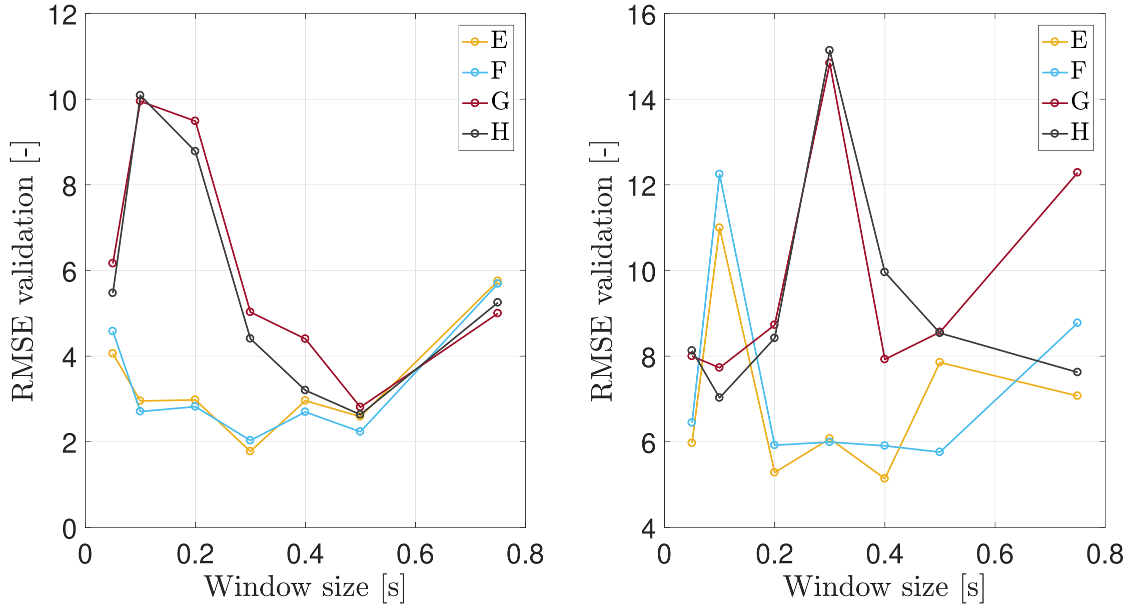

Figure 3.9: RMSE for the prediction of the knee angle from a person-specific machine learning model of two participants (P08 on the left and P07 on the right) using treadmill data, illustrating prediction error as a function of window size (x-axis).

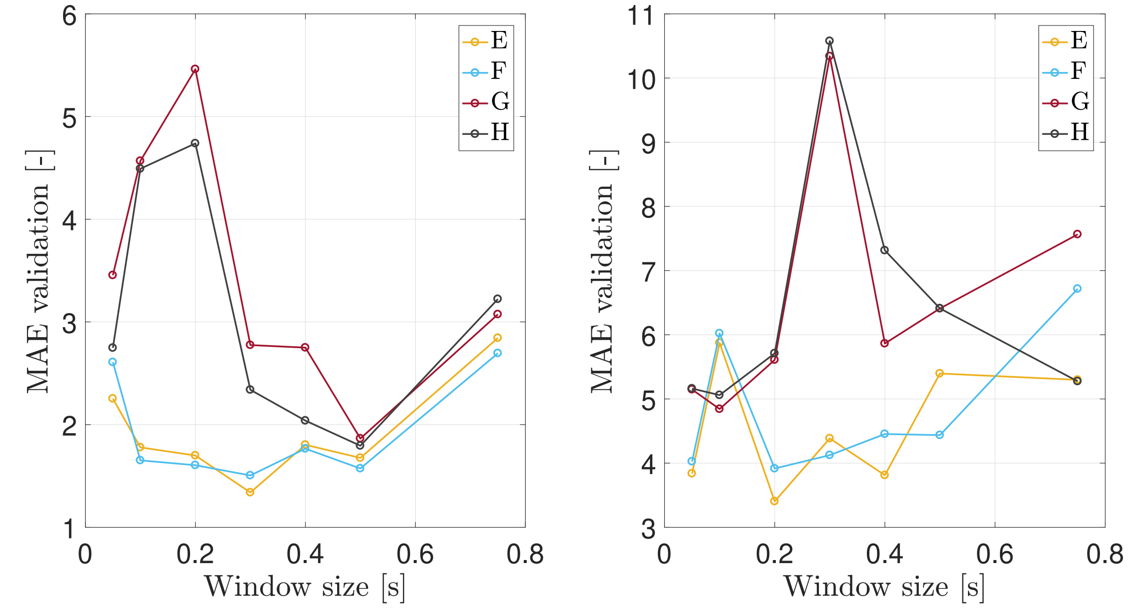

Figure 3.10: MAE for the prediction of the knee angle from a person-specific machine learning model of two participants (P08 on the left and P07 on the right) on a treadmill for different window size and for different cases.

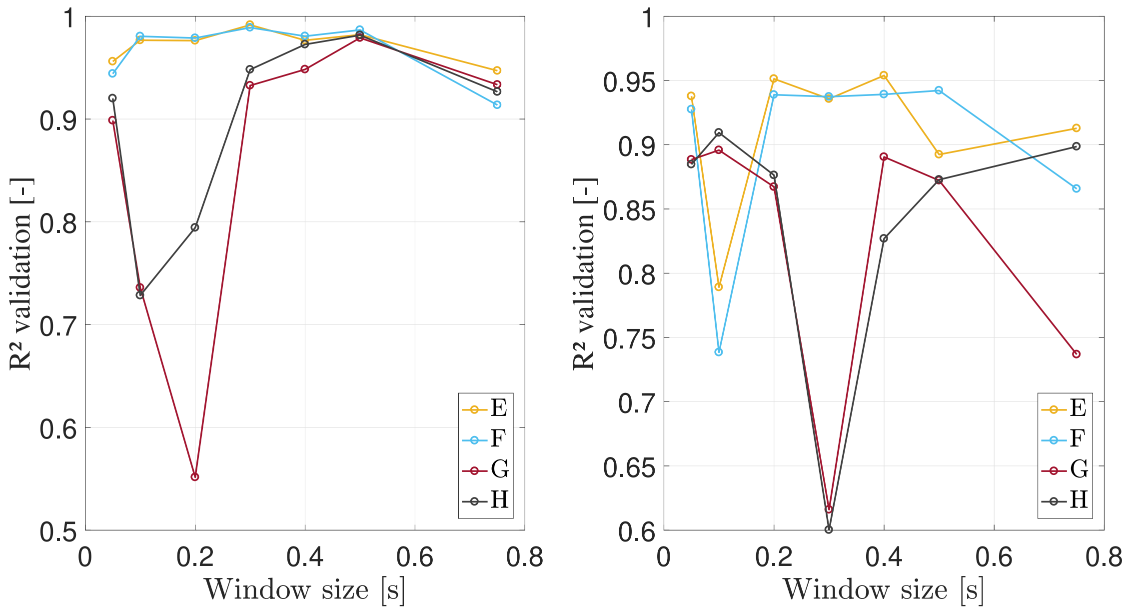

Figure 3.11: Correlation coefficients for the prediction of the knee angle from a person-specific machine learning model of two participants (P08 on the left and P07 on the right) on a treadmill for different window size and for different cases.

From these figures, one can notice that there are two models that performed much better than the others: Models E and F. Compared to the previous models (A to D), Models E and F (no downsampling of OMC data) perform almost as good as the previous models with the similar length of training data (Models A and C). By contrast, Models G and H (no downsampling of OMC data and use of the resultant for IMU data) present higher errors and smaller correlation, especially at smaller window sizes ($\simeq 0.2 - 0.3$ seconds). Therefore, a combination of raw OMC data with the use of the resultant should be rejected. The effect of filtering is not distinguishable since Models E and F give similar results, as well as Models G and H.

To resume this sensitivity analysis, here are the different observations made:

- The window size does not allow us to reject a peculiar range since it is not repeatable between the two reference individuals;
- The longer the sequence of data used for training, the better the model performance but the slower the process time to train it;
- In the case of downsampled data, using the resultant decreases slightly the model performance but since the resultant vector is not altered by the rotation of the sensors, the model is more robust;
- In the case of raw data, using the resultant decrease the model performance more sharply;
- The effect of filtering is not distinguishable.

Given that this sensitivity analysis was performed on the treadmill data, more homogeneous and longer sequence of data than for overground, the combination of data processing and parameters selected to generate personalised models is Model C. Indeed, despite the better performance of the longer sequence of training data, using only 10 seconds is closer to the 70% of data that will be used to train overground models (the overground sequence of data is not long enough). In any case, the errors and correlation coefficients obtained in the case of a shorter sequence of data are still acceptable. Moreover, as explained previously, the use of the resultant increases the robustness much more than it decreases the performance of the model. Data were also downsampled and filtered since it has been proven not to be prejudicial for the model accuracy and that it allows to decrease considerably the central processing unit (CPU) time. The sample rate chosen was of 100 Hz for both IMU and OMC data. Finally, a window size of ± 1 second has been selected - corresponding to 100 frames for a dowsampling to 100 Hz.

3.2.2 Influence of IMUs's Orientation on Person-Specific Models

An investigation on the effect of IMU orientation on the model's estimations was also achieved for the combination A and B (longer and shorter sequence of training data) for the same two individuals (P08 and P07). A rotation about the y-axis (long axis of the leg) of several degrees (from 0° to 10°) was performed on the acceleration and gyroscope vectors, as follows:

$$\begin{bmatrix} A'_x \\ A'_y \\ A'_z \end{bmatrix} = \begin{bmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ -\sin\alpha & 0 & \cos\alpha \end{bmatrix} \begin{bmatrix} A_x \\ A_y \\ A_z \end{bmatrix} \tag{3.4}$$

where α is the angle of rotation between 0° and 10° and $\mathbf{A}$ is the vector to rotate (acceleration or angular velocity). Therefore, the sensor data was systematically rotated to synthetically generate new data and then these new data were put into the model. Results obtained from this investigation have also been compared to combinations C and D which use the resultant IMU data instead of the raw data (now denoted R in the legend). This analysis enabled to highlight the influence of a small rotation (up to 10°) in the orientation of the sensors on the model performance.

Figures 3.12 to 3.17 illustrate this investigation. The model was robust to these rotations, and even with a 10° rotation, the predictions were better than what was achieved using the resultant IMU data. Therefore, even if the sensor is not perfectly positioned, it should not influence the results obtained. This finding is highly important since it means that our algorithm is robust enough to take into account the variability encountered with the positioning of the sensors, especially when they are not always positioned by the same person. Note the preferred window size of ± 1 second.

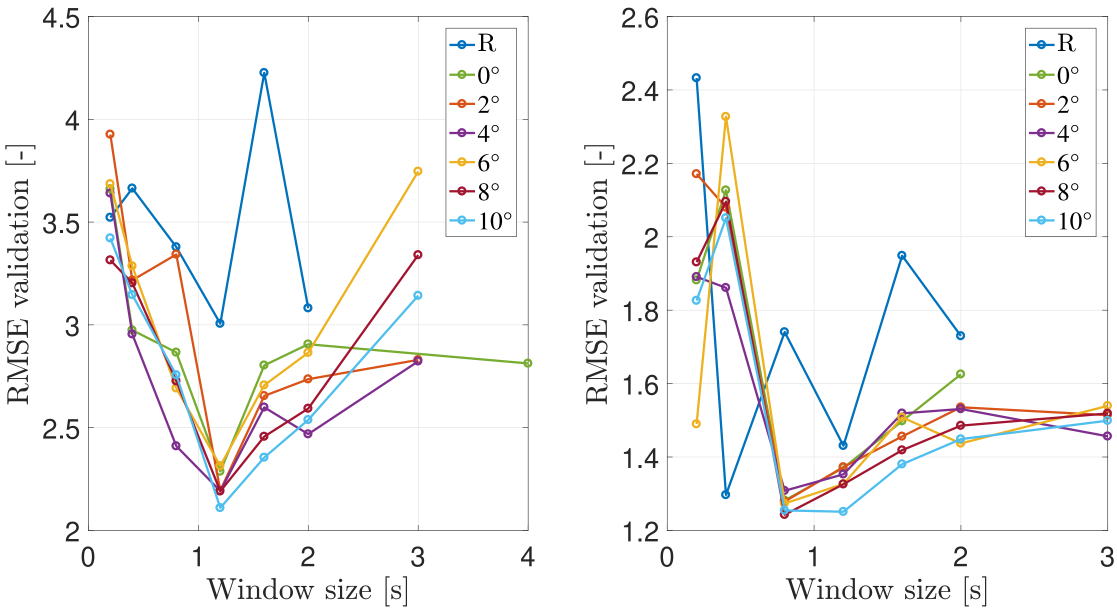

Figure 3.12: RMSE for the prediction of the knee angle from a person-specific machine learning model of participant P08 on a treadmill for different window size and for different IMUs's orientations (and R represents the use of the resultant). Left: training of ± 10 seconds and prediction for ± 50 seconds ; Right: training of ± 5 seconds and prediction of ± 20 seconds.

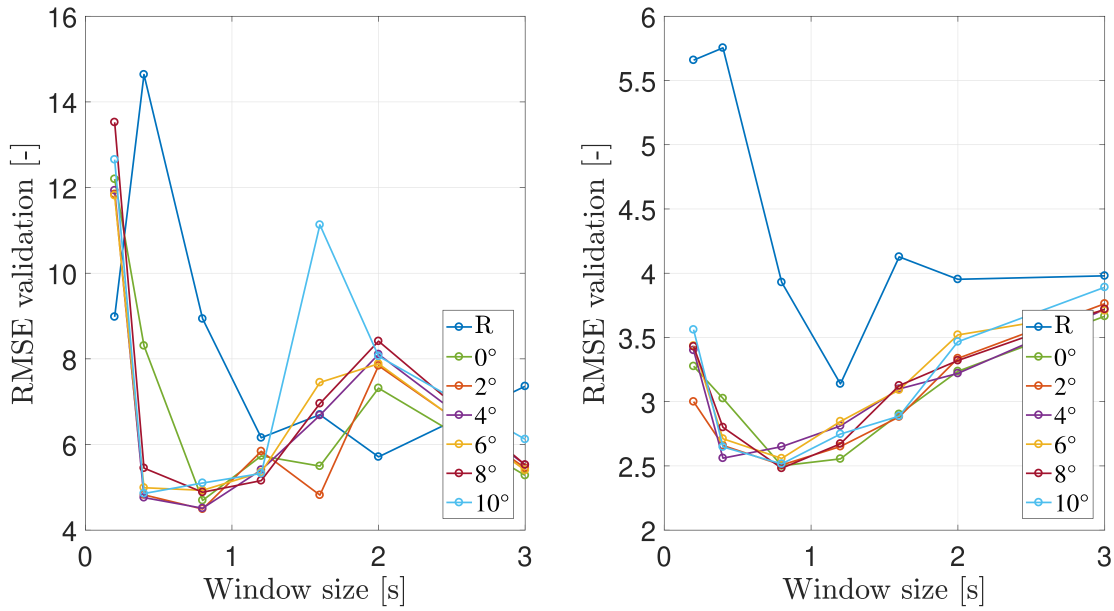

Figure 3.13: RMSE for the prediction of the knee angle from a person-specific machine learning model of participant P07 on a treadmill for different window size and for different IMUs's orientations (and R represents the use of the resultant). Left: training of ± 10 seconds and prediction for ± 50 seconds ; Right: training of ± 5 seconds and prediction of ± 20 seconds.

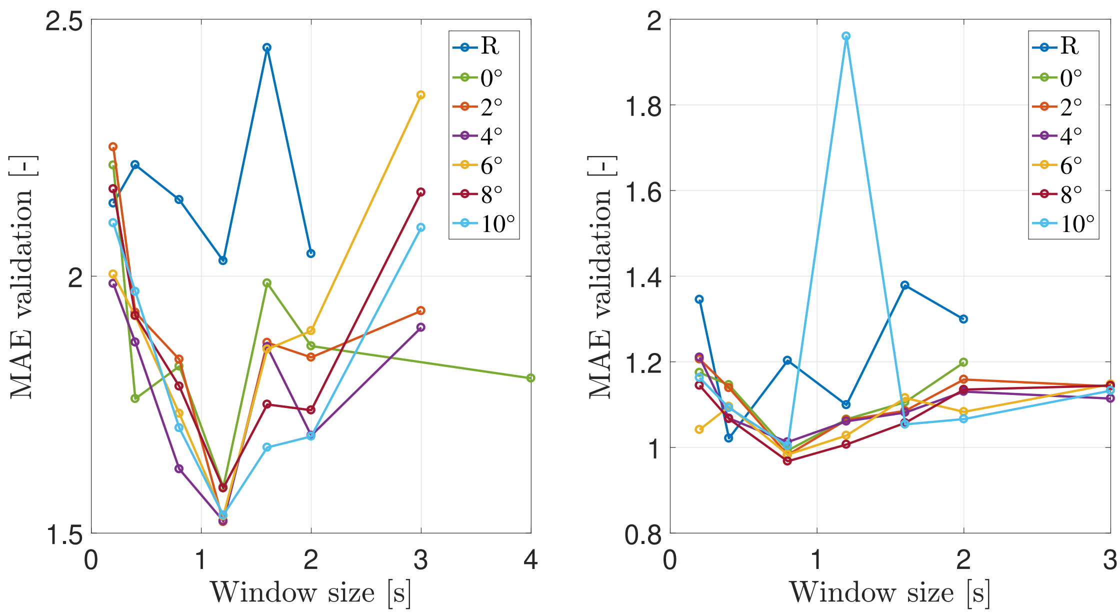

Figure 3.14: MAE for the prediction of the knee angle from a person-specific machine learning model of participant P08 on a treadmill for different window size and for different IMUs's orientations (and R represents the use of the resultant). Left: training of ± 10 seconds and prediction for ± 50 seconds ; Right: training of ± 5 seconds and prediction of ± 20 seconds.

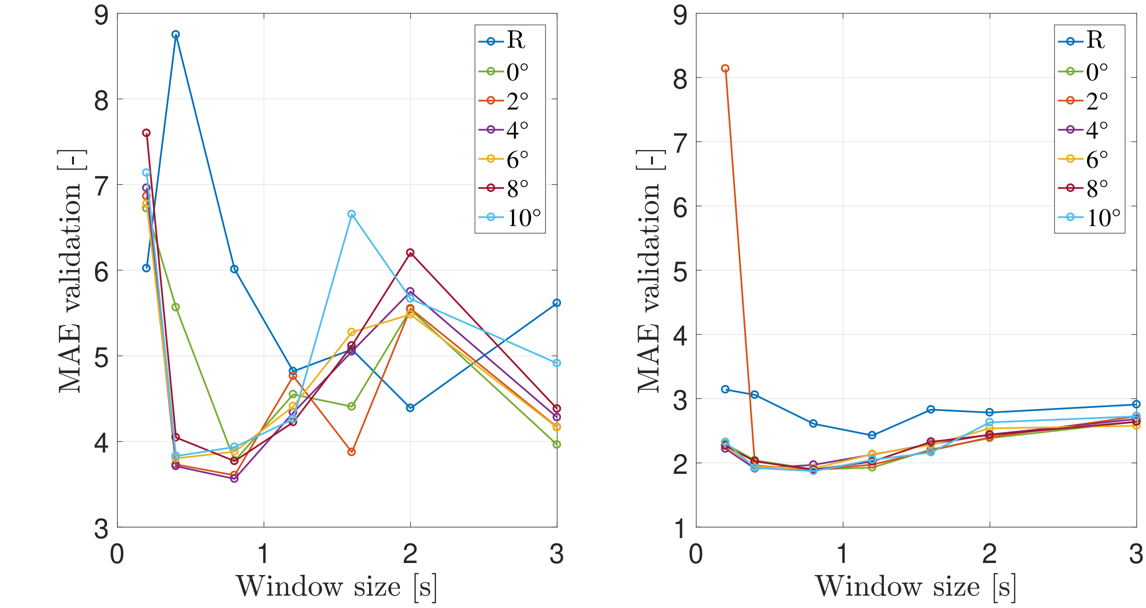

Figure 3.15: MAE for the prediction of the knee angle from a person-specific machine learning model of participant P07 on a treadmill for different window size and for different IMUs's orientations (and R represents the use of the resultant). Left: training of ± 10 seconds and prediction for ± 50 seconds ; Right: training of ± 5 seconds and prediction of ± 20 seconds.

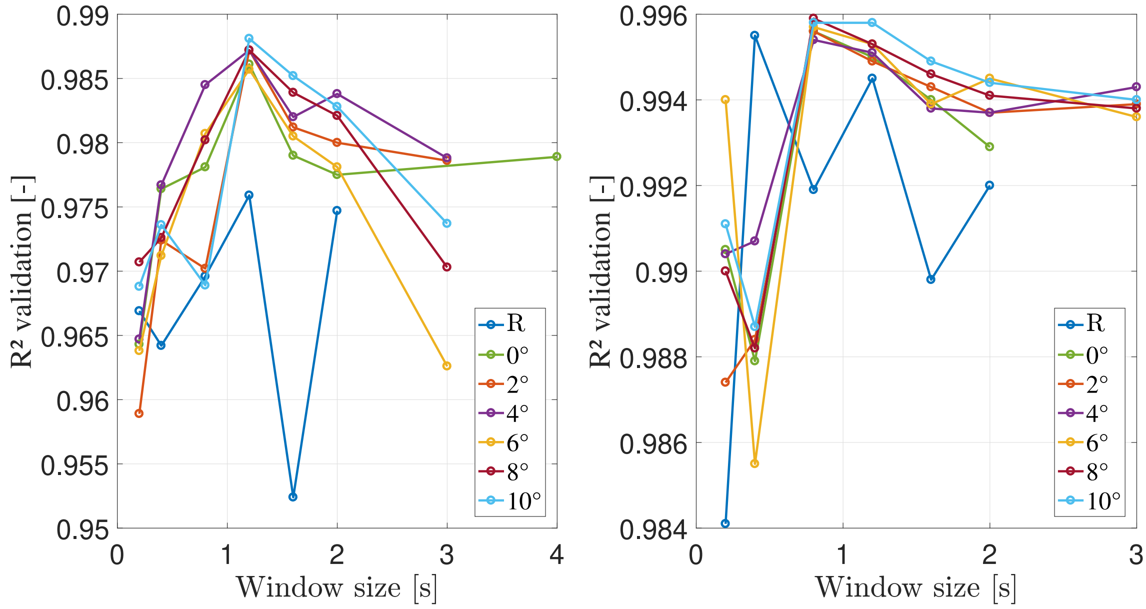

Figure 3.16: Correlation coefficients for the prediction of the knee angle from a person-specific machine learning model of participant P08 on a treadmill for different window size and for different IMUs's orientations (and R represents the use of the resultant). Left: training of ± 10 seconds and prediction for ± 50 seconds ; Right: training of ± 5 seconds and prediction of ± 20 seconds.

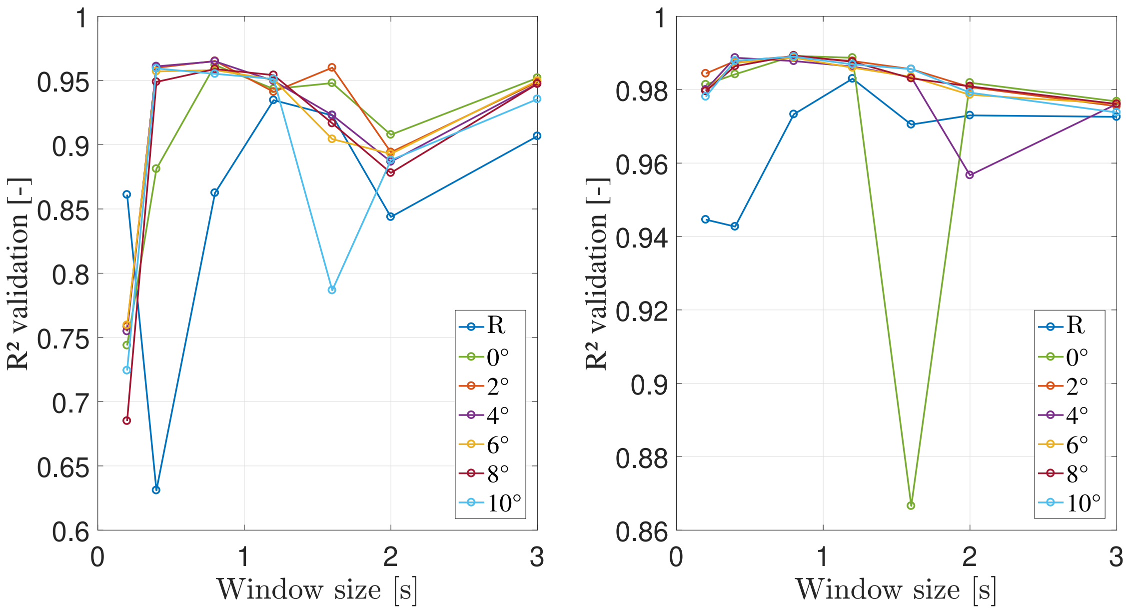

Figure 3.17: Correlation coefficients for the prediction of the knee angle from a person-specific machine learning model of participant P07 on a treadmill for different window size and for different IMUs's orientations (and R represents the use of the resultant). Left: training of ± 10 seconds and prediction for ± 50 seconds ; Right: training of ± 5 seconds and prediction of ± 20 seconds.

3.2.3 Analysis of Personalised Models

RMSE, MAE and correlation coefficients obtained after comparing OMC-derived and IMU-derived knee kinematics for the combination of data processing and parameters selected after the sensitivity analysis are listed in Table 3.5. For overground walking, IMU difference (left minus right IMU) was added to the data in order to increase its amount for training to the small step count.

Overground walking (Patient and Healthy)			
ID	**RMSE [°]**	**MAE [°]**	**R^2**
1	5.35	3.81	0.92
2	4.61	3.56	0.95
3	3.54	2.43	0.95
5	4.56	3.01	0.91
6	21.74	17.36	0.00
7	18.12	11.65	0.40
8	20.28	15.98	0.01
H05	5.22	4.15	0.95
H06	8.11	6.40	0.91
H11	12.8	9.55	0.55
H13	6.45	4.64	0.92
Average	10.07	7.50	0.68
S.D.	6.92	5.35	0.38

Treadmill walking			
ID	**RMSE [°]**	**MAE [°]**	**R^2**
P07	4.98	2.90	0.94
P08	3.67	1.83	0.96
P09	3.64	1.68	0.96
P10	3.99	2.48	0.95
P11	3.81	1.79	0.96
P12	6.23	3.19	0.90
P13	3.45	1.92	0.96
P15	3.01	1.79	0.97
P16	3.41	1.71	0.96
P17	4.29	2.49	0.94
Average	4.05	2.18	0.95
S.D.	0.94	0.55	0.02

Table 3.5: Comparison between OMC-derived and IMU-derived knee kinematics. The healthy participants are H05, H06, H11, H13.

Figures 3.18 to 3.20 illustrate the prediction of knee kinematics obtained from individual P08, Patient 3 and healthy individual H05. One should notice than the depicted straight lines represent periods with no joint kinematics (between walking trials) from the OMC data. They were skipped to reduce the input data.

Errors found for treadmill walking are all lower than 5 degrees, which is considered as our threshold. In the overground dataset, comparable results were observed with a MAE lower than the threshold. However, a few datasets (Patients 6, 7, 8 and H11) go well over this threshold, indicating that there might be issues with those data. Removing them from the cohort enabled to find a mean RMSE of 5.41°, a mean MAE of 4° and a mean correlation coefficient of 0.93. As expected in view of the heterogeneity of overground dataset, these values are higher than those of the treadmill dataset but still acceptable. As an example, the selected features for the treadmill personalised model of individual P08 are provided in Appendix F.

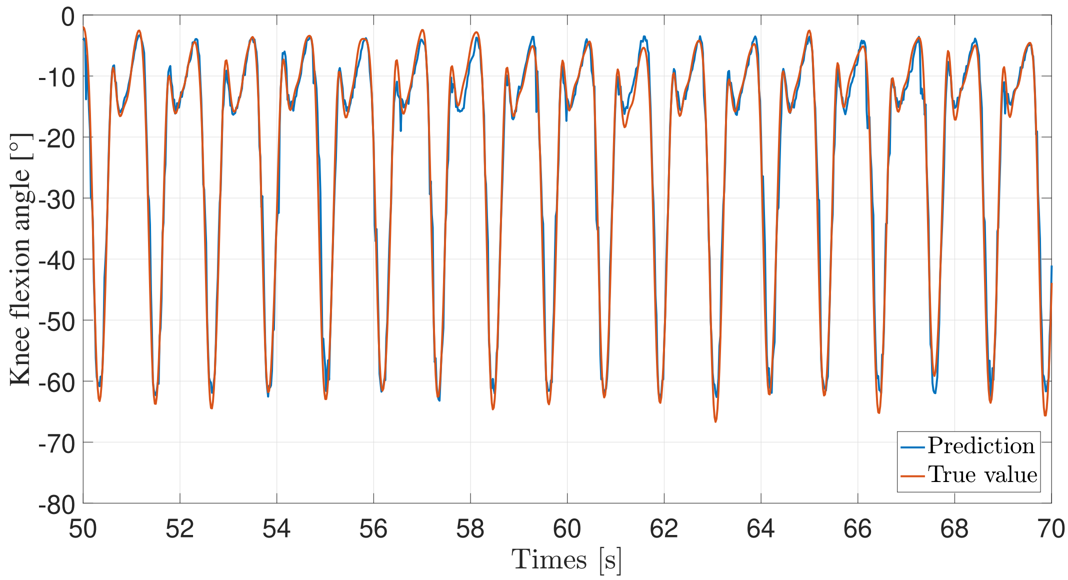

Figure 3.18: Predicted knee flexion from IMUs as compared to knee angles from OMC for treadmill walking (P08).

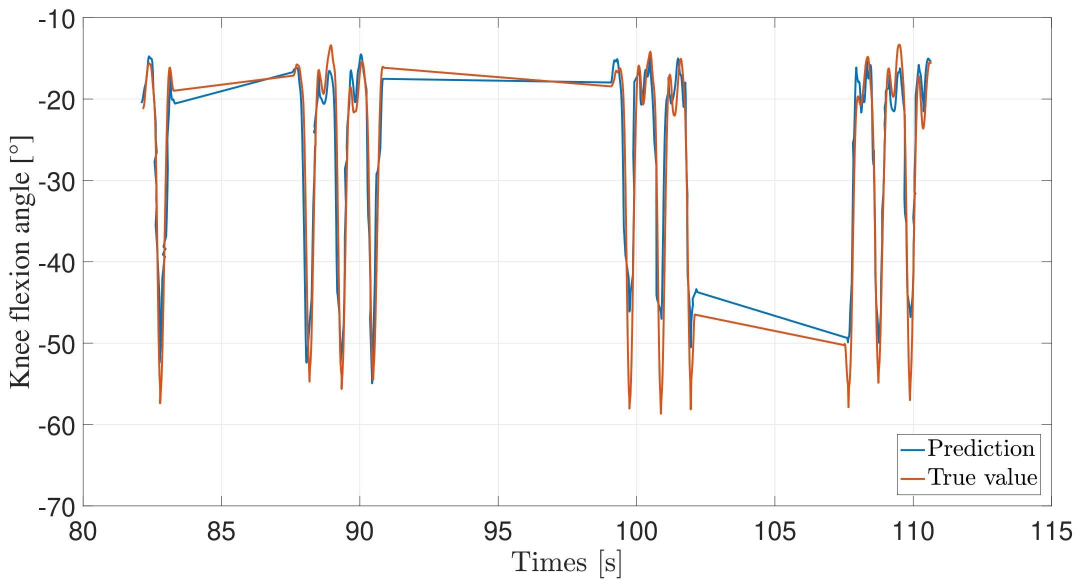

Figure 3.19: Predicted knee flexion from IMUs as compared to knee angles from OMC for overground walking (Patient 3).

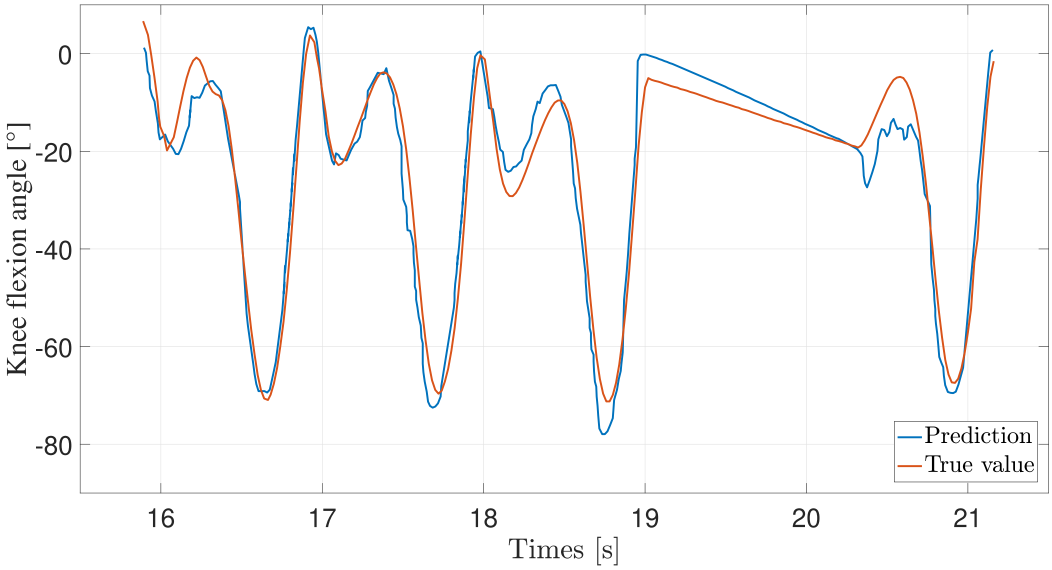

Figure 3.20: Predicted knee flexion from IMUs as compared to knee angles from OMC for overground walking (Healthy individual H11).

3.3 Generalised Models

Three generalised models were generated: the first one for treadmill walking, one for patient overground walking and the last one for the whole cohort of overground walking. They were trained on data from a group of individuals walking at their own self-selected speed. These models were tested by performing a Leave-One-Out analysis. Generalised models are expected to have less prediction accuracy compared to the personalised model, particularly given the relatively small datasets that is currently used to train the model. Moreover, generalised treadmill model is anticipated working better than the generalised overground ones, due to the increased sample size and homogeneity and provided us with a proof-of-concept for what a generalised model might look like with a greater dataset.

For the treadmill generalised model, the window size has been reduced from $\pm$ 1 second to 0.25 second. Indeed, the sample size was too large to be handled by the computer with the largest window size. As expected and listed in Table 3.6, it shows greater errors than the personalised models, with a mean RMSE of 6.68° and a mean MAE of 2.32° (against 4.05° and 2.18° for the mean RMSE and mean MAE of the treadmill personalised models). However, the difference observed between the MAE is much smaller than the one between the RMSE. It means that despite having quite a similar difference between the OMC-derived and IMU-derived knee kinematics, the generalised model presents larger errors than the personalised model, which affect its performance. Figure 3.21 illustrates an example of predicted knee flexion angle from IMUs (blue curve) as compared to knee angles from OMC (orange curve) for an individual (P08).

Treadmill walking			
ID	**RMSE [°]**	**MAE [°]**	**R2**
P07	7.19	2.44	0.97
P08	6.42	2.35	0.98
P09	9.30	2.91	0.97
P10	6.78	2.45	0.98
P11	4.88	1.89	0.98
P12	5.63	2.14	0.96
P13	6.60	2.14	0.94
P15	10.61	3.00	0.93
P16	4.99	2.00	0.97
P17	4.42	1.84	0.98
Average	6.68	2.32	0.97
S.D.	1.97	0.40	0.02

Table 3.6: Leave-One-Out analysis of a generalised model. Parameters: Top 100 features (selected using tsFresh) and a window of 0.25 second.

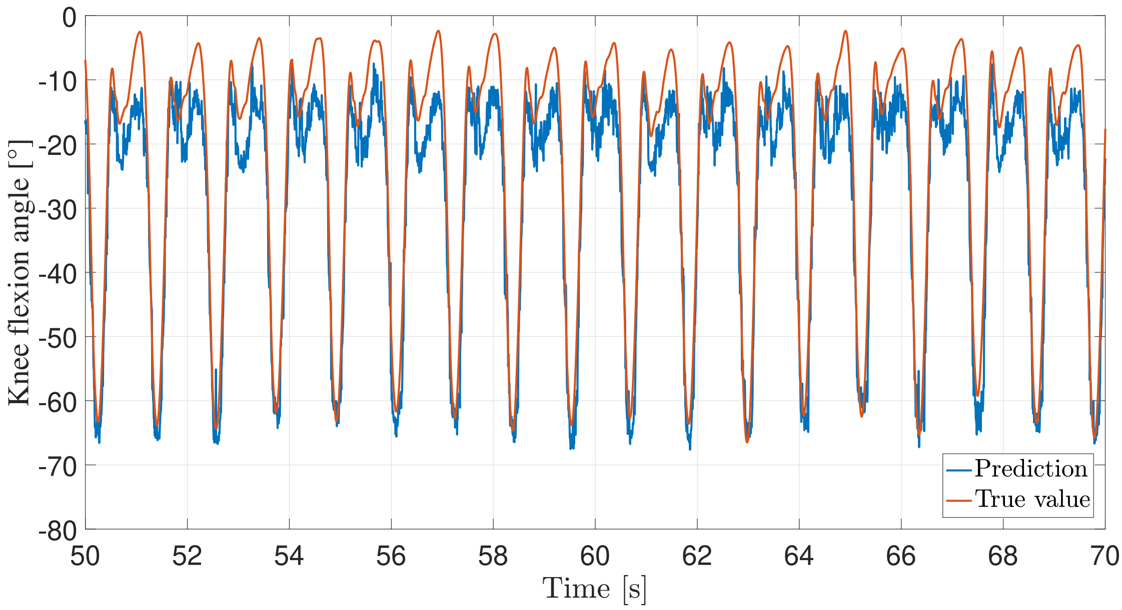

Figure 3.21: Example of predicted knee flexion from IMUs as compared to knee angles from OMC for a generalised treadmill model (P08).

The overground generalised models were trained to use a window size of 1 seconds, like the personalised models. Indeed, the sample size and thus number of steps were much smaller than for the treadmill model. The errors as well as the correlation coefficients obtained highlight the need for more data to get an acceptable generalised model (see Tab. 3.7 (left)). RMSE and MAE even increase when the healthy overground cohort is added to the patient data, suggesting that the gait profile of the healthy and patient cohort is different enough to disturb the model (see Tab. 3.7 (right)).

Patient Overground walking			
ID	**RMSE [°]**	**MAE [°]**	**R2**
1	20.15	15.80	0.16
2	22.16	18.81	0.03
3	18.73	16.00	0.23
5	20.67	16.52	0.21
6	20.43	11.77	0.11
7	23.77	21.50	0.19
8	23.32	18.24	0.09
Average	21.32	16.95	0.14
S.D.	1.83	3.03	0.07

Patient + Healthy Overground walking			
ID	**RMSE [°]**	**MAE [°]**	**R2**
1	18.25	14.10	0.31
2	23.41	18.27	0.06
3	20.52	16.91	0.15
5	22.77	17.07	-0.28
6	23.20	17.23	-0.01
7	24.02	21.44	0.20
8	24.81	19.03	-0.20
H05	18.95	15.14	0.58
H06	21.16	16.87	0.57
H11	18.38	12.10	0.66
H13	25.26	20.76	-0.02
Average	21.88	17.18	0.18
S.D.	2.57	2.74	0.32

Table 3.7: Summary of leave-one-out analysis for overground resultant IMU patient data with IMU difference. Left: only patient cohort. Right: Patient and healthy overground cohorts.

In the sensitivity analysis, it has been proven that using the resultant of the IMU data decreased the model's performance since the information available for training was reduced. However, in that case, it was not too prejudicial. This is why we also tried to generate the overground generalised models using the raw IMU data. These results are illustrated in Table 3.8. As expected, the errors decreased and the correlation increased, but these changes are only minor. The Leave-One-Out analysis illustrates the variation in model performance across the cohort (see Figure 3.22), highlighting again the need for more data, notably to capture the variability within the overground dataset in order to obtain a generalised model. For example, the model trained with Patient 2's data had the highest correlation, as it also obtained the most number of steps to train on. In contrast, the model which contained Patient 8's data had the lowest correlation, and not surprisingly, was trained on the least number of steps.

Patient Overground walking			
ID	**RMSE [°]**	**MAE [°]**	**R2**
1	17.86	13.86	0.32
2	17.27	14.95	0.63
3	18.96	15.11	0.22
5	20.67	16.52	0.21
6	20.99	15.76	0.15
7	23.61	21.25	0.31
8	22.26	16.96	0.12
Average	20.23	16.35	0.28
S.D.	2.32	2.10	0.17

Table 3.8: Summary of Leave-One-Out analysis for overground resultant IMU patient data with IMU difference.

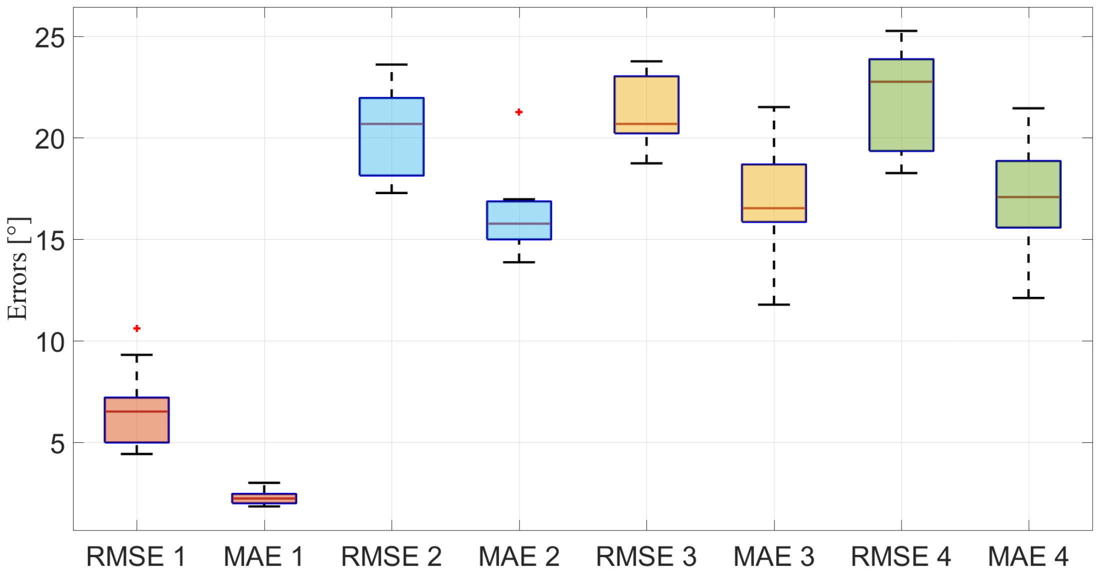

Figure 3.22: Errors (RMSE and MAE) obtained for the Leave-One-Out analysis of the generalised models. 1: Treadmill. 2: Overground Patient raw data. 3: Overground Patient resultant data. 4: Overground Patient+Healthy data.

3.4 Summary

This machine learning pipeline uses *tsFresh*, a time series feature extraction tool to automatically define features in the IMU data to predict knee kinematics during walking gait. We identified that 100 features from the *tsFresh* trained on a sample of at least 10 seconds of walking data (preferably 50 seconds) was capable of reconstructing knee kinematics with errors less than 5°. For a potential clinical application, this as a reasonable workflow, given that a personalised model can be tuned at a clinical visit, where 'ground truth' kinematics can be determined using optical motion capture. When this model was trained on the overground dataset, both patient and healthy participants had only 10 to 30 steps at different speeds with varying gait profiles. The number of steps per person was not enough to capture the variability within the dataset. This was demonstrated by the large errors seen in the generalised patient model (an average of 17° for the MAE). The Leave-One-Out analysis illustrates that when there is enough data to train the model, it is capable of reasonable predictions. Indeed, it was capable of predicting knee flexion kinematics to within 2-3° (MAE) for the treadmill dataset.

In view of the lack of data for the preoperative gaits, the comparison before and after knee replacement to assess the rehabilitation was not possible. This will be investigated as soon as new patients data will be recorded. All in all, the approach followed works very well for personalised models but is more challenging to generate generalised model.

Chapter 4

Patients Follow-up Analysis

The machine learning algorithm has been proven to predict accurately knee flexion when enough data was available to train the model. However, in view of the lack of usable information for patients' preoperative gait sessions, the evolution of outcome following joint surgery could not be assessed. Moreover, this objective measurement of their physical activity could not be contrasted with subjective PROMs accumulated through the clinical follow-up. Conversely, they can be studied and compared to other PBOMs such as the 6-min walk test, or direct monitoring of performance, like the range of motion. Thus, this chapter aims at analysing and contrasting all data collected during the patient's follow-up. All of this information can be found in Appendix A.3.

4.1 General Tendency

Firstly, the general tendency is investigated in order to examine how a typical rehabilitation following knee arthroplasty goes. The evolution of the different metrics recorded at each clinical appointment is analysed. An enhancement is expected to appear between the preoperative gait, before the surgery, and the postoperative gait, at the end of the rehabilitation process.

In terms of Oxford Knee Score (see Appendix C), the mean tendency presented in Figure 4.1 reports an improving OKS over the recovery period, with two slight drops at recovery Week 2 (just after undergoing the surgery) and 7. The first one could be attributed to patients coming off strong medication. There is a moderate standard deviation around the mean of 7–8 units on average. Subjects indicated an intermediate score of 25.5, which increased to 39.75, meaning almost excellent at the end of the rehabilitation. As shown in Figure 4.2 (left), the amount of patients reporting a poor knee function ([0–19]) decreased between the preoperative and postoperative gait until reaching zero at the $\approx$ 4 months later check-up. At Week 2, the dominant proportions of patients presented a moderate function ([20–29]), which improved to good ([30–39]) at the postoperative gait and to excellent ([40–48]) $\approx$ 4 months later. This improvement can also be observed in Figure 4.2 (right), with the presence of a small drop at Week 2. One should notice that the first two boxplots are much larger than the last two. It means that subjects reported quite different scores at the beginning, whereas they all reached a better knee function and the deviation is thus less pronounced at the end of the recovery period.

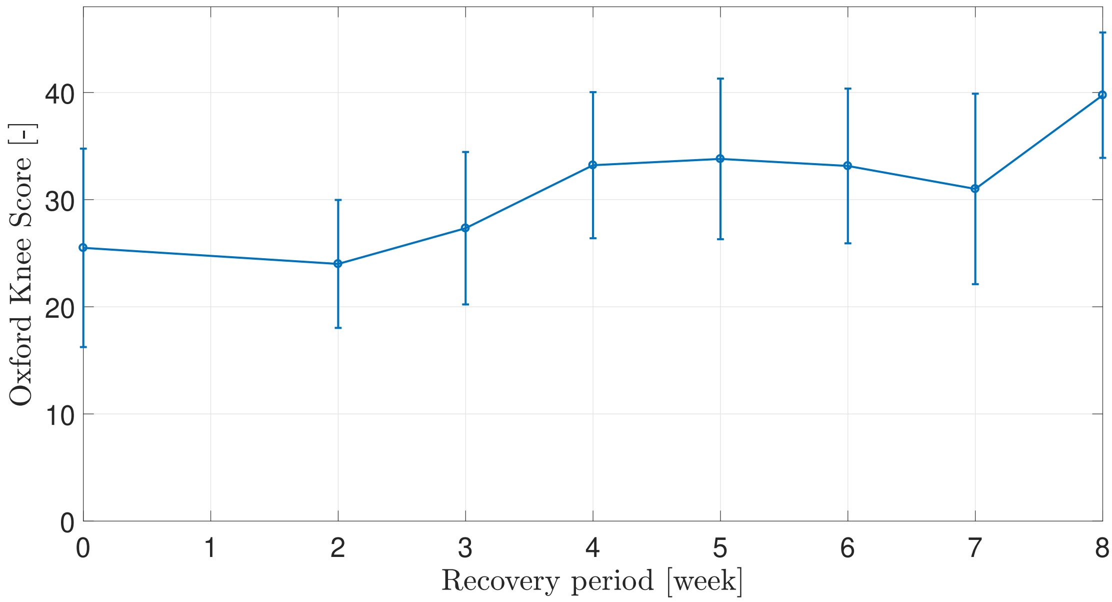

Figure 4.1: Mean Oxford Knee Score metric and its standard deviation over the average recovery period.

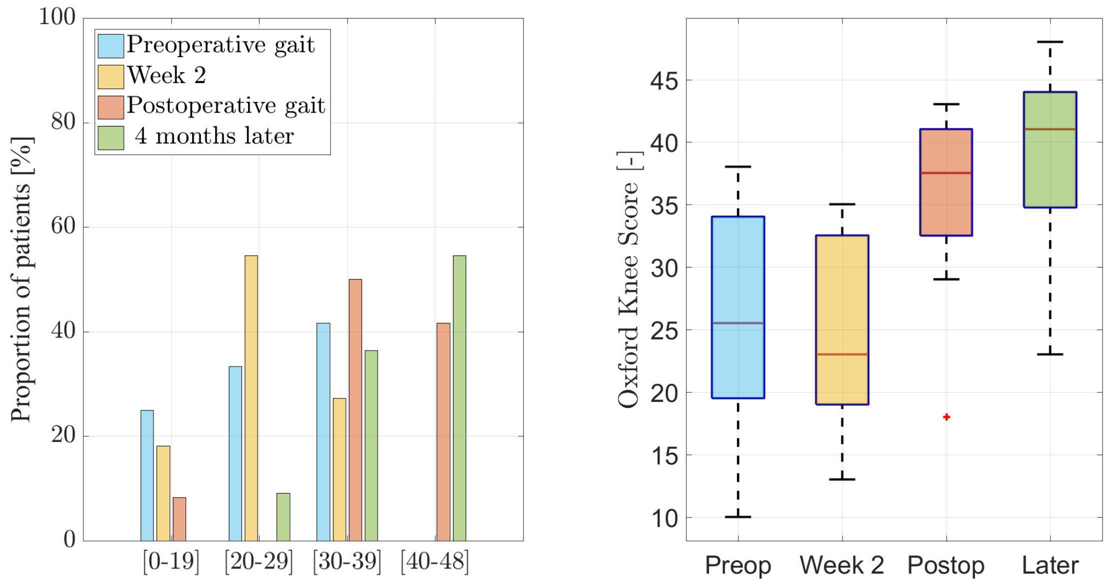

Figure 4.2: Left: Proportion of patients belonging to each OKS group (from poor grading [0–19] to excellent grading [40–48]) at each time point of the recovery period. Right: Boxplots representing the evolution of OKS over the time points of the recovery period.

The same tendency is observed for the Forgotten Joint Score (see Appendix E), in Figure 4.3 (right). Note that, unlike OKS, lower score means better results. By contrast with OKS, patients seem more in agreement at the beginning of the recovery period, with a bad score, than at the end. Almost all subjects were constantly aware of their knee joint at Week 2. At the postoperative gait, there is still a large proportion of patients that reported a bad score. Even $\approx$ 4 months later, while some subjects seem to be less aware of the artificial joint, some keep reporting a bad score.

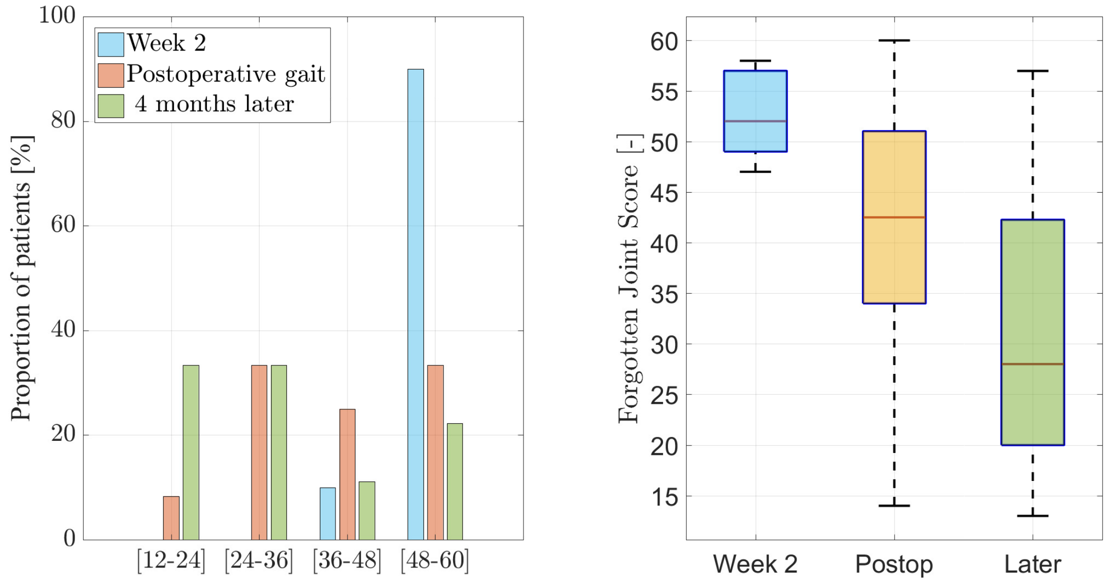

Figure 4.3: Left: Proportion of patients belonging to each FJS group (from poor grading [48–60] to excellent grading [12–24]) at each time point of the recovery period. Right: Boxplots representing the evolution of FJS over the time points of the recovery period.

In terms of range of motion (ROM), included in the direct monitoring of patients' performance category, there is also a clear improvement between the first and last physiotherapy sessions, as illustrated in Figure 4.4 (left).

The last PROMs used in this study is the EQ-5D (see Appendix D). It comprises a visual analogue scale (VAS) ranging from 0 (worst) to 100 (best), representing the patient's health at the day of this test. This measure is completely subjective and reflects the patient's own judgement, which is probably why there is a moderate deviation, especially before the surgery (see Fig. 4.4 and Fig. 4.5). Indeed, while a patient reported a score of 100, another only stated 55. Overall, Figure 4.5 shows that the mean increases slightly over the recovery period (from 78.75 to 93.5).

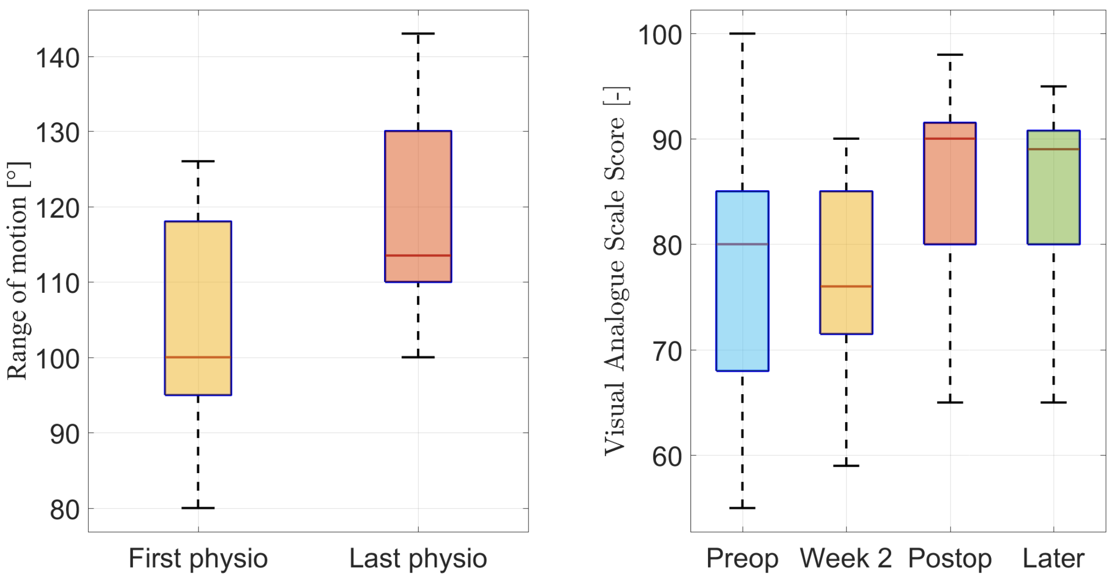

Figure 4.4: Left: Boxplots representing the evolution of ROM between the first and last physiotherapy sessions. Right: Boxplots representing the evolution of VAS score over the time points of the recovery period.

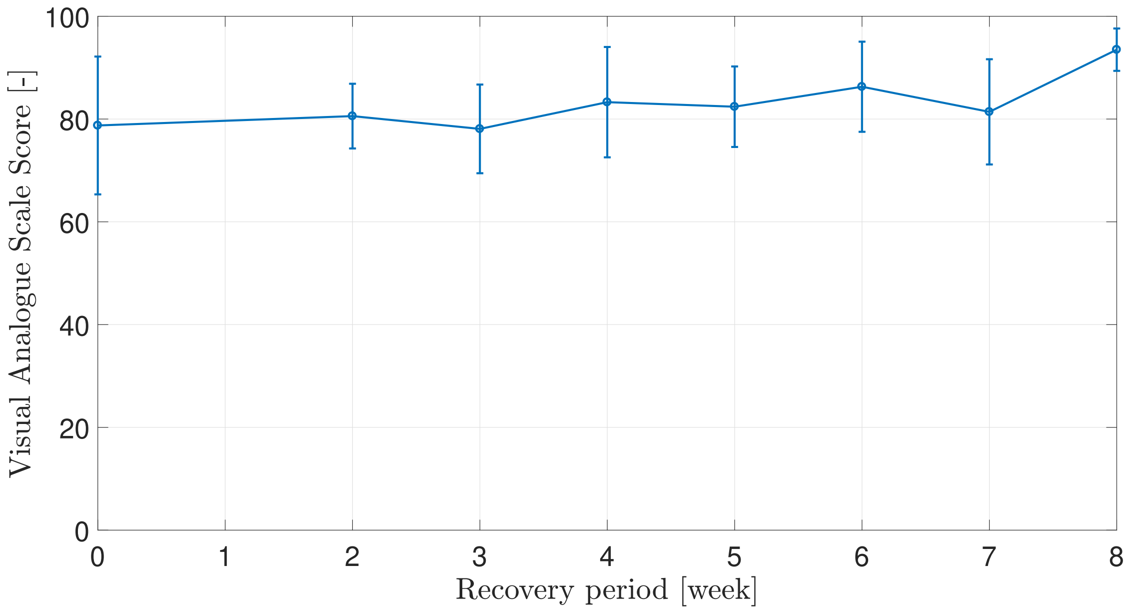

Figure 4.5: Mean Visual Analogue Scale score and its standard deviation over the average recovery period.

As reported in Figure 4.6 (left), on average, the VAS score improved during the rehabilitation. At the preoperative gait, the predominant proportion of patients considered to have a good health at this day ([80–90]). However, at the beginning of the physiotherapy sessions (Week 2), a decrease can be observed, with the largest proportion in the [70–80] grading. At the end of their clinical follow-up, despite reporting the best range ([90–100]) for more than half of the patients at the postoperative gait, the majority of them are divided between the two ranges [80–90] and [90–100] at the $\approx$ 4-month check-up. This could be explained by the fact that this meeting coincides with the first weeks of confinement due to the covid-19 pandemic. Therefore, the overall health of patients could be impacted by the anxiety generated.

In Figure 4.6 (right), 4.7 and 4.8, the spreading of patients according to their answer to the different themes is illustrated. In general, for all of them, an improvement is observed over the recovery period. For the mobility, on average, patients reported moderate problems before and just following the surgery. At the end of their follow-up, they were divided between two answers: no problem or slight problems. Most patients did not present any problems washing or dressing during the whole rehabilitation process. For their usual activities, again severe, moderate and slight problems were reported at the beginning (preoperative gait and Week 2), whereas patients had almost no trouble at the end. More than 50% of patients complained of moderate pain before the surgery and at Week 2. This discomfort disappeared for most of them at the end of the six weeks. Finally, the great majority of patients did not report any anxiety or depression whenever. This is not in agreement with our previous hypothesis that stated that the Covid-19 pandemic could explain the slight decrease in the VAS score between the postoperative gait and the $\approx$ 4-month check-up. However, one should notice that usually, people do not like to express their level of anxiety or depression to strangers.

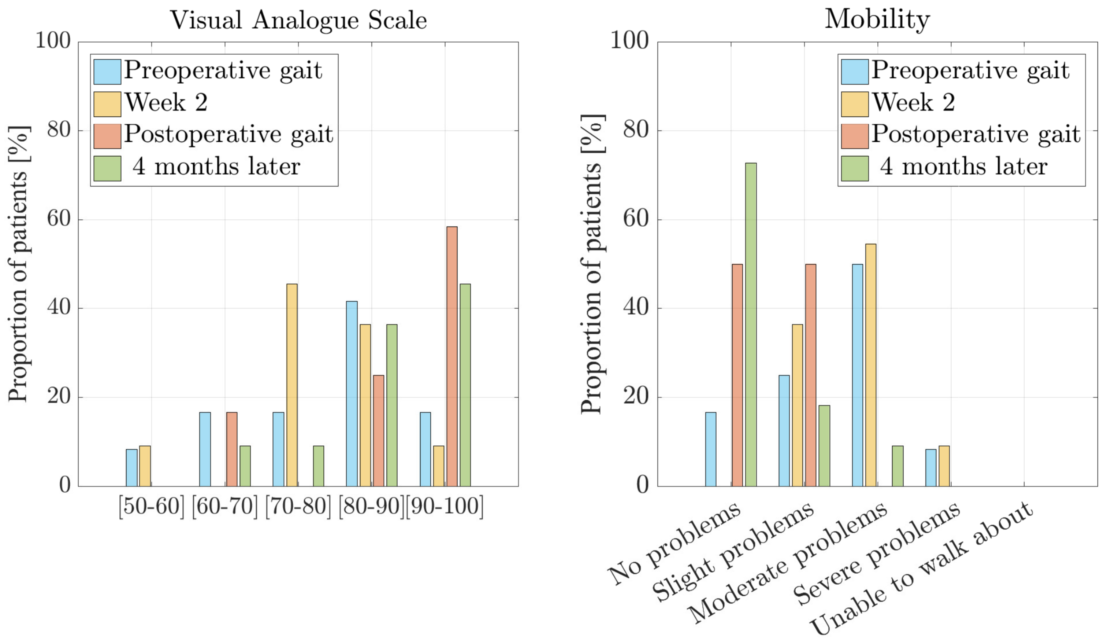

Figure 4.6: Left: Proportion of patients belonging to each score at each time point of the recovery period. Right: Proportions of patients corresponding to each proposition at each time point of the recovery period for the theme 'mobility'.

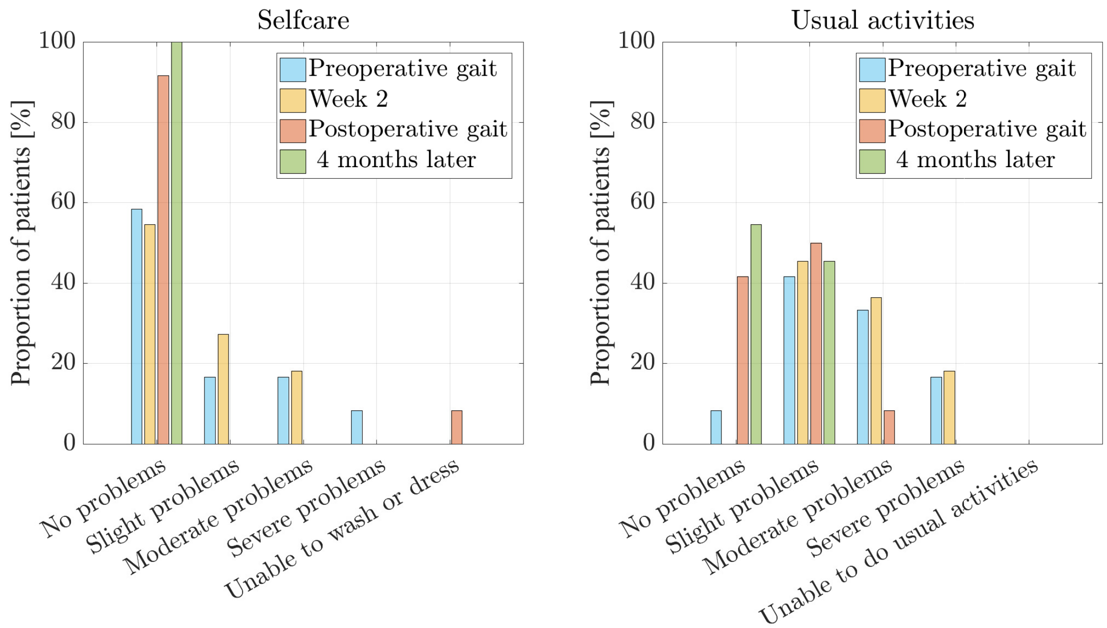

Figure 4.7: Left: Proportions of patients corresponding to each proposition at each time point of the recovery period for the theme 'selfcare'. Right: Proportions of patients corresponding to each proposition at each time point of the recovery period for the theme 'usual activities'.

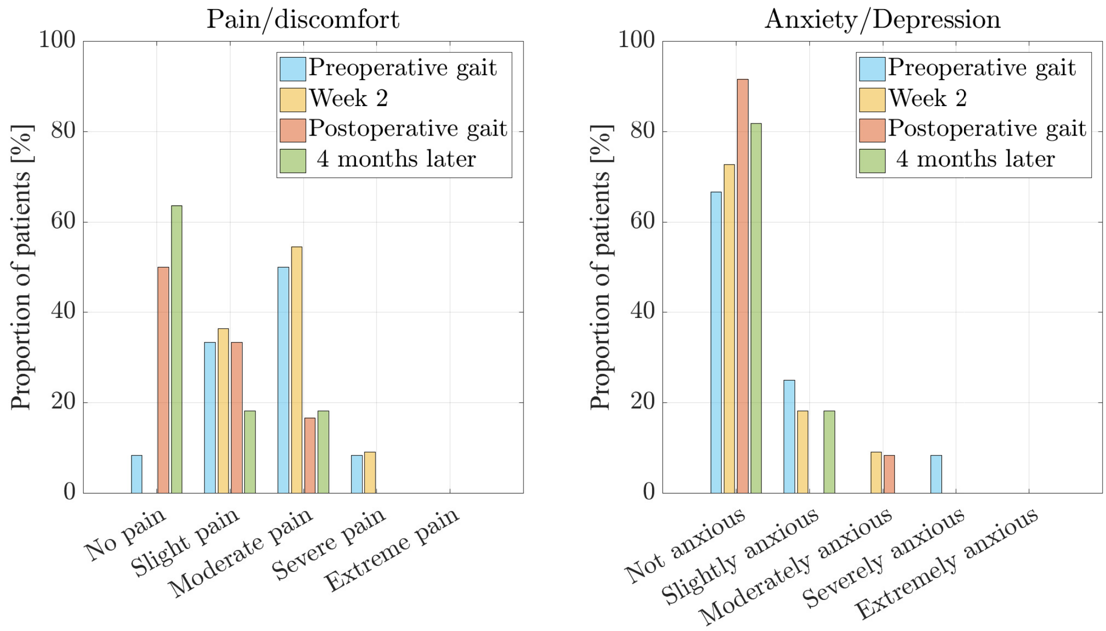

Figure 4.8: Left: Proportions of patients corresponding to each proposition at each time point of the recovery period for the theme 'pain/discomfort'. Right: Proportions of patients corresponding to each proposition at each time point of the recovery period for the theme 'anxiety/depression'.

The 6-min walk test achieved at the two gait meetings also shows an improved average for the speed (see Fig. 4.9 and Tab. 4.1). A paired t-test was performed to analyse the difference observed and assess its significance. The null hypothesis tested is that the true mean difference is zero, whereas the alternative hypothesis assumes that it is not equal to zero. The test statistic, denoted t, is calculated as follows:

$$t = \frac{\bar{d} - \mu_d}{s_d/\sqrt{n}} \tag{4.1}$$

where $\bar{d}$ is the sample mean of differences, s_d is the sample standard deviation of differences, $\mu_d = 0$ represents the null hypothesis and n is the sample size. One should notice that the standard error of $\bar{d}$ is $\frac{s_d}{\sqrt{n}}$. The pair of data containing one non-recorded speed were removed from the sample, which leaves $n = 10$. Therefore,

$$t = \frac{-0.12 - 0}{0.25/\sqrt{9}} = -1.44 \tag{4.2}$$

The calculated t value (in absolute value) is then compared to the critical one with the degrees of freedom $df = n - 1 = 8$ from the t distribution table for a confidence level chosen to 95%. Since $t = |-1.44| < 2.306 = t_{\text{crit}}$, the null hypothesis is not rejected, meaning that the means are not significantly different. The probability to observe this test statistic under the null hypothesis is $p = 0.19 = 19\%$ which is definitely not lower than 5%. Therefore, the improved average speed for the postoperative test is probably just due to luck [109] [110] [111].

ID	Pre 6-min walk [m/s]	Post 6-min walk [m/s]	Difference d
1	0.94	1.10	-0.16
2	1.08	1.46	-0.38
3	1.36	1.28	0.08
4	n/a	1.50	n/a
5	1.10	1.38	-0.28
6	0.86	0.58	0.28
7	0.88	0.72	0.16
8	1.07	1.56	-0.49
9	n/a	0.49	n/a
10	0.93	1.11	-0.18
12	0.63	0.74	-0.11
13	0.69	Covid	n/a
Average	0.95	1.08	-0.12
S.D.	0.21	0.39	0.25

Table 4.1: Speed recorded during the 6-min walk test at the preoperative and postoperative gaits.

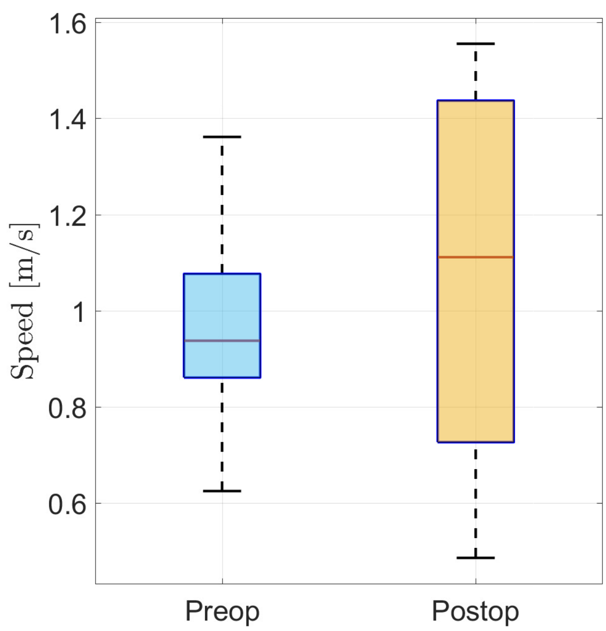

Figure 4.9: Boxplots representing the evolution of the speed during the 6-min walk test between the preoperative and postoperative gaits.

4.2 Individual

In this second part, patient rehabilitation has been looked at individually. As explained in Chapter 1 Section 1.3.6, there are some discrepancies between the different metrics used to assess the recovery process. Thus, the goal in this section is to detect if some measures provide contrasting information or if they are in agreement.

Figure 4.10 represents each individual evolution of the OKS along the recovery period. All patients have reported an increase between preoperative and postoperative gaits, approximately six weeks later (sometimes more if they required more physiotherapy sessions), thus it should not be a concern for the clinicians. However, this growth has been classified into three responses. Patient 13 reported only a slight increase with less than 10% change between the initial and last score measured (orange curve). Five patients presented a moderate growth with a change comprised between 10 and 30% (blue curve). The last six patients reported a significant increase where the difference exceeded 30% (green curve). One should notice that Patient 13 which presented the least important increase, reported a relatively high score to begin with (38, good function). Therefore, it is understandable that he could not improve it as much as patients who started with a moderate or poor knee function.

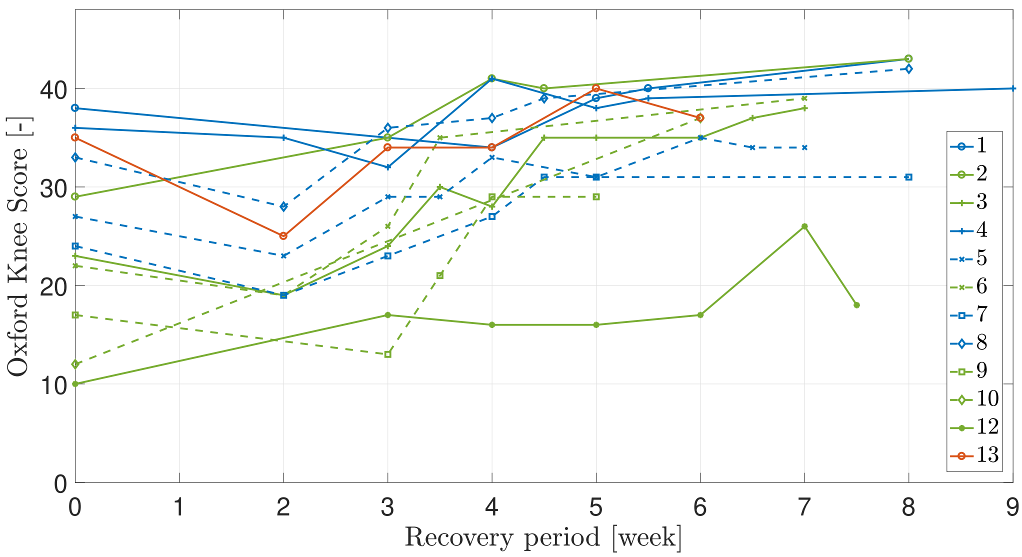

Figure 4.10: Individual Oxford Knee Score metric curves over the average recovery period.

In terms of range of motion, which was measured at each physiotherapy sessions, again, all patients reported an improvement between the first and last meetings. Moreover, except for Patient 3, they all increased their range of motion at each session irrevocably. Depending on their increase, they were grouped into three categories. Patient 2 only reported an increase of less than 10%. Seven patients presented an improvement ranging between 10 and 20%. Finally, Patients 5 and 9 reported an increase of more than 20%. One should notice that Patient 2, who reported the least increase, already reported a high ROM at the beginning (125°). Consequently, as for Patient 13 for OKS, it is understandable that he could not improve it as much as others. Patients 10 and 13 were not represented in this graph. Indeed, Patient 10 missed the follow-up in physiotherapy and Patient 13 finished at the moment of the lockdown so we could not collect his data in time (see notes in Tab. A.1 from Appendix A).

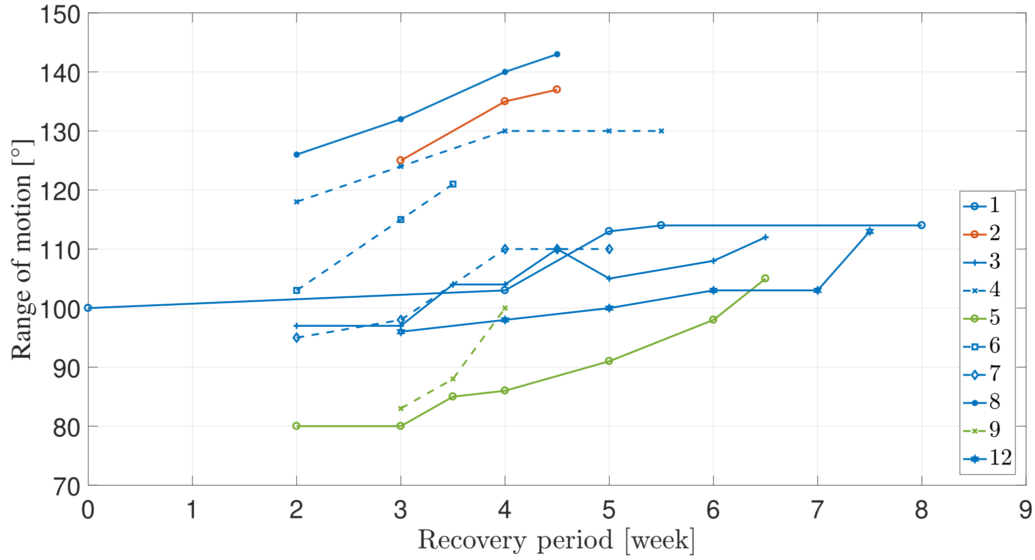

Figure 4.11: Individual range of motion curves over the average recovery period.

Unlike the two previous measurements, VAS score was not improved for all patients between the preoperative and postoperative gaits (see Fig. 4.12). Indeed, three patients reported a moderate decrease (until 10% of decrease). Four patients presented consistent score between the two time points (between less than 5% decrease and less than 5% increase). Finally, the last five patients improved their score more than 5%. Specifically, Patients 7 and 8 reported an increase of 64 and 51% respectively. However, these two are also the ones which started with the lowest score, unlike the three patients who reported a drop and began with the highest score.

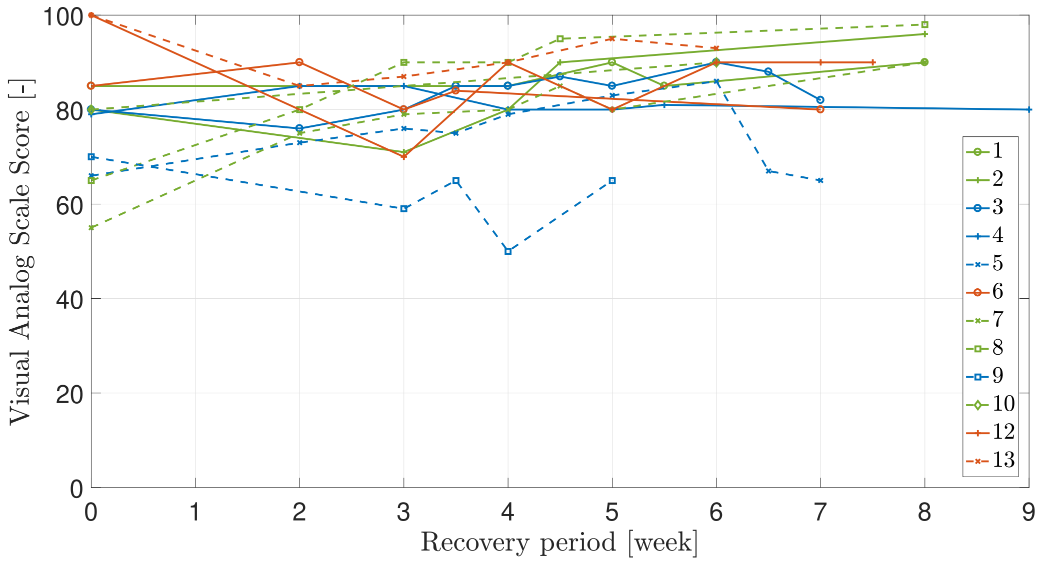

Figure 4.12: Individual Visual Analogue Scale Score metric curves over the average recovery period.

Patients' range of motion for knee flexion was also evaluated pre- and post-surgery during the gait sessions. Figure 4.13 illustrates the knee angle of each patient's affected side at the preoperative and postoperative gaits. Note that only the patients with no issues in the two OMC files were represented. Moreover, the experimental curves have been shifted and mirrored about the x-axis to start around zero and count the joint angle in positive numbers. This shifting does not influence the results, it is just a change of reference to measure the angle. Patients 1, 3, 6, and 8 did not recover their ROM post-surgery compared to pre-surgery baseline, unlike Patients 5 and 7 who improved it. However, since all patients improved their ROM at the physiotherapy sessions, we should not take any hasty conclusion about these results.

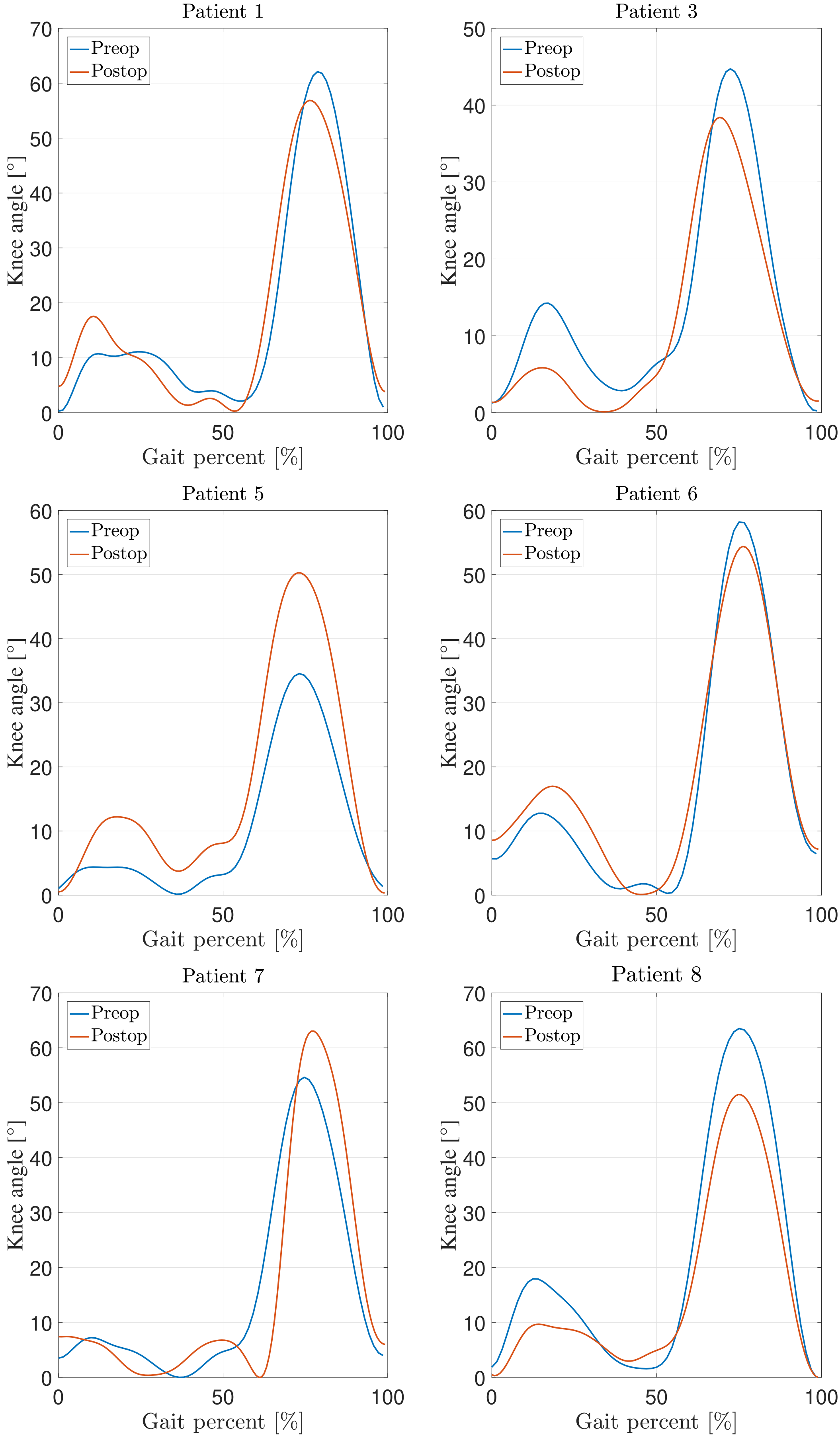

Figure 4.13: Patient pre-surgery (blue) and post-surgery (orange) knee flexion range of motion.

Pre- and post-surgery self-selected walking speeds were also computed by dividing the distance covered by the markers positioned on the sacral by the time taken to travel this distance. The beginning and end of each trial were not taken into the calculation in order to avoid any acceleration and keep a constant speed for the whole test. For each subject, for whom both Mocap files were usable, the speed has been averaged on all trials for both gait sessions and its standard deviation was computed. All this information is presented in Table 4.2. Patients' walking speed ranges from 0.76 m/s to 1.04 m/s and has a very tight standard deviation across the walking trials. In order to determine the significance of the average difference observed between the pre- and post-surgery sessions, a two-sample t-test has been used. Its principle is similar to the paired t-test at the only exception that the two-sample t-test requires independent groups for every sample [112]. The *Matlab* function `ttest2` was used for this purpose for each patient. As a result, only Patient 5 reported an increase self-selected speed post-surgery. Patients 1, 3, and 6 showed no significant change in walking speed and Patients 7 and 8 presented a reduced speed.

ID	**Preoperative Gait**		**Postoperative Gait**		**Change**
	Mean speed [m/s]	**S.D.**	**Mean speed [m/s]**	**S.D.**	
1	0.82	0.05	0.85	0.06	=
3	1.01	0.09	0.96	0.04	=
5	0.76	0.06	1.03	0.04	↗
6	0.82	0.03	0.79	0.09	=
7	0.94	0.03	0.80	0.08	↘
8	1.04	0.06	0.95	0.07	↘

Table 4.2: Mean speeds recorded during the walking trials at the preoperative and postoperative gaits and two-sample t-test significance.

Finally, from the 6-min walking test, six patients reported an increase walking speed, with a percent increase ranging from 16 to 45%. The last three patients decreased the speed from 6 to 32% (see Table 4.3).

To summarise, the percent change [%] of each metric recorded at the beginning and at the end of the recovery period are listed in Table 4.3. Three colours are used to highlight either the significance in increase (for OKS and ROM physiotherapy) or the difference between the increase, decrease and consistent measures: orange represents either the least improvement or the drop observed, blue is for the moderate increase or consistent measurements recorded, and green means significant enhancements. This table demonstrates the important contrasts among different metrics. For example, Patient 8 presented a significant decrease for its self-selected speed during the walking trials and for its ROM post-surgery at the gait sessions. However, he also reported a major increase in his speed for the 6-min walking test as well as for the score attributed to his general health. Moreover, his ROM measured in physiotherapy was also improved. It was highlighted that patients who presented very low range of motion and pain will often show major improvements after joint replacement surgery. By contrast, subjects who often have a very high range of movement and good knee function and health prior to surgery may sometimes not fully recover their pre-surgery ROM or Oxford Knee Score. It does not mean that their outcome is bad, though. In fact, one should notice that, in view of these results, no patients were identified for further review during rehabilitation.

ID	OKS	VAS	ROM physiotherapy	ROM kinematic	Mean Speed Change	6-min Walk Speed
1	13.16	5.88	10.68	-9.56	3.66	16.18
2	48.28	20	9.6	n/a	n/a	35.48
3	65.22	2.5	15.46	-16.2	-4.95	-6.12
4	11.11	1.27	10.17	n/a	n/a	n/a
5	25.93	-1.52	31.25	17.65	35.53	25.32
6	77.27	-5.88	17.48	-4.37	-3.66	-32.26
7	29.17	63.64	15.79	24.74	-14.89	-17.46
8	27.27	50.77	13.49	-22.07	-8.65	45.45
9	70.59	-7.14	20.48	n/a	n/a	n/a
10	208.33	12.5	n/a	n/a	n/a	19.40
12	80	-10	17.71	n/a	n/a	18.44
13	5.71	-7	n/a	Covid	n/a	n/a

Table 4.3: Percent change [%] obtained for each metric used between the initial and last measurements.

Chapter 5

Discussion

This concluding chapter aims at providing a link between our main findings and other literature in the field. Moreover, the limitations and issues encountered during this project are highlighted. Based on this, perspectives for future work are proposed.

5.1 Comparison with previous studies

Our machine learning algorithm is based on the idea to automatically extract time series features. In the field of human motion analysis, this type of workflow has already been exploited for other applications.

For example, Kempa-Liehr et al. [113] used `tsfresh` for the case of a user-specific running-walking classification, as well as for its generalisation to a multi-user multi-activity classification, reaching an accuracy of 92%. They also employed two ankle-worn IMUs for the first experiment, but increased this number to nine to generalise it to other activities.

Derie et al. [114] applied a similar machine learning algorithm to time continuous data generated from bilateral tri-axial accelerometers placed on the ankles in order to estimate the vertical instantaneous loading rate of runners. For the learning, the same approaches as our work were considered: a subject-independent model based, which was trained on the data of all runners but one used for the prediction, and a subject-dependent model, personalised for each individual. They also reported a better predictive performance for the subject-dependent model with a mean absolute percentage error of 6.08% against 11.09% for the subject-independent model. The reason was again the variability between participants.

Another study based on an artificial neural network was performed by Wouda et al. [115] to estimate lower-body joint angles at first and vertical ground reaction forces in a second time. Three body-worn IMUs are positioned on the pelvis and on the two lower legs and the cohort consisted in eight individuals running on an instrumented treadmill. Once again, two scenarios were evaluated: a single subject and a multiple subject. Knee flexion angle was estimated with an accuracy of less than 5° (mean RMSE) for the personalised model. By contrast, for the generalised model, despite reporting a good correlation coefficient for most participants (more than 0.9), variations in accuracy were found which may be due to their different landing patterns. Therefore, including more individuals is required to obtain a better performance, which is the same conclusion as for our results.

Gholami et al. [11] did a similar work using a convolutional neural network architecture to monitor lower extremity running kinematics with a single IMU positioned on the shoe. Just like our study, reference joint angles were measured by an optical motion capture system and two evaluation methods were used: intra-participant, i.e. personalised models, and inter-participant (i.e. generalised models). They managed to reach an accuracy (RMSE) of less than 3.5° and 6.5° in intra- and inter-participant scenarios respectively for their cohort of ten healthy subjects running on a treadmill. These results are relatively close to our treadmill dataset errors, even if we predicted walking knee flexion by using a different machine learning workflow and parameters.

Other studies about the estimation of lower joint kinematics exist but using different techniques than data-driven machine learning, such as orientation-based methods or musculoskeletal model-based approaches [80] [83]. Moreover, from all the works based on machine learning to predict knee angles, there is currently only one, to our knowledge, which concerns patients affected by osteoarthritis (De Brabandere et al. [116]). However, the focus of this study is to monitor changes during the disease progression for hip osteoarthritis patients using a single IMU from a mobile phone attached to the hip. Their algorithm is also based on feature construction using `TSFuse` Python package to estimate hip and knee angles during various exercises. Their cohort was composed of twenty participants, but a subset of only ten was used due to incorrect recording of mobile phone measurements. Two different scenarios were considered again: applying the model to seen patients, i.e. for whom some labelled data is already available, and to an unseen one. Nonetheless, even for the seen patients, they included data from all other patients, unlike us. For the unseen patient (corresponding to our generalised model), the average mean absolute percentage error (MAPE) for the hip angle was 29 and 36% for the left and right side respectively, and 32 and 48% for the left and right knee angle respectively. These results are not accurate enough for valid clinical use. However, where our weakness lies in the lack of data for the generalised overground patients model, De Brabandere and collaborators want to investigate more sensors at different locations, as well as the robustness to the position of the sensors, which they think to be the cause of their inaccurate results.

In brief, the basic idea of our algorithm has already been used in various fields of human motion capture lately, but rarely for the case of osteoarthritis patients. To our knowledge, this is the first time that it has been applied to monitor subject rehabilitation following knee replacement surgery. The principle of investigating inter- and intra-participants scenarios is not new and has been widely employed to assess the machine learning pipeline in previous studies. Personalised models have the advantage of providing more accurate results but require to record training data for every new patient. One can imagine this collection implemented in a clinic setting, prior to tracking and predicting knee flexion data outside the clinic. Generalised models, despite decreasing the performance due to the lack of personalisation and high variability among subjects, can be applied to an unseen patient for whom no labelled data is already available. Ideally, we could generate enough data across a population to create this inter-patient model that could predict accurately knee flexion angles for individuals who are not part of it.

5.2 Limitations and perspectives

For the machine learning models, data is the limiting factor on how well they can perform. This is especially true for the generalised machine learning models as already explained in Chapter 3 Section 3.3. Therefore, the continuation of this study should increase the patient dataset, especially now that the restrictions imposed by the Covid-19 pandemic are no longer valid. The increase of preoperative data will also enable to create the generalised model for this step too.

Not only is the amount of data important but the quality of the data matters. If there is a misalignment in the synchronisation or a mismatch between the IMU and OMC data then large errors will result. Many data were not usable for this reason. However, now that we are aware of this issue, more attention should be payed to this for future data collection.

Other tasks than walking has been recorded but not exploited. Indeed, it was easier to investigate cyclic movements like walking at first. Future work should also develop models for these other activities or maybe a more generic one comprising all of them, like De Brabandere et al. [116] did in their work.

Another limitation has been encountered due to the file size of the IMUs tracking at home. Indeed, the recording lasted the whole day starting at the physiotherapy session in the morning. Working remotely for most part of this project with a computer of limited resources handicapped us for handling such large files. Consequently, addressing this issue, for example by hooking up to supercomputer centres and having the data processed there, is required to obtain knee kinematics in a home setting and investigate the evolution of patients' range of motion over a longer duration.

Once all of these limitations will be met, information about the outcome collected in patients' environments will be compared to PROMs (OKS, FJS, EQ-5D) and ROM gathered at the physiotherapy sessions. As mentioned previously, subjective measures are not enough to assess postoperative physical function efficiently. For example, during this first stage of the study, some patients made comments when they were asked the questions, notably about their lack of clarity. One of them asked if we should take into account painkillers when evaluating pain since hr progressively stopped taking them, which increased his pain. For the question concerning the duration a patient can walk before feeling severe pain, another one told us that he was only doing thirty minutes by order of the doctor, so he did not know if he could do more. Moreover, there is no way to skip the question if it does not apply. This last comment was especially made for the FJS questionnaire. Finally, for the question about anxiety and depression in the EQ-5D questionnaire, patients mostly reported that they were not feeling anxious or depressed the whole period of recovery. We think that this can be explained by the fact that people usually do not like to express their concern or bad mood level to total strangers. Subjective measures are therefore not enough to ensure the successful outcome of knee arthroplasty.

Conclusion

The current clinical gold standard for routine follow-up is a minimum face-to-face consultation and collection of PROMs. The lack of quantitative information about the outcome following knee replacement surgery has suggested to monitor patients' physical activity over a long capture duration by using wearable sensors. The purpose of this study was to present a new paradigm of arthroplasty follow-up with continuous monitoring of patients in their own environment.

A workflow was presented to assess joint function following knee replacement surgery using only two ankle-worn IMUs. To this end, a machine learning algorithm based on `tsFresh` Python package has been implemented to predict knee flexion time series kinematics during walking. The idea was to use statistically significant time-series features to train the multivariate nonlinear regression. To evaluate the algorithm, estimations of joint kinematics were compared to 'ground truth' joint kinematics recorded from optical motion capture.

At this stage, a cohort of twelve patients undergoing knee arthroplasty were enrolled in this study. They participated in two gait sessions before and around six weeks after their surgery during which optical marker trajectories and acceleration and angular velocity from IMUs were recorded. After being processed and synchronised, these data were used to assess the machine learning algorithm. However, in view of the issues encountered with their quality and quantity, two other datasets from the database of the Auckland Bioengineering Institute were exploited too. One was composed of ten healthy individual walking on a treadmill, whereas the other concerned four healthy participants performing overground walking.

Two types of model were generated and evaluated: a personalised model, trained on a portion of a subject's data and predicting the remaining part, and a generalised model, trained on every individual of the cohort but one used for prediction. For person-specific models, a sensitivity analysis was performed on the treadmill dataset to select the most optimal parameters and data processing ways. A window size of 1 second and a number of features of 100 frames were chosen. Moreover, downsampling, filtering and using the resultant were proven not to be too prejudicial concerning the accuracy and advantageous in terms of robustness and processing time. We have also demonstrated that the model was robust to a slight rotation of sensors. The combination of data processing and parameters chosen, trained on a sample of at least 10 seconds of walking data (preferably more) was capable of predicting knee kinematics with more than 95% accuracy (MAE). This also holds for the treadmill generalised model. However, for the overground generalised model, a major limitation was encountered: the number of steps per person was not enough to capture the variability within the dataset. Increasing the patient dataset will definitely improve this model performance. Therefore, the models presented here give us confidence that knee flexion kinematics can be obtained from only two IMU data, which is an exciting opportunity to monitor knee flexion kinematics in real-world settings.

Other metrics were also collected at the clinical follow-up such as PROMs and ROM. Broadly speaking, all of them (OKS, FJS, ROM, EQ-5D) presented an improvement over the recovery period. By contrast, when inspecting each subject individually, a contrast between the different metrics was observed. Nonetheless, further investigation is required with more subjects. These results will also be compared to the objective biomechanical variable collected in patients' own environment that our machine learning algorithm now allows us to obtain. Ultimately, this may help clinicians to identify potential complications during recovery and provide the opportunity for early intervention.

Our method enables to obtain high quality functionally data, to follow and guide patients through their rehabilitation journey. It practically allows the monitoring of patients real world mobility and activity, without the need for time-consuming and subjective assessments. All of this could be undertaken remotely, helping out in a post-Covid 19 where we are reminded of the possibility that face-to-face monitoring may be jeopardised by future pandemics.

Appendices

Appendix A

Patients data

A.1 General information

ID	Gender	Ethnicity	Date of birth	Age	Height	Weight	BMI[1]	Notes
1	M	NZ	06/06/1941	78.51	170	68	23.53	
2	F	NZ	14/07/1949	70.40	168.5	84.5	29.76	
3	F	NZ	31/10/1949	70.12	151.5	53.6	23.35	
4	M	NZ	22/07/1954	65.36	167.5	74.9	26.7	
5	F	NZ	25/11/1960	59.04	168	100	35.43	
6	M	NZ	20/10/1946	73.16	179	105.2	32.83	
7	M	NZ	16/02/1942	77.81	168.5	73.5	25.89	
8	M	NZ	23/04/1957	62.62	175	92	30.04	
9	F	Indian	04/04/1957	62.71	171	133	45.48	
10	F	NZ	23/02/1953	66.96	161.4	80.4	30.86	Missed follow up from weeks 2 to 6 in physiotherapy
12	F	Indian	22/02/1959	60.98	152	80.9	35.02	
13	M	NZ	10/10/1949	70.37	172.5	86	28.9	Did not perform 6 week gait lab due to coronavirus

[1] Body Mass Index

Table A.1: General information about the participants.

One should notice that there is no Patient 11. Indeed, despite giving his consent at the information session, he was supposed to be accompanied by his daughter for each session for translation purpose. However, she was not available so he was not able to follow the timeline.

A.2 Surgical information

ID	Surgery date	Surgeon	Comorbidities	Pre-operative pain medications	Previous surgery	Planned surgery	Side	Blood loss	Time in hospital
1	20/11/2019	Ferguson	Heart attach, Altered Balance, Dyslipidaemia	Celocoxab 1/day in the morning	n/a	Total	Left	100	4
2	21/11/2019	Monk	Hyperlipidaemia	Panacodeine (1-2 tabs po prn/daily); voltaren LA 1/week	None	Uni	Right	n/a	2
3	28/11/2019	Monk	HTN	Not discussed	Meniscectomy 80's	Total	Right	n/a	2
4	14/11/2019	Monk	Dyslipidaemia, HTN	Voltaren 25mg po prn 1/month	Arthroscopy 2015 (planning to do other knee)	Uni	Left	150	1
5	27/11/2019	Monk	None	Mr. Monk prescribes - take on big days 2/week	None	Uni	Right	250	5
6	29/11/2019	Munro	Asthma, HTN	Not discussed	Ankle reconstruction (previous TKJR other side)	Total	Right	n/a	4
7	19/11/2019	Ferguson	Dyslipidaemia, HTN, CKD, Fractured back 2012	Occasionally uses (probably paracetamol)	Arthroscopy 15-20 years ago	Total	Right	n/a	3
8	19/11/2019	Ferguson	None	None	None (other knee recent TJR 20/03/19)	Uni	Right	n/a	2
9	3/12/2019	Munro	HTN, Hyperthyroid, Dyslipidaemia, Asthma	Not discussed	(TKJR other side)	Total	Left	0	6
10	23/01/2020	Monk	Previous DVT, asthma, HTN		Previous bilat 1st MTPJ arthrodesis	Uni	Right	n/a	n/a
12	31/01/2020	Monk	T2DM, HTN	Paracetamol and Tramadol	None	Total	Left	25	n/a
13	14/02/2020	Monk	Previous AVR, hypothryoidism, HPT	Celocoxab 1/day in the morning	None	Total	Left	500	3

Table A.2: Surgical information about the participants.

A.3 Information from Gait lab and physiotherapy sessions

ID	Week	OKS	EQ-5D Mobility	EQ-5D Selfcare	EQ-5D Activities	EQ-5D Pain	EQ-5D Anxiety	VAS	FJS	ROM	6-min walk Distance [m]	Downhill treadmill Speed reached [km/h]
1	0	38	1	1	2	3	2	85	/	100	340	n/a
	4	34	2	1	2	2	1	85	49	103		
	5	39	1	1	1	2	1	90	/	113		
	5.5	40	1	1	1	1	1	85	/	114		
	8	43	1	1	1	1	1	90	34	114	395	3.7
	4 months later	44	1	1	1	1	2	90	31	/		
2	0	29	3	1	2	2	1	80	/	/	387.5	n/a
	3	35	2	1	2	2	1	71	59	125		
	4	41	2	1	1	2	1	80	/	135		
	4.5	40	3	1	1	2	1	90	/	137		
	8	43	1	1	1	1	1	96	14	/	525	5
	4 months later	48	1	1	1	1	1	95	13	/		
3	0	23	3	2	3	3	1	80	/	/	490	n/a
	2	19	3	2	3	3	1	76	57	97		
	3	24	3	1	3	3	1	80	/	97		
	3.5	30	3	1	3	1	1	85	/	104		
	4	28	2	1	3	3	1	85	/	104		
	4.5	35	2	1	2	2	1	87	/	110		
	5	35	2	1	2	3	1	85	/	105		
	6	35	2	1	1	2	1	90	48	108		
	6.5	37	1	1	2	2	1	88	45	112		
	7	38	2	1	2	2	1	82	45	/	460	5
	4 months later	37	2	1	1	2	1	91	38	/		

ID	Week	OKS	EQ-5D Mobility	EQ-5D Selfcare	EQ-5D Activities	EQ-5D Pain	EQ-5D Anxiety	VAS	FJS	ROM	6-min walk Distance [m]	Downhill treadmill Speed reached [km/h]
4	0	36	2	1	2	3	1	79	/	/	n/a	n/a
	2	35	2	1	2	3	1	85	52	118		
	3	32	2	1	2	3	1	85	/	124		
	4	41	2	1	2	3	1	80	/	130		
	5	38	2	1	2	2	1	80	/	130		
	5.5	39	2	1	2	2	1	81	/	130		
	9	40	2	1	2	2	1	80	33	/	540	5
	4 months later	41	1	1	2	1	1	85	28	/		
5	0	27	3	2	4	2	1	66	/	/	395	n/a
	2	23	3	2	3	3	1	73	58	80		
	3	29	2	1	2	2	1	76	/	80		
	3.5	29	2	1	4	2	1	75	/	85		
	4	33	2	1	3	2	1	79	/	86		
	5	31	3	1	3	3	1	83	/	91		
	6	35	2	1	2	2	1	86	50	98		
	6.5	34	1	1	2	2	1	67	46	105		
	7	34	1	1	3	1	1	65	/	/	495	4.4
	4 months later	38	1	1	2	1	1	89	/	/		
6	0	22	3	1	2	3	1	85	/	/	310	n/a
	2	19	3	1	2	3	1	90	58	103		
	3	26	2	1	1	2	1	80	/	115		
	3.5	35	2	1	1	3	1	84	/	121		
	7	39	2	1	1	2	1	80	35	/	210	1.2
	4 months later	42	1	1	1	1	1	79	17	/		
7	0	24	3	3	3	2	2	55	/	/	315	n/a
	2	19	3	1	4	3	2	75	53	95		
	3	23	2	1	3	3	1	79	/	98		
	4	27	3	1	2	2	1	80	/	110		
	4.5	31	2	1	3	2	1	85	/	110		
	5	31	2	1	2	2	1	80	55	110		
	8	31	2	1	2	2	1	90	50	/	260	4.3
	4 months later	34	1	1	2	2	1	90	55	/		

ID	Week	OKS	EQ-5D Mobility	Selfcare	Activities	Pain	Anxiety	VAS	FJS	ROM	6-min walk Distance [m]	Downhill treadmill Speed reached [km/h]
8	0	33	2	1	2	3	2	65	/	/	385	n/a
	2	28	2	1	2	2	1	80	47	126		
	3	36	2	1	2	2	1	90	/	132		
	4	37	1	1	2	2	1	90	/	140		
	4.5	39	2	1	1	2	1	95	/	143		
	8	42	1	1	2	1	1	98	40	/	560	5
	4 months later	46	1	1	1	1	1	93	21	/		
9	0	17	3	3	3	3	4	70	/	/	n/a	n/a
	3	13	3	2	4	4	2	59	/	83		
	3.5	21	2	2	3	3	2	65	/	88		
	4	29	2	2	2	2	3	50	/	100		
	5	29	2	5	2	3	3	65	54	/		
	4 months later	29	2	1	2	3	2	65	/	/	175	1.2
10	0	12	2	1	3	2	1	80	/	/	335	n/a
	6	37	1	1	1	1	1	90	34	/	400	2
	4 months later	44	1	1	1	1	1	80	25	/		
12	0	10	4	4	4	4	3	100	/	/	225	n/a
	3	17	4	3	3	3	1	70	52	96		
	4	16	3	1	2	3	1	90	/	98		
	5	16	3	3	3	3	1	80	/	100		
	6	17	3	1	3	3	1	90	59	103		
	7	26	3	1	1	2	1	90	58	103		
	7.5	18	2	1	1	3	1	90	60	113	266.5	4
	13	23	3	1	2	3	1	80	57	/		
13	0	35	1	1	1	1	1	100	/	/	250	3
	2	25	3	3	3	2	1	85	48			
	3	34	3	1	1	2	1	87	/			
	4	34	1	1	1	2	1	90	/			
	5	40	2	1	1	1	1	95	/			
	6	37	1	1	2	1	1	93	52		Covid	Covid

Table A.3: Information from Gait lab and physiotherapy sessions about the participants.

Appendix B

Consent form

CONSENT FORM

Evaluating outcomes following knee arthroplasty using IMU

I have read the Participant Information Sheet, understood the nature of the study and why I have been selected. I have had an opportunity to discuss my concerns (if any) with the study investigators. I am satisfied with the answers I have been given. I understand that taking part in this study is voluntary (my choice), that I am free to withdraw my participation at any time and to withdraw any data traceable to me up to 13 months from signing this consent form, and that I may request a copy of the results from my own data.

I understand that my movements will be recorded by an Inertial Measurement Unit (IMU) sensor while performing daily tasks, which will be used to assist with identifying my average loading during my recovery period.

I understand that the data will be kept for a period of six (06) years after which it will be destroyed.

I understand that my participation in this study is confidential and that no material, which could identify me, will be used in any reports or presentations. The data will be stored in a Seagate hard-drive and only available to the study investigators, whom are mentioned above. Volunteer's names will not be used to identify any data.

I understand that data acquired will be primarily used to create population models of loadings in the lower limbs and conduct biomechanical analyses, and that the de-identified group statistics will be used in publications, such as journal papers, posters and conference presentations. I understand my averaged data from this study may be compared against past averaged de-identified patient data who are aged matched and have had the same joint replacement surgery. This will help the investigators evaluate their computational model and sensor solution.

I understand that my consent to take part does not alter my legal rights or any standard care I may receive. I am assured that my choice of participation or non-participation will not affect my future healthcare, and I have been assured by my surgeon and the investigators in this study. I have had time to consider whether to participate and I know whom to contact if I have any questions regarding this study.

Name: ________________________________

Signature: _____________________________   Date: ______________

APPROVED BY THE AUCKLAND HEALTH RESEARCH ETHICS COMMITTEE ON ______________ FOR THREE (03) YEARS.
REFERENCE NUMBER ____________

Optional 3-D Gait (Walking) Lab Study at University of Auckland - Newmarket Campus (only if requested)

1/2

I agree to participate in a 3-D gait (walking) analysis study at the Newmarket Campus of the Department of Exercise Science, University of Auckland (if requested by research team)	Yes	No

I understand that I have been be asked to visit the gait lab at University of Auckland (Department of Exercise Science) Newmarket Campus. I agree to participate in a 3-D gait analysis that will will take approximately 60-90 minutes. This will involve me walking along a specialised mat with surrounding sensors/cameras, only as I am able to tolerate. I understand I may be asked to participate in 3 visits to the gait lab: before my surgery and then 2 & 6 weeks after my surgery.

I understand the research team will cover the transport costs with a petrol voucher from Greenlane Clinical Centre to Newmarket Campus, University of Auckland.

I understand I can still participate in the study if I do not want to attend this part. I also understand I can opt out of this part of the study at any point.

Name: _______________________________

Signature: ____________________________ Date: ______________

APPROVED BY THE AUCKLAND HEALTH RESEARCH ETHICS COMMITTEE ON
_____________ FOR THREE (03) YEARS.
REFERENCE NUMBER ___________

Appendix C

Oxford Knee Score questions

The Oxford Knee Score questionnaire consists of 12 questions covering function and pain of the knee. Each question is scored from 0 to 4 and the overall score is the sum of all items ranging from 0 (worst outcome) to 48 (best outcome). A grade from 0 to 19 may indicate severe knee arthritis and therefore a high probability of surgical intervention. A score between 20 and 29 may suggest moderate to serious knee osteoarthritis and requires an appointment with an orthopaedic surgeon for assessment. 30 to 39 represents mild to moderate knee arthritis, a consultation with a physician is recommended for examination and nonsurgical treatment might be prescribed. Finally, a score from 40 to 48 refers to a satisfactory knee function [117].

1. How would you describe the pain you usually have in your knee?
   - None
   - Very mild
   - Mild
   - Moderate
   - Severe
2. Have you had any trouble washing and drying yourself (all over) because of your knee?
   - No trouble at all
   - Very little trouble
   - Moderate trouble
   - Extreme difficulty
   - Impossible to do
3. Have you had any trouble getting in and out of the car or using public transport because of your knee? (With or without a stick)
   - No trouble at all
   - Very little trouble
   - Moderate trouble
   - Extreme difficulty
   - Impossible to do
4. For how long are you able to walk before the pain in your knee becomes severe? (With or without a stick)

- No pain, more than 60 min
- 16 - 60 minutes
- 5 - 15 minutes
- Around the house only
- Not at all

5. After a meal (sat at a table), how painful has it been for you to stand up from a chair because of your knee?
    - Not at all painful
    - Slightly painful
    - Moderately pain
    - Very painful
    - Unbearable
6. Have you been limping when walking, because of your knee?
    - Rarely / never
    - Sometimes or just at first
    - Often, not just at first
    - Most of the time
    - All of the time
7. Could you kneel down and get up again afterwards?
    - Yes, easily
    - With little difficulty
    - With moderate difficulty
    - With extreme difficulty
    - No, impossible
8. Are you troubled by pain in your knee at night in bed?
    - Not at all
    - Only one or two nights
    - Some nights
    - Most nights
    - Every night
9. How much has pain from your knee interfered with your usual work? (including housework)
    - Not at all
    - A little bit
    - Moderately
    - Greatly
    - Totally

10. Have you felt that your knee might suddenly "give away" or let you down?
    - Rarely / Never
    - Sometimes or just at first
    - Often, not at first
    - Most of the time
    - All the time
11. Could you do household shopping on your own?
    - Yes, easily
    - With little difficulty
    - With moderate difficulty
    - With extreme difficulty
    - No, impossible
12. Could you walk down a flight of stairs?
    - Yes, easily
    - With little difficulty
    - With moderate difficulty
    - With extreme difficulty
    - No, impossible

Appendix D

EQ-5D

EQ-5D is a common tool to measure generic health status developed by the EuroQol group. It has been widely employed in clinical trials, population studies and real-world settings. It explores five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Depending on the version, it has different levels. The one used in this work is EQ-5D-5L based on 5 levels: no problems, slight problems, moderate problems, severe problems and extreme problems. Each dimension is scoring with a one-digit number representing the level selected, 1 characterising the best outcome and 5 the worst. The result of this test is a five-digit number describing the patient's health state. The second part of the test comprises a visual analogue scale (VAS) from 0 to 100, 0 being the worst health possible and 100 the best one. This is used as a quantitative measure of health outcomes that reflects the patient's own judgement [118].

Please thick the **one** box that best describes your health **Today**.

1. Mobility
   - I have no problems in walking about
   - I have slight problems in walking about
   - I have moderate problems in walking about
   - I have severe problems in walking about
   - I am unable to walk about
2. Self-care
   - I have no problems washing or dressing myself
   - I have slight problems washing or dressing myself
   - I have moderate problems washing or dressing myself
   - I have severe problems washing or dressing myself
   - I am unable to wash or dress myself
3. Usual activities (e.g. work, study, housework, family or leisure activities)
   - I have no problems doing my usual activities
   - I have slight problems doing my usual activities
   - I have moderate problems doing my usual activities
   - I have severe problems doing my usual activities
   - I am unable to do my usual activities

4. Pain/Discomfort
   - I have no pain or discomfort
   - I have slight pain or discomfort
   - I have moderate pain or discomfort
   - I have severe pain or discomfort
   - I have extreme pain or discomfort
5. Anxiety/Depression
   - I am not anxious or depressed
   - I am slightly anxious or depressed
   - I am moderately anxious or depressed
   - I am severely anxious or depressed
   - I am extremely anxious or depressed
6. We would like to know how good or bad your health is **today**. This scale is numbered from 0 to 100, 100 meaning the best health you can imagine and 0 the worst health. Mark an X on the scale to indicate how your health is Today.

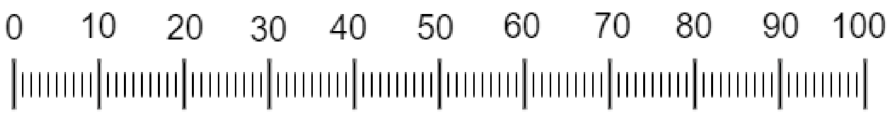

Figure D.1: Scale from 0 to 100 (adapted from [118]).

Appendix E

Forgotten Joint Score

The Forgotten Joint Score is composed of 12 questions concerning how often the patient is aware of his affected knee joint in everyday life. Each question is scored from 1 (never) to 5 (mostly). The raw grade is the sum of all items ranging from 12 (best) to 60 (worst). It is then summarised and transformed into a scale from 0 to 100 where the highest values indicate that the patient is less aware of the affected joint.

1. Are you aware of your knee joint in bed at night?
   - Never
   - Almost never
   - Seldom
   - Sometimes
   - Mostly
2. Are you aware of your knee joint when you are sitting on a chair for more than one hour?
   - Never
   - Almost never
   - Seldom
   - Sometimes
   - Mostly
3. Are you aware of your knee joint when you are walking for more than 15 minutes?
   - Never
   - Almost never
   - Seldom
   - Sometimes
   - Mostly
4. Are you aware of your knee joint when you are taking a bath/shower?
   - Never
   - Almost never
   - Seldom
   - Sometimes

- Mostly

5. Are you aware of your knee joint when you are traveling in a car?
   - Never
   - Almost never
   - Seldom
   - Sometimes
   - Mostly
6. Are you aware of your knee joint when you are climbing stairs?
   - Never
   - Almost never
   - Seldom
   - Sometimes
   - Mostly
7. Are you aware of your knee joint when you are walking on uneven ground?
   - Never
   - Almost never
   - Seldom
   - Sometimes
   - Mostly
8. Are you aware of your knee joint when you are standing up from a low-sitting position?
   - Never
   - Almost never
   - Seldom
   - Sometimes
   - Mostly
9. Are you aware of your knee joint when you are standing for long periods of time?
   - Never
   - Almost never
   - Seldom
   - Sometimes
   - Mostly
10. Are you aware of your knee joint when you are doing housework or gardening?
    - Never
    - Almost never
    - Seldom
    - Sometimes

- Mostly

11. Are you aware of your knee joint when you are taking a walk/hiking?
    - Never
    - Almost never
    - Seldom
    - Sometimes
    - Mostly
12. Are you aware of your knee joint when you are doing your favorite sport?
    - Never
    - Almost never
    - Seldom
    - Sometimes
    - Mostly

Appendix F

Example of Selected Features

Accelerometer features	Gyroscope features
agg_autocorrelation__f_agg_"var"__maxlag_40	agg_autocorrelation__f_agg_"mean"__maxlag_40
agg_linear_trend__f_agg_"max"__chunk_len_10__attr_"rvalue"	agg_linear_trend__f_agg_"max"__chunk_len_10__attr_"intercept"
agg_linear_trend__f_agg_"max"__chunk_len_10__attr_"slope"	agg_linear_trend__f_agg_"min"__chunk_len_10__attr_"intercept"
agg_linear_trend__f_agg_"mean"__chunk_len_10__attr_"slope"	agg_linear_trend__f_agg_"min"__chunk_len_5__attr_"slope"
agg_linear_trend__f_agg_"min"__chunk_len_10__attr_"stderr"	agg_linear_trend__f_agg_"var"__chunk_len_10__attr_"stderr"
agg_linear_trend__f_agg_"min"__chunk_len_50__attr_"intercept"	approximate_entropy__m_2__r_0.5
agg_linear_trend__f_agg_"min"__chunk_len_5__attr_"intercept"	approximate_entropy__m_2__r_0.7
agg_linear_trend__f_agg_"var"__chunk_len_5__attr_"intercept"	approximate_entropy__m_2__r_0.9
agg_linear_trend__f_agg_"var"__chunk_len_5__attr_"slope"	autocorrelation__lag_3
approximate_entropy__m_2__r_0.3	autocorrelation__lag_4
approximate_entropy__m_2__r_0.5	autocorrelation__lag_6
approximate_entropy__m_2__r_0.7	autocorrelation__lag_7
augmented_dickey_fuller__autolag_"AIC"__attr_"pvalue"	change_quantiles__f_agg_"mean"__isabs_False__qh_1.0__ql_0.2
autocorrelation__lag_6	change_quantiles__f_agg_"mean"__isabs_False__qh_1.0__ql_0.4
change_quantiles__f_agg_"var"__isabs_True__qh_0.8__ql_0.2	change_quantiles__f_agg_"var"__isabs_False__qh_0.6__ql_0.4
cid_ce__normalize_True	count_above_mean
cwt_coefficients__widths_(2, 5, 10, 20)__coeff_10__w_20	cwt_coefficients__widths_(2, 5, 10, 20)__coeff_10__w_20
cwt_coefficients__widths_(2, 5, 10, 20)__coeff_11__w_20	cwt_coefficients__widths_(2, 5, 10, 20)__coeff_13__w_10
cwt_coefficients__widths_(2, 5, 10, 20)__coeff_12__w_20	cwt_coefficients__widths_(2, 5, 10, 20)__coeff_8__w_20
cwt_coefficients__widths_(2, 5, 10, 20)__coeff_13__w_20	cwt_coefficients__widths_(2, 5, 10, 20)__coeff_9__w_20
cwt_coefficients__widths_(2, 5, 10, 20)__coeff_14__w_10	energy_ratio_by_chunks__num_segments_10__segment_focus_2
cwt_coefficients__widths_(2, 5, 10, 20)__coeff_3__w_20	energy_ratio_by_chunks__num_segments_10__segment_focus_5
energy_ratio_by_chunks__num_segments_10__segment_focus_0	energy_ratio_by_chunks__num_segments_10__segment_focus_7
energy_ratio_by_chunks__num_segments_10__segment_focus_2	energy_ratio_by_chunks__num_segments_10__segment_focus_9
energy_ratio_by_chunks__num_segments_10__segment_focus_4	fft_aggregated__aggtype_"kurtosis"
energy_ratio_by_chunks__num_segments_10__segment_focus_5	fft_aggregated__aggtype_"variance"
energy_ratio_by_chunks__num_segments_10__segment_focus_7	fft_coefficient__coeff_10__attr_"imag"
energy_ratio_by_chunks__num_segments_10__segment_focus_8	fft_coefficient__coeff_13__attr_"abs"
fft_aggregated__aggtype_"kurtosis"	fft_coefficient__coeff_13__attr_"imag"
fft_aggregated__aggtype_"skew"	fft_coefficient__coeff_14__attr_"imag"
fft_coefficient__coeff_17__attr_"abs"	fft_coefficient__coeff_16__attr_"imag"
fft_coefficient__coeff_1__attr_"real"	fft_coefficient__coeff_1__attr_"real"
fft_coefficient__coeff_28__attr_"abs"	fft_coefficient__coeff_26__attr_"real"
fft_coefficient__coeff_2__attr_"angle"	fft_coefficient__coeff_27__attr_"real"
fft_coefficient__coeff_2__attr_"real"	fft_coefficient__coeff_29__attr_"angle"
fft_coefficient__coeff_48__attr_"real"	fft_coefficient__coeff_30__attr_"real"
fft_coefficient__coeff_49__attr_"imag"	fft_coefficient__coeff_33__attr_"real"
fft_coefficient__coeff_50__attr_"angle"	fft_coefficient__coeff_38__attr_"real"
first_location_of_maximum	fft_coefficient__coeff_3__attr_"abs"
first_location_of_minimum	fft_coefficient__coeff_44__attr_"abs"
friedrich_coefficients__m_3__r_30__coeff_1	fft_coefficient__coeff_4__attr_"abs"

kurtosis	fft_coefficient__coeff_4__attr_"imag"
last_location_of_maximum	fft_coefficient__coeff_5__attr_"imag"
linear_trend__attr_"intercept"	fft_coefficient__coeff_6__attr_"imag"
skewness	friedrich_coefficients__m_3__r_30__coeff_2
spkt_welch_density__coeff_8	index_mass_quantile__q_0.9
	kurtosis
	linear_trend__attr_"slope"
	partial_autocorrelation__lag_2
	spkt_welch_density__coeff_8
	sum_values
	symmetry_looking__r_0.15000000000000002
	time_reversal_asymmetry_statistic__lag_1
	time_reversal_asymmetry_statistic__lag_3

Table F.1: Example of the overview of all 100 features extracted and selected using `tsFresh`.

Bibliography

[1] D. J. Hunter and S. Bierma-Zeinstra. Osteoarthritis. *The Lancet*, 393(10182):1745 – 1759, 2019.

[2] S. Glyn-Jones, A. J. R. Palmer, R. Agricola, A. J. Price, T. L. Vincent, H. Weinans, and A. J. Carr. Osteoarthritis. *The Lancet*, 386(9991):376 – 387, 2015.

[3] A. J. Carr, O. Robertsson, S. Graves, A. J. Price, N. K. Arden, A. Judge, and D. J. Beard. Knee replacement. *The Lancet*, 379(9823):1331–1340, 2012.

[4] I. E. Luna, H. Kehlet, B. Peterson, H. R. Wede, S. J. Hoevsgaard, and E. K. Aasvang. Early patient-reported outcomes versus objective function after total hip and knee arthroplasty. *The Bone & Joint Journal*, 99-B(9):1167–1175, 2017. PMID: 28860396.

[5] A. Gabr, R. Tansey, and F. S. Haddad. *Outcome Measures in Total Knee Arthroplasty*, pages 79–87. Springer International Publishing, Cham, 2015.

[6] E. Palaiochorlidis and T. Karachalios. *Total Knee Arthroplasty. Evaluating Outcomes*, pages 39–47. Springer London, London, 2015.

[7] N. Artz, S. Dixon, V. Wylde, A. Beswick, A. Blom, and R. Gooberman-Hill. Physiotherapy Provision Following Discharge after Total Hip and Total Knee Replacement: A Survey of Current Practice at High-Volume NHS Hospitals in England and Wales. *Musculoskeletal Care*, 11(1):31–38, 2013.

[8] S. Konan, F. Hossain, S. Patel, and F. S. Haddad. Measuring function after hip and knee surgery. *The Bone & Joint Journal*, 96-B(11):1431–1435, 2014. PMID: 25371452.

[9] A. K. Steinhoff and W. D. Bugbee. Knee Injury and Osteoarthritis Outcome Score has higher responsiveness and lower ceiling effect than Knee Society Function Score after total knee arthroplasty. *Knee Surgery, Sports Traumatology, Arthroscopy*, 24:2627–2633, 2014.

[10] C. Schwartz and Brüls O. Human movement analysis slideshow, 2019.

[11] M. Gholami, C. Napier, and C. Menon. Estimating Lower Extremity Running Gait Kinematics with a Single Accelerometer: A Deep Learning Approach. *Sensors*, 20:2939, 5 2020.

[12] K. Woo. Modelling Patellofemoral Implant Design. *[Master Thesis], Auckland Bioengineering Institute, University of Auckland*, 2 2020.

[13] J. Abulhasan and M. Grey. Anatomy and Physiology of Knee Stability. *Journal of Functional Morphology and Kinesiology*, 2:34, 9 2017.

[14] J. Goldblatt and J. Richmond. Anatomy and biomechanics of the knee. *Operative Techniques in Sports Medicine*, 11(3):172 – 186, 2003. The Multiple Ligament Injured Knee, Part I.

[15] Physiopedia. Knee. `https://www.physio-pedia.com/Knee`. Accessed: 14.03.2020.

[16] C. Andrew. Knee Anatomy, Function and Common Problems. `https://www.healthpages.org/anatomy-function/knee-joint-structure-function-problems/`, 2019. Accessed: 14.03.2020.

[17] D. Ruffoni. Biomechanics slideshow, 2018.

[18] Federation University Australia. Joints and movement. In *Anatomy and Physiology for Health Professionals 1*. 2018. Accessed: 15.03.2020.

[19] Rayner and Smale. Patellofemoral Joint Pain 2013 Consensus. https://www.raynersmale.com/blog/2014/2/25/patellofemoral-joint-pain-2013-consensus, 2013. Accessed: 15.03.2020.

[20] Merriam-Webster online dictionary. Valgus. https://www.merriam-webster.com/dictionary/valgus. Accessed: 28.03.2020.

[21] Merriam-Webster online dictionary. Varus. https://www.merriam-webster.com/dictionary/varus. Accessed: 28.03.2020.

[22] Evans Cycles. Cycling and Knee Pain - Part 1. https://www.evanscycles.com/coffeestop/advice/cycling-and-knee-pain-part-1. Accessed: 30.03.2020.

[23] New Zealand Orthopaedic Association. The New Zealand Joint Registry - Twenty years report January 1999 to December 2018. 2019. https://nzoa.org.nz/nzoa-joint-registry.

[24] ORTHROpride. Belgian Hip and Knee Arthroplasty Registry - 2018 Annual Report. 2019. https://www.ehealth.fgov.be/fr/esante/professionnels-de-la-sante/qermidorthopride/presentation-generale.

[25] B. Jolles-Haeberli. Chirurgie du genou arthrosique. In *Manuel Pratique de Chirurgie Orthopédique*, pages 297 – 321. Elsevier Masson, Paris, 2014.

[26] J. Buckwalter and J. Martin. Osteoarthritis. *Advanced Drug Delivery Reviews*, 58(2):150 – 167, 2006. Drug Delivery in Degenerative Joint Disease.

[27] Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study. *The Lancet*, 380(9844):815 – 823, 2012.

[28] G. S. Man and G. Mologhianu. Osteoarthritis pathogenesis — A complex process that involves the entire joint. *Journal of medicine and life*, 7:37–41, 03 2014.

[29] K. Mohit and M. Nizar N. Pathogenesis of Osteoarthritis. In *Osteoarthritis: Pathogenesis, Diagnosis, Available Treatments, Drug Safety, Regenerative and Precision Medicine*, pages 1–28. ADIS, Springer International Publishing Switzerland, 2015.

[30] G. Barreto, M. Manninen, and K. Eklund. Osteoarthritis and Toll-Like Receptors: When Innate Immunity Meets Chondrocyte Apoptosis. *Biology*, 9:65, 3 2020.

[31] S. Jos. knee-oa-xr-nml-vs-griii. http://www.docjoints.com/knee-replacement/knee-oa-xr-nml-vs-griii-2/. Accessed: 04.04.2020.

[32] eOrthopod. Tibial Osteotomy. https://eorthopod.com/tibial-osteotomy/. Accessed: 07.04.2020.

[33] E. Castiello and S. Affatato. 7 - The first surgical approach for total knee arthroplasty (TKA). In *Surgical Techniques in Total Knee Arthroplasty and Alternative Procedures*, pages 109 – 122. Woodhead Publishing, Oxford, 2015.

[34] OrthoInfo. Unicompartmental Knee Replacement. https://orthoinfo.aaos.org/en/treatment/unicompartmental-knee-replacement. Accessed: 08.04.2020.

[35] OrthoInfo. Knee Replacement Implants. https://orthoinfo.aaos.org/en/treatment/knee-replacement-implants. Accessed: 25.04.2020.

[36] B. Hatten. The Total Knee Replacement Prosthesis. https://www.mykneeguide.com/the-knee/the-knee-prosthesis. Accessed: 25.04.2020.

[37] S. Affatato. 1 - The history of total knee arthroplasty (TKA). In S. Affatato, editor, *Surgical Techniques in Total Knee Arthroplasty and Alternative Procedures*, pages 3 – 16. Woodhead Publishing, Oxford, 2015.

[38] C. Sinai. Prosthesis Types. https://www.cedars-sinai.edu/Patients/Programs-and-Services/Imaging-Center/For-Physicians/Musculoskeletal-Radiology/Exhibits-and-Presentations/Knee-Arthroplasty/Prosthesis-Types.aspx. Accessed: 26.04.2020.

[39] N. Li, Y. Tan, Y. Deng, and L. Chen. Posterior cruciate-retaining versus posterior stabilized total knee arthroplasty: A meta-analysis of randomized controlled trials. *Knee surgery, sports traumatology, arthroscopy : official journal of the ESSKA*, 22, 11 2012.

[40] M. J. Bercik, A. Joshi, and J. Parvizi. Posterior Cruciate-Retaining Versus Posterior-Stabilized Total Knee Arthroplasty: A Meta-Analysis. *The Journal of Arthroplasty*, 28(3):439 – 444, 2013.

[41] K. Yagishita, T. Muneta, Y.-J. Ju, T. Morito, J. Yamazaki, and I. Sekiya. High-flex posterior cruciate-retaining vs posterior cruciate-substituting designs in simultaneous bilateral total knee arthroplasty: a prospective, randomized study. *The Journal of arthroplasty*, 27 3:368–74, 2012.

[42] Mark Karadsheh. Tka Prosthesis Design. https://www.orthobullets.com/recon/5019/tka-prosthesis-design. Accessed: 26.04.2020.

[43] M. Varacallo, T. D. Luo, and N. A. Johanson. *Total Knee Arthroplasty (TKA) Techniques*. StatPearls Publishing, Treasure Island (FL), 2020.

[44] eOrthopod. Artificial Joint Replacement of the Knee. https://eorthopod.com/artificial-joint-replacement-of-the-knee/. Accessed: 27.04.2020.

[45] S. Comitini, D. Tigani, and M. Commessatti. 6 - Preoperative planning in total knee arthroplasty (TKA). In *Surgical Techniques in Total Knee Arthroplasty and Alternative Procedures*, pages 95 – 108. Woodhead Publishing, Oxford, 2015.

[46] L. Rahman and S. Oussedik. *Patient Preparation for Total Knee Arthroplasty: Reducing Blood Loss, Thromboprophylaxis and Reducing Infection Risk*, pages 57–67. Springer International Publishing, Cham, 2015.

[47] D. A. Parker and V. A. Mhaskar. *Surgical Technique of Total Knee Arthroplasty: Basic Concepts Including Surgical Approaches, Minimally Invasive Surgery and Simultaneous Bilateral Arthroplasty*, pages 111–126. Springer International Publishing, Cham, 2015.

[48] A. J. Cien. Minimally invasive (Subvastus) muscle sparing TKA. https://www.dradamcien.com/minimally-invasive-subvastus-muscle-sparing-tka-orthopaedic-joint-replacement-surgeon-south-bend-in.html. Accessed: 03.05.2020.

[49] Zimmer Biomet. Persona®, The Personalized Knee - Surgical Technique. https://www.zimmerbiomet.com/medical-professionals/knee/product/persona-knee-system.html. Accessed: 08.05.2020.

[50] K. Bozic. Total Knee Replacement Surgical Procedure. https://www.arthritis-health.com/surgery/knee-surgery/total-knee-replacement-surgical-procedure. Accessed: 05.05.2020.

[51] B. Parcells. Tka Bone cuts. https://hipandkneebook.com/tja-publication-blog/2017/3/14/tka-bone-cuts, 2017. Accessed: 04.05.2020.

[52] G. R. Scuderi and A. J. Tria. *Knee Arthroplasty Handbook: Techniques in Total Knee and Revision Arthroplasty*. Springer, New York, NY, 2006.

[53] P. L. Sousa and M. P. Abdel. *Technological Aids in Total Knee Arthroplasty: Navigation, Patient-Specific Instrumentation, and Robotics*, pages 127–135. Springer International Publishing, Cham, 2015.

[54] Brüls O. Biomedical robotics and active prostheses slideshow, 2019.

[55] J. H. Lonner and J. F. Fraser. *A Brief History of Robotics in Surgery*, pages 3–12. Springer International Publishing, Cham, 2019.

[56] M. W. Allen and D. J. Jacofsky. *Evolution of Robotics in Arthroplasty*, pages 13–25. Springer International Publishing, Cham, 2019.

[57] OrthoInfo. Activities After Knee Replacement. `https://orthoinfo.aaos.org/en/recovery/activities-after-knee-replacement/`. Accessed: 10.05.2020.

[58] Healthline. Recovery Timeline for TKR: Rehabilitation Stages and Physical Therapy. `https://www.healthline.com/health/total-knee-replacement-surgery/rehabilitation-timeline#weeks-4-to-6`. Accessed: 10.05.2020.

[59] M. Khan, G. Green, A. Gabr, and F. Haddad. *Principles of Revision Total Knee Arthroplasty: Incisions, Approaches, Implant Removal and Debridement*, pages 235–247. 1 2015.

[60] O. Garin. *Ceiling Effect*, pages 631–633. Springer Netherlands, Dordrecht, 2014.

[61] P. Harding, A. E. Holland, C. Delany, and R. S. Hinman. Do Activity Levels Increase After Total Hip and Knee Arthroplasty? *Clinical orthopaedics and related research*, 472 (5):1502-1511, 2014.

[62] J. E. Stevens-Lapsley, M. L. Schenkman, and M. R. Dayton. Comparison of Self-Reported Knee Injury and Osteoarthritis Outcome Score to Performance Measures in Patients After Total Knee Arthroplasty. *PM&R*, 3(6):541–549, 2011.

[63] W. Tao, T. Liu, R. Zheng, and H. Feng. Gait Analysis Using Wearable Sensors. *Sensors (Basel, Switzerland)*, 12:2255–83, 12 2012.

[64] C. Pao. What is an IMU Sensor? `https://www.ceva-dsp.com/ourblog/what-is-an-imu-sensor/`, 2018. Accessed: 05.06.2020.

[65] Yunyi Jia, Longxiang Guo, and Xin Wang. 4 - Real-time control systems. In L. Deka and M. Chowdhury, editors, *Transportation Cyber-Physical Systems*, pages 81 – 113. Elsevier, 2018.

[66] D. Roetenberg. *Inertial and magnetic sensing of human motion*. PhD thesis, 01 2006.

[67] Watson. Vibrating Structure Gyro (VSG) Principles of Operation. `https://watson-gyro.com/support-service/legacy-products/vibrating-structure-gyro-vsg-principles-of-operation/`. Accessed: 05.06.2020.

[68] V. Mabe. How Does a Magnetometer Work? `https://sciencing.com/about-5397128-magnetometer.html`, 2017. Accessed: 06.06.2020.

[69] L. Atallah, G. G. Jones, R. Ali, J. J. H. Leong, B. Lo, and G.-Z. Yang. Observing Recovery from Knee-Replacement Surgery by Using Wearable Sensors. In *Proceedings of the 2011 International Conference on Body Sensor Networks*, BSN '11, page 29–34, USA, 2011. IEEE Computer Society.

[70] K. Aminian and B. Najafi. Capturing human motion using body-fixed sensors: Outdoor measurement and clinical applications. *Journal of Visualization and Computer Animation*, 15:79–94, 05 2004.

[71] D. Trojaniello, A. Cereatti, E. Pelosin, L. Avanzino, A. Mirelman, J. M. Hausdorff, and U. Della Croce. Estimation of step-by-step spatio-temporal parameters of normal and impaired gait using shank-mounted magneto-inertial sensors: application to elderly, hemiparetic, parkinsonian and choreic gait. *Journal of NeuroEngineering and Rehabilitation*, 11(152), 11 2014.

[72] P. Esser, H. Dawes, J. Collett, M. G. Feltham, and K. Howells. Assessment of spatio-temporal gait parameters using inertial measurement units in neurological populations. *Gait Posture*, 34(4):558 – 560, 2011.

[73] D. Trojaniello, A. Ravaschio, J. M. Hausdorff, and A. Cereatti. Comparative assessment of different methods for the estimation of gait temporal parameters using a single inertial sensor: application to elderly, post-stroke, parkinson's disease and huntington's disease subjects. *Gait Posture*, 42(3):310 – 316, 2015.

[74] F. A. Storm, C. J. Buckley, and C. Mazzà. Gait event detection in laboratory and real life settings: Accuracy of ankle and waist sensor based methods. *Gait Posture*, 50:42 – 46, 2016.

[75] H. Chang, Y. Hsu, S. Yang, J. Lin, and Z. Wu. A Wearable Inertial Measurement System with Complementary Filter for Gait Analysis of Patients with Stroke or Parkinson's Disease. *IEEE Access*, 4:8442–8453, 2016.

[76] H. Houdijk, F. M. Appelman, J. M. Van Velzen, L.H. Van der Woude, and C.A. Van Bennekom. Validity of DynaPort GaitMonitor for assessment of spatiotemporal parameters in amputee gait. *Journal of Rehabilitation Research and Development*, 45(suppl 9):1335–1342, 2008.

[77] S. Hundza, W. Hook, C. Harris, S. Mahajan, P. Leslie, C. Spani, L. Spalteholz, B. Birch, D. Commandeur, and N. Livingston. Accurate and Reliable Gait Cycle Detection in Parkinson's Disease. *IEEE transactions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society*, 22, 10 2013.

[78] M. Boutaayamou, C. Schwartz, J. Stamatakis, V. Denoël, D. Maquet, B. Forthomme, J.-L. Croisier, B. Macq, J. Verly, G. Garraux, and O. Brüls. Development and validation of an accelerometer-based method for quantifying gait events. *Medical Engineering and Physics*, 37, 1 2015.

[79] I. Weygers, M. Kok, M. Konings, H. Hallez, H. De Vroey, and K. Claeys. Inertial Sensor-Based Lower Limb Joint Kinematics: A Methodological Systematic Review. *Sensors*, 20(3):673, 1 2020.

[80] G. Cooper, I. Sheret, L. McMillian, K. Siliverdis, N. Sha, D. Hodgins, L. Kenney, and D. Howard. Inertial sensor-based knee flexion/extension angle estimation. *Journal of Biomechanics*, 42(16):2678 – 2685, 2009.

[81] S. Tadano, R. Takeda, and H. Miyagawa. Three Dimensional Gait Analysis Using Wearable Acceleration and Gyro Sensors Based on Quaternion Calculations. *Sensors*, 13(7):9321–9343, 7 2013.

[82] P. Picerno. 25 years of lower limb joint kinematics by using inertial and magnetic sensors: A review of methodological approaches. *Gait Posture*, 51, 11 2016.

[83] E. Dorschky, M. Nitschke, A.-K. Seifer, A. J. Van den Bogert, and B. M. Eskofier. Estimation of gait kinematics and kinetics from inertial sensor data using optimal control of musculoskeletal models. *Journal of Biomechanics*, 95:109278, 2019.

[84] B. H. W. Koning, M. M. Van der Krogt, C. T. M. Baten, and B. F. J. M. Koopman. Driving a musculoskeletal model with inertial and magnetic measurement units. *Computer Methods in Biomechanics and Biomedical Engineering*, 18(9):1003–1013, 2015.

[85] J. Y. Goulermas, A. H. Findlow, C. J. Nester, P. Liatsis, X. Zeng, L. P. J. Kenney, P. Tresadern, S. B. Thies, and D. Howard. An Instance-Based Algorithm with Auxiliary Similarity Information for the Estimation of Gait Kinematics from Wearable Sensors. *IEEE Transactions on Neural Networks*, 19(9):1574–1582, 2008.

[86] A. Findlow, J.Y. Goulermas, C. Nester, D. Howard, and L.P.J. Kenney. Predicting lower limb joint kinematics using wearable motion sensors. *Gait Posture*, 28(1):120 – 126, 2008.

[87] F. J. Wouda, M. Giuberti, G. Bellusci, E. Maartens, J. Reenalda, B.-J. F. Van Beijnum, and P. H. Veltink. Estimation of Vertical Ground Reaction Forces and Sagittal Knee Kinematics During Running Using Three Inertial Sensors. *Frontiers in Physiology*, 9:218, 2018.

[88] P. B. Gastin, O. C. Mclean, R. V.P. Breed, and M. Spittle. Tackle and impact detection in elite australian football using wearable microsensor technology. *Journal of Sports Sciences*, 32(10):947–953, 2014.

[89] L. C. Benson, T. J. Tait, K. Befus, J. Choi, C. Hillson, C. Stilling, S. Grewal, K. MacDonald, K. Pasanen, and C. A. Emery. Validation of a commercially available inertial measurement unit for recording jump load in youth basketball players. *Journal of Sports Sciences*, 38(8):928–936, 2020.

[90] G. Bravo-Illanes, R. Halvorson, R. Matthew, D. Lansdown, C. Ma, and R. Bajcsy. IMU Sensor Fusion Algorithm for Monitoring Knee Kinematics in ACL Reconstructed Patients. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5877–5881, 2019.

[91] M. E. Kayaalp, A. N. Agres, J. Reichmann, M. Bashkuev, G. N. Duda, and R. Becker. Validation of a Novel Device for the Knee Monitoring of Orthopaedic Patients. *Sensors*, 19(23):5193, 11 2019.

[92] R. M. Chapman, W. E. Moschetti, and D. W. Van Citters. Stance and swing phase knee flexion recover at different rates following total knee arthroplasty: An inertial measurement unit study. *Journal of Biomechanics*, 84:129 – 137, 2019.

[93] Y.-P. Huang, Y.-Y. Liu, W.-H. Hsu, L.-J. Lai, and M. S. Lee. Monitoring and Assessment of Rehabilitation Progress on Range of Motion After Total Knee Replacement by Sensor-Based System. *Sensors*, 20(6):1703, 3 2020.

[94] S. Yeung. *Population-based Modelling in Clinical Musculoskeletal Biomechanics*. PhD thesis, University of Auckland, 2019.

[95] Vicon. Blue Trident: Measuring high-fidelity limb and body movement. `https://www.vicon.com/cms/wp-content/uploads/2019/09/PS3439_Blue-Trident-factsheet_Final.pdf`, 2020. Accessed: 20.03.2020.

[96] T. F. Besier, D. L. Sturnieks, J. A. Alderson, and D. G. Lloyd. Repeatability of gait data using a functional hip joint centre and a mean helical knee axis. *Journal of Biomechanics*, 36(8):1159 – 1168, 2003.

[97] G. Wu, S. Siegler, P. Allard, C. Kirtley, A. Leardini, D. Rosenbaum, M. Whittle, D. D'Lima, L. Cristofolini, H. Witte, O. A. Schmid, and I. Stokes. ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion - Part I: Ankle, hip, and spine. *Journal of Biomechanics*, 35:543–548, 4 2002.

[98] G. Wu, F. Van der Helm, D. Veeger, M. Makhsous, P. Roy, C. Anglin, J. Nagels, A. Karduna, K. McQuade, X. Wang, F. Werner, and B. Buchholz. 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*, 38:981–992, 6 2005.

[99] T. L. Jakobsen, H. Kehlet, and T. Bandholm. Reliability of the 6-min walk test after total knee arthroplasty. *Knee surgery, sports traumatology, arthroscopy: official journal of the ESSKA*, 21(11):2625 – 2628, 2013.

[100] A. Matsas, N. Taylor, and H. McBurney. Knee joint kinematics from familiarised treadmill walking can be generalised to overground walking in young unimpaired subjects. *Gait Posture*, 11(1):46 – 53, 2000.

[101] A. V. Wiik, A. Aqil, S. Tankard, A. A. Amis, and J. P. Cobb. Downhill walking gait pattern discriminates between types of knee arthroplasty: improved physiological knee functionality in uka versus tka. *Knee surgery, sports traumatology, arthroscopy*, 23:1748–1755, 2015.

[102] M. Christ, N. Braun, J. Neuffer, and A. W. Kempa-Liehr. Time Series FeatuRe Extraction on basis of Scalable Hypothesis tests (tsfresh – A Python package). *Neurocomputing*, 307:72 – 77, 2018.

[103] M. Christ, A. W. Kempa-Liehr, and M. Feindt. Distributed and parallel time series feature extraction for industrial big data applications. *CoRR*, abs/1610.07717, 2016.

[104] OpenSim documentation. `https://simtk-confluence.stanford.edu:8443/display/OpenSim/Documentation`. Accessed: 25.06.2020.

[105] M. Christ, A. W. Kempa-Liehr, and M. Feindt. Tsfresh documentation. `https://tsfresh.readthedocs.io/en/latest/text/introduction.html`, 2020. Accessed: 25.06.2020.

[106] S. Satyan. What is Mean Squared Error, Mean Absolute Error, Root Mean Squared Error and R Squared? `https://www.studytonight.com/post/what-is-mean-squared-error-mean-absolute-error-root-mean-squared-error-and-r-squared`, 2019. Accessed: 26.07.2020.

[107] DataTechNotes. Regression Model Accuracy (MAE, MSE, RMSE, R-squared) Check in R. https://www.datatechnotes.com/2019/02/regression-model-accuracy-mae-mse-rmse.html, 2019. Accessed: 26.07.2020.

[108] R. Bartlett. *Introduction to Sports Biomechanics: Analysing Human Movement Patterns*. London: Routledge, 2 edition, 2007.

[109] Kent State University. SPSS Tutorials: Paired Samples T-Test. https://libguides.library.kent.edu/SPSS/PairedSamplestTest, 2020. Accessed: 2.08.2020.

[110] Boston University School of Public Health. Paired t-test. https://sphweb.bumc.bu.edu/otlt/MPH-Modules/BS/SAS/SAS4-OneSampleTtest/SAS4-OneSampleTtest7.html, 2016. Accessed: 2.08.2020.

[111] Statistics Solution. Paired Sample T-Test. https://www.statisticssolutions.com/manova-analysis-paired-sample-t-test/, 2020. Accessed: 2.08.2020.

[112] Minitab Blog. Understanding t-tests: 1-sample, 2-sample, and Paired t-tests. https://blog.minitab.com/blog/adventures-in-statistics-2/understanding-t-tests-1-sample-2-sample-and-paired-t-tests, 2016. Accessed: 4.08.2020.

[113] A. W. Kempa-Liehr, J. Oram, A. Wong, M. Finch, and T. F. Besier. Feature Engineering Workflow for Activity Recognition from Synchronized Inertial Measurement Units. In *Pattern Recognition - ACPR 2019 Workshops, Auckland, New Zealand, November 26, 2019, Proceedings*, volume 1180 of *Communications in Computer and Information Science*, pages 223–231. Springer, 2019.

[114] R. Derie, P. Robberechts, P. Van den Berghe, J. Gerlo, D. De Clercq, V. Segers, and J. Davis. Tibial Acceleration-Based Prediction of Maximal Vertical Loading Rate During Overground Running: A Machine Learning Approach. *Frontiers in Bioengineering and Biotechnology*, 8:33, 2020.

[115] F. Wouda, M. Giuberti, G. Bellusci, E. Maartens, J. Reenalda, B.-J. Beijnum, and P. Veltink. Estimation of Vertical Ground Reaction Forces and Sagittal Knee Kinematics During Running Using Three Inertial Sensors. *Frontiers in Physiology*, 9, 3 2018.

[116] A. De Brabandere, J. Emmerzaal, A. Timmermans, I. Jonkers, B. Vanwanseele, and J. Davis. A Machine Learning Approach to Estimate Hip and Knee Joint Loading Using a Mobile Phone-Embedded IMU. *Frontiers in Bioengineering and Biotechnology*, 8:320, 2020.

[117] Oxford Knee Score questionnaire. http://www.orthopaedicscore.com/scorepages/oxford_knee_score.html. Accessed: 03.02.2020.

[118] EuroQol. EQ-5D-5L. https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/, 2017. Accessed: 20.05.2020.

[119] Forgotten Joint Score. http://www.forgotten-joint-score.info/. Accessed: 03.02.2020.

[120] I. Faisal, T. Purboyo, and A. Ansori. A Review of Accelerometer Sensor and Gyroscope Sensor in IMU Sensors on Motion Capture. *Journal of Engineering and Applied Sciences*, 15:826–829, 11 2019.