Biomechanical function in knee osteoarthritis and post-total knee replacement: comparing subjective and objective outcomes and predicting gait function post-total knee arthroplasty

Thesis submitted in fulfilment of the requirement for the degree of Doctor of Philosophy in Biomedical Engineering

Marina De Vecchis

M.Sc.

School of Engineering - Cardiff University

UK

July 2023

Abstract

Patient-reported outcomes (PROMs) have been widely utilised to evaluate the TKR outcome and to predict it. Although PROMs inform on the patient's perception of function, they are mostly influenced by pain levels, and relate poorly to what patients can achieve objectively. For this reason, both subjective and objective function should be measured to comprehensively quantify the impairments pre-TKR, the improvement post-TKR and to aid in TKR outcome prediction. However, there is no gold standard to measure function objectively. This research aimed to advance the application of the Cardiff classifier, a measure of gait biomechanics, to subjects with severe knee OA and post-TKR, to compare the Classifier to similar measures of gait biomechanics, identify factors predictive of the gait biomechanics post-TKR, and investigate the relationship between biomechanics, patient-reported outcomes and physical performance pre- and post-TKR.

3D gait analysis was performed in two cohorts of non-pathological subjects (NPs) and patients pre- and one-year post-TKR (Cardiff and Karolinska cohorts). The Cardiff classifier's Belief of OA (BOA), Gait Deviation Index (GDI) and the GDI-kinetics were utilised to evaluate patients' objective gait function at each time point in both cohorts. The BOA had a large responsiveness to change, which was greater than the GDI and GDI-kinetic in 39 patients from the UK and 29 from Sweden. While the correlation between BOA-GDI and BOA-GDI-kinetic was moderate pre-TKR in both cohorts, the two gait indexes and their change pre to post-TKR showed poor or mixed agreement with the BOA post-TKR or its change score. By comparing the outputs of the classifiers developed from each cohort, it was found that about 55% of the highest-ranking gait features discriminating patients pre-TKR to their references were the same or similar between Cardiff and Karolinska patients. Gait biomechanics improved in both patients' groups but mostly did not return to normal one-year post-TKR.

In the Cardiff cohort mentioned above, it was demonstrated that when comparing the patients gait function to NPs of similar age (NP50 classifier), the BOA was significantly lower (=better gait) pre- and one-year post-TKR versus comparing patients to a younger group of NPs (mixed-age classifier), but the change in gait function was comparable between the NP50 and mixed-age classifiers one-year post-surgery. Pre-surgical and surgical factors did not correlate to the change in BOA one-year post-TKR (NP50 classifier). A regression model revealed that the objective gait function pre-TKR, sex and BMI explained 56% of the variance of the gait function one-year post-TKR; there was a significant association between a worse gait function pre-surgery and a worse gait biomechanics one-year post-TKR, irrespective of sex and BMI. A patient sub-group analysis also showed that a greater knee ROM pre-TKR was associated with a better gait function post-TKR.

3D gait analysis data, performance-based tests (PBTs) (timed-up and go, 40m fast-paced walk test, stair climb test and 30s chair test), Oxford Knee Score and Knee Injury and Osteoarthritis Outcome Score were collected from patients pre, three and six months post-TKR. It was found that trunk kinematics in the frontal, sagittal and transverse planes were not relevant in aiding in the discrimination of gait biomechanics between 9 NPs (n = 18 knees) and 18 subjects with late-stage OA (n = 20 knees) within the Cardiff classifier. Results showed a correlation, or trends of association, between gait biomechanics and the core PBTs suggested by OARSI (40m fast-paced walk test, stair climb test and 30s chair test), pre, three and six months post-TKR. However, no correlation, nor a trend of association, could be found between PROMs and gait biomechanics or between PROMs and PBTs pre- or three, six and approximately twelve months post-TKR.

Employing the Cardiff classifier to assess *in vivo* knee kinematics during a step-up motion showed an 83.8% accuracy in discriminating between severe knee OA (n = 18) and NP knee function (n = 19). The novel application of the Cardiff classifier to knee kinematics data collected via single-plane fluoroscopy showed that two years post-TKR, the knee function improved but was not comparable to NPs. The level of classification uncertainty was higher than previous studies employing the classifier, suggesting the need to include additional knee arthrokinematics features. Additionally, the results showed that knee kinematics was not associated with the OKS pre- or post-TKR or satisfaction, nor the patient's perception of their knee "feeling like a normal knee" two years post-TKR.

This work aided in expanding the application of the Cardiff classifier beyond the assessment of gait biomechanics, supported the use of PBTs for the assessment of function, reinforced the evidence that objective measures of function and PROMs measure different constructs and should be utilised together in evaluating OA or TKR outcomes.

Acknowledgements

I would like to express my sincere gratitude to my supervisors, Professor Cathy Holt and Dr Gemma Whatling, for giving me the opportunity to embark on this exciting project. Thank you, Cathy, for providing me with invaluable knowledge, technical and emotional support throughout. Thank you, Gemma, for your supportive feedback and guidance throughout the research project.

This study was embedded within the work of the wider Biomechanics and Bioengineering Research Centre Versus Arthritis, and therefore, thank you to our Centre Director Professor Valerie Sparkes, our research manager Cheryl Cleary, Dr Paul White, the Research Officers and the orthopaedic surgeons at CAVOC who were essential in the organisation of the study and patients' recruitment.

Thank you to Versus Arthritis for supporting our research at the Biomechanics and Bioengineering Research Centre Versus Arthritis.

Thank you to all the patients and volunteers who gave their time so kindly and made this research possible.

I would like to thank Natasha, Judit, and Kirsty for contacting the patients and volunteers, and for being so supportive during the data collection. Laboratory sessions were long and complex, and they would have been much more difficult without the help of Zornitza and Dionne, I would like to thank them sincerely for making sure the laboratory was working perfectly before the sessions and for helping me collecting the data and going through the tedious task of processing it.

Thank you to Dr Paul Biggs for his teaching and for always being available to help me understanding the complexities of the Cardiff classifier.

Thank you to Dr Josefine Naili, Professor Anders Holsgaard-Larsen for being great research collaborators, Professor Mark Pickering for teaching me how to use Orthovis, Dr Joe Lynch for always being supportive during my work with fluoroscopic data.

A special thank you goes to all my colleagues at the Musculoskeletal Biomechanics Research Facility, David, Jenny, Becca, to my PhD buddies, particularly Polly Blaikie, your bright mind has always been inspiring, Jake, Alonso and Lauren who, in these years, have been great companions and with whom I created fond memories.

Another special thank you goes to my friends and family for the long chats and support even if living far away; without knowing that you are there, I would have not been able to go through this crazy journey.

Thank you to my managers Dr Russ Coppack and Gp Capt Alexander Bennett at the Academic Department of Military Rehabilitation who have been so supportive during the completion of my PhD, and to all my new colleagues who kept me in a good mood in the final stages of my PhD.

The most special thank you goes to my partner Rob, who has always been there for me, motivated me, made me laugh when I was stressed and kept me going in the most loving way.

Table of Contents

List of abbreviations

2D	Two-dimensional
30sCT	30-second Chair Test
3D	Three-dimensional
40mFPWT	40m fast-paced Walk Test
6MWT	6-Minute Walk Test
ACL	Anterior Cruciate Ligament
ASA	American Society of Anaesthesiologists physical status
BMI	Body Mass Index
BNP	Belief of Non-Pathological
BOA	Belief of Osteoarthritis
BOE	Body of Evidence
CBOE	Combined Body of Evidence
CAD	Computer-Aided Design
CT	Computed Tomography
DRR	Digitally Reconstructed Radiograph
GDI	Gait Deviation Index
GRF	Ground Reaction Force
HSS	Hospital for Special Surgery Score
HTO	High Tibial Osteotomy
ICC	Intraclass Correlation Coefficient
ISB	International Society of Biomechanics
K-L	Kellgren-Lawrence
KOOS	Knee Injury and Osteoarthritis Outcome Score
mSv	milliSievert
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NJR	National Joint Registry
OA	Osteoarthritis
OARSI	OsteoArthritis Research International
OKS	Oxford Knee Score
PBTs	Performance-Based Tests
PC	Principal Component
PCA	Principal Component Analysis
PCL	Posterior Cruciate Ligament

Chapter 1: Introduction

Osteoarthritis (OA) is a musculoskeletal condition that has a considerable impact on the quality of life of people that are affected by it. The Global Burden of Diseases, Injuries and Risk Factors Study found that OA affected 7% of people worldwide in 2019 and classed OA as the 15th primary cause of years lived with disability globally (Vos et al. 2020). In England, it was estimated that knee OA affects about 18% of adults over the age of 45, with roughly 6% of cases classed as severe (Arthritis Research UK 2019). Knee OA is characterised by pain, and changes in the biomechanics of the knee and the joints surrounding it, resulting in loss of function. Total knee replacement surgery (TKR) is often recommended for patients experiencing severe pain, functional limitations, and instability, as there is no cure for joint degeneration. The main goals of TKR are to alleviate pain, improve physical function and, therefore, the quality of life. The number of TKR procedures performed every year is rapidly increasing, with projections of a 401% growth by 2040 (Singh et al. 2019). With the rising amount of TKR surgeries, it is paramount to assess the outcome with appropriate measures that provide insights into the expected level of recovery, identify factors influencing outcomes, and what could be targeted to improve the success of the surgery.

Currently, the NHS in the UK measures outcomes pre to post-TKR via patient-reported outcomes (PROMs) of pain and function (NHS Digital 2017), which only take a few minutes to complete. Subjective outcomes, such as pain levels, quality of life, and patient-reported function, reflect the individual's perception and experience. There is a growing interest in understanding the factors that may influence achieving an optimal outcome, and most of the current research is centred on predicting PROMs. While PROMs provide insights into the patient's perspective and how they perceive the impact of OA and TKR on their daily life and well-being, they are poorly associated with objective physical function (Gandhi et al. 2009b; Stratford et al. 2010; Mizner et al. 2011; Hamilton et al. 2012b; Graff et al. 2016). Hence, the Osteoarthritis Research International (OARSI) suggested that PROMs should be utilised alongside objective measures of function both in clinical and research settings to provide a more comprehensive understanding of the patient’s condition (Dobson et al. 2013). The tests of physical performance suggested by OARSI, such

as the Timed-up and go, do not inform on the quality of the movement, as they only measure the time needed to cover a certain distance (Naili et al. 2017a). Conversely, the assessment of joint biomechanics via three-dimensional (3D) motion analysis provides quantifiable data on patients' compensations in presence of knee OA and allows for an accurate evaluation of the TKR surgery's effectiveness in correcting lower limb biomechanics. 3D motion analysis produces a large amount of data that are complex to interpret. Principal Component Analysis (PCA) enables the reduction of the data employed in the analysis whilst retaining the relevant information on gait biomechanics, and highlights pattern differences between severe knee OA or post-TKR and non-pathological gait (Deluzio et al. 1997; Deluzio et al. 1999; Robertson et al. 2013; Hubley-Kozey and Astephen Wilson 2018).

The current PhD studentship was funded by Versus Arthritis (the leading charity in the UK providing support to individuals with arthritis) and Cardiff University and developed from earlier research conducted within the Biomechanics and Bioengineering Research Centre Versus Arthritis, where a method to classify gait function objectively has been developed (Jones 2004) and refined over time (Biggs 2016), the Cardiff classifier, which is used in combination with PCA. The Cardiff classifier can summarise the gait features extracted via PCA into a single measure of gait function, providing an accurate classification between patients and non-pathological subjects (NPs), informing on which gait biomechanics best discriminate between severe knee OA and NP biomechanics, and assessing changes in gait biomechanics post-TKR as evidenced by studies conducted by Biggs et al., 2019; Jones, 2004; Jones et al., 2008; Metcalfe, 2014; Whatling et al., 2022; Worsley, 2011.

Further research is needed to expand the application of the Cardiff classifier, to understand how it relates to other measures of gait biomechanics and objective physical function utilised in research and clinically, and to the patient's perception of their functional capabilities pre- and post-TKR. Being able to measure function pre- and post-TKR objectively can contribute to the identification of factors that influence postoperative objective outcomes, which may be different from those affecting PROMs, and there is limited research in this area (Devasenapathy et al. 2019). This knowledge can explain potential challenges that patients may encounter after having had a TKR, and aid advancements in objective outcome assessment, to improve

patient care. To achieve these research goals and strengthen the research findings, collaborations were established with researchers from the same field, facilitating access to supplementary datasets including individuals pre- and post-TKR and NPs from Karolinska Institutet in Stockholm, Sweden, and the Australian National University in Canberra, Australia.

Aims and objectives

This PhD thesis aims to advance and refine the application of the Cardiff classifier to subjects with severe knee OA and post-TKR, to compare the Classifier to similar measures of gait biomechanics, as well as to explore factors predictive of the biomechanics post-TKR, and the relationship between biomechanics, patient-reported outcomes and physical performance pre- and post-TKR. The aims were achieved by meeting the following objectives.

Objective 1 – Chapter 4: *To compare the measurement abilities of the Cardiff classifier to the Gait Deviation Index and the Gait Deviation Index-kinetics pre and one-year post-TKR, and to identify the key gait features discriminating knee OA from NP biomechanics.*

Chapter 4 utilises 3D gait analysis on two cohorts of NPs and patients pre- and one-year post-TKR (Cardiff and Karolinska cohorts). The Cardiff classifier, Gait Deviation Index (GDI) and the GDI-kinetics are utilised to evaluate the patients' objective gait function at each time point in both cohorts. The internal responsiveness of the classifier pre- to post-TKR is explored and compared to that of the GDI and GDI-kinetic in both patient groups. The associations between the three measures at each time point and their change are assessed in each patient group. The classifier was utilised to reveal the key gait features best discriminating patients from their relative NPs within each cohort and to explore the patients' gait biomechanics changes one-year post-TKR in each group.

Objective 2 – Chapter 5: *To determine the appropriateness of using a mixed-age reference group to assess the gait biomechanics of people with severe knee OA*

within the Cardiff classifier, and to identify pre-surgical and surgical factors predictive of the gait biomechanics and gait biomechanics changes one-year post-TKR.

Solely the Cardiff cohort from Chapter 4 is utilised within Chapter 5. In Chapter 4, the cohort includes an NP group of mixed-age, younger than the patients. In Chapter 5, a sub-group of Cardiff NPs aged fifty or older is utilised to evaluate the gait biomechanics of Cardiff patients from Chapter 4 via the Cardiff classifier, both pre and one-year post-TKR. The outputs of the classifiers for the Cardiff patients in Chapter 4 and those obtained in Chapter 5 are compared pre- and one-year post-TKR to reveal whether the objective gait function may be underestimated when comparing patients to younger NPs. The association between the change in gait biomechanics post-TKR, pre-surgical, and surgical factors is explored. A regression model is developed to determine if gait biomechanics one-year post-TKR can be predicted via gait biomechanics and BMI pre-surgery, and sex.

Objective 3 – Chapter 6: *To investigate the merit of adding gait trunk kinematics to aid in classifying NPs and individuals with severe knee OA via the Cardiff classifier, and to explore the interrelationship between the gait biomechanics, performance-based tests, and patient-reported outcomes before and after TKR surgery.*

3D gait analysis of trunk and lower limb kinematics is carried out on a group of patients pre-, three-, six-, twelve+ months post-TKR and NPs. Gait biomechanics of the trunk and the affected lower limb are utilised within the Cardiff classifier to reveal if trunk kinematics may contain key features aiding in discriminating severe knee OA gait from that of NPs. Additionally, participants are assessed at each time point via the performance-based tests (PBTs) suggested by OARSI. The interrelationship between patients' gait biomechanics, PBTs and PROMs is explored pre-, three-, six-, and twelve+ months post-TKR to understand if lower PBT scores may be associated with poorer biomechanics and if objective outcomes of function are associated with the patient's perception of their function.

Objective 4 – Chapter 7: *To explore the ability of the Cardiff classifier to discriminate between NP and OA subjects using in vivo knee kinematics during a*

loaded activity (measured via dynamic single-plane fluoroscopy and image registration), and to investigate the association between knee kinematics and patient-reported outcomes pre- and post-TKR.

Chapter 7 utilises knee kinematics data measured via dynamic single-plane fluoroscopy during a step-up movement collected on NPs and patients pre- and two years post-TKR. The Cardiff classifier is utilised to evaluate whether the knee osteo and arthrokinematics can discriminate between patients with end-stage knee OA and NPs, to advance the application of the classifier technique beyond gait analysis. The patients' change in overall knee function (i.e., knee kinematics) is assessed post-TKR to understand the extent of recovery of non-pathological knee kinematics. The relationship between knee kinematics and PROMs is explored pre-and post-TKR to understand how knee osteo and arthrokinematics may affect the patient's perception of their function. Moreover, the association between the knee kinematics post-TKR and satisfaction/their knee "feeling like a normal knee" is explored.

Chapter 2: Literature review

2.1 KNEE OSTEOARTHRITIS

2.1.1 Definition and symptoms

Among arthritic diseases, osteoarthritis (OA) is the most frequent and knee OA is one of the main causes of disability worldwide (Cross et al. 2014). OA is the chronic inflammation and gradual alteration of joint tissues, and it occurs over time (National Clinical Guideline Centre 2014). Knee OA can affect individually the medial, lateral compartments of the joint or the patellofemoral joint. In the presence of severe knee OA, multiple compartments of the knee joint are usually affected.

Knee OA is the result of hyaline cartilage erosion following the alteration of the anabolic and catabolic activity that creates and breaks, respectively, the cartilage matrix (Michael et al. 2010). It is not clear why hyaline cartilage starts to degenerate, OA is multifactorial, and it is believed that a combination of metabolic and biomechanical influences may play a role (Michael et al. 2010).

The major complaint in people suffering from knee OA is pain that is exacerbated by movement and gradually disappears with rest (Michael et al. 2010). Pain tends to become more persistent as the pathology advances and it affects about 75% of people with OA (Versus Arthritis 2019).

The pain is usually described as dull (Michael et al. 2010) and swelling of the knee may occur due to tissue inflammation. Knee OA can also cause joint stiffness that lasts less than thirty minutes when starting to move after a period of inactivity (NICE 2014). Knee OA affects the quality of life and can cause disability. As the pathology progresses, pain becomes continuous, and irreversible structural changes of the joint occur, limiting the knee's range of motion (ROM), affecting the way the person moves and the ability to carry out everyday life activities. In some cases, the

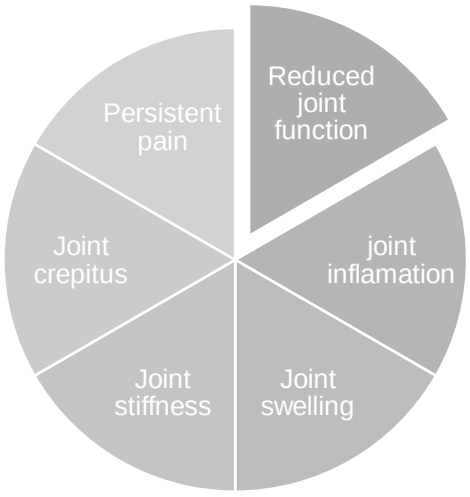


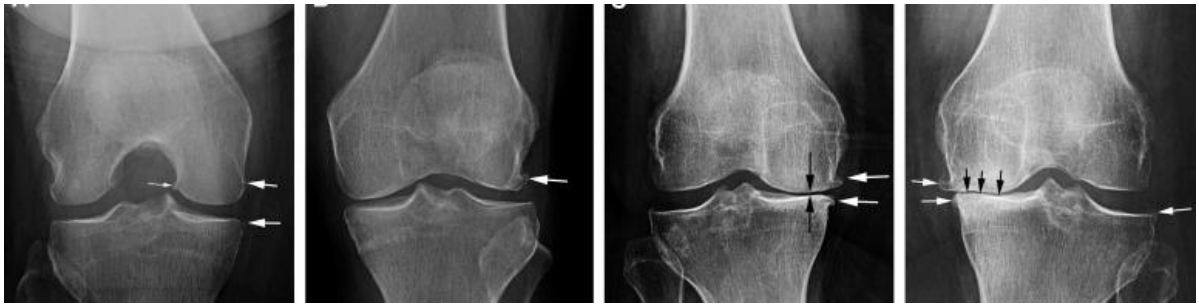
Figure 1 Most common severe knee OA symptoms.

use of a walking aid, such as a walking stick or crutches, becomes necessary to be able to walk.

Pain and limited mobility can decrease independence and be associated with depression, reduced ability to work, and limited physical activity, possibly playing a role in causing frailty in older people (Osteoarthritis Research Society International 2016). Moreover, having OA increases the risk of cardiovascular complications and the likelihood of having ischemic heart disease or heart failure is three times higher than people that do not suffer from OA (Hall et al. 2016).

2.1.2 Diagnosis

Knee OA is often clinically diagnosed when the condition is symptomatic, macro changes in the knee joint have occurred and can be seen objectively through radiographs (Braun and Gold 2012). The radiographic severity of knee OA is determined by the presence of joint space narrowing, indicating thinning of the cartilage, osteophytes, sclerosis of the bone and altered alignment of the tibia relative to the femur on radiographs (Kellgren and Lawrence 1957). Radiographic findings are usually graded with the Kellgren-Lawrence (KL) grading scheme (Kellgren and Lawrence 1957). The KL score has five grades, ranging from 0 (no signs of osteoarthritis) to IV (presence of several osteophytes, joint space narrowing, joint deformity) (Kellgren and Lawrence 1957), grades 1 to 4 are shown in Figure 2.



I II III IV

Figure 2 Radiographs of the knee with varying KL grades of OA: I) Small osteophytes; II) Clear presence of osteophyte but normal joint space; III) Several osteophytes and moderate loss of joint space; IV) Numerous osteophytes, severe joint space narrowing, subchondral bone sclerosis, bone shape deformity (Kellgren and Lawrence 1957).

The radiographic presence of knee OA is usually defined when the KL grade ≥ II (moderate knee OA), meaning that a distinct osteophyte can be identified (Kellgren and Lawrence 1957). Interestingly, not all the people who display radiographic signs of knee OA experience symptoms (Suri et al. 2012), and minimal joint alterations can be associated with the presence of considerable pain (National Clinical Guideline Centre 2014). This is why the radiographic definition of knee OA should be coupled with a clinical investigation to assess the presence of the typical signs of advanced-stage knee OA such as stiffness, movement restriction, gonalgia (otherwise known as knee pain) and joint crepitus (Luyten et al. 2012) and to help to define which interventions are most appropriate.

2.1.3 Epidemiology

OA has a rising burden both on individuals and society and in the classification of the leading causes of disability, it was in the fifth position for the increase in years lived with a disability between 1990 and 2015, on a global scale (Kiadaliri et al. 2018). The prevalence of knee OA is disproportionate between genders. A study carried out in the US found that the lifetime risk of developing knee OA tends to be higher in women (47%) than men (40%) (Murphy et al. 2008). This is in line with recent findings, displaying that roughly 18% of women older than 60 years had symptomatic OA, globally, but only 9.6% of men were affected (WHO 2019).

The epidemiological literature showed that knee OA prevalence and incidence tend to increase with age (Felson et al. 2000; Prieto-Alhambra et al. 2014). (Yu et al. 2015) found that the incidence of OA in England tended to raise steeply in the age range 45-64, reaching the highest values between 75 and 84 years. The authors (Yu et al. 2015) also reported that knee OA incidence (3.5%) was higher than hip (1.4%) and hand (1.3%) OA. Analogously, (Losina et al. 2013) conducted a study retrieving data from the National Health Interview Survey in the US and they found that the incidence of symptomatic knee OA was higher between the age of 55 and 64. An important aspect to consider is that among the people who suffer from OA, more than half are younger than sixty-five years old (Deshpande et al. 2016). Since these people will most likely live for 20 years or longer, they will possibly suffer from a progressive level of disability as they age.

The OA global prevalence increased steeply by 48% between 1990 and 2019 (Vos et al. 2020), and it was estimated that 25% of people over the age of 18 will have OA by 2030. One of the possible reasons why OA prevalence is rising is the increase of some of the major risk factors for the disease, such as obesity and physical inactivity.

2.1.4 Risk factors for OA

There are several risk factors involved in the development of OA. Because of the complex, multifactorial nature of OA, the initiating mechanisms are not clear. Age has been described as one of the main risk factors for OA, but the reasons for it are unclear. With ageing, the reparative processes and metabolism are slower, and several risk factors may have been developed throughout an individual's life and facilitating the onset of OA (Neogi and Zhang 2013). Having OA within the family increases the risk to develop the pathology (Kerkhof et al. 2014), and a review highlighted that knee OA pain intensity seems to be genetically modulated (Neogi and Zhang 2013).

As previously reported, knee OA is more prevalent in women, and since the incidence of knee OA tends to increase in the postmenopausal period when estrogen concentrations decrease (Boyan et al. 2013), it was hypothesised that hormones may play a role in developing knee OA. A study on 2621 women, found that those with a low level of estradiol (an estrogen hormone) had a higher risk of developing knee OA (Hussain et al. 2014). Another study on 4766 women after menopause found that those utilising hormone therapy had a lower prevalence of radiographic and symptomatic knee OA, compared to those not utilising such treatment (Jung et al. 2019). It was also found that women naturally tend to have a reduced cartilage thickness compared to men, a higher risk of cartilage issues at the tibiofemoral level, and a higher rate of cartilage thinning over time compared to men, potentially resulting in a higher susceptibility to knee OA (Hanna et al. 2009).

Another major risk factor for developing knee OA is a joint injury. Early OA is characterised by a change in cell activity that seems to be associated with an inflammatory reaction within the joint, more often caused by ligament injuries and meniscal tears (Paschos 2017). It was demonstrated that the cartilage matrix starts to reorganise after an Anterior Cruciate Ligament (ACL) injury, and these changes

continue for months after the damage (Bigoni et al. 2013). These modifications could produce cartilage catabolism (i.e., the breakdown of the tissue) and start tissue degeneration (Bigoni et al. 2013). A systematic review demonstrated that the prevalence of knee OA was between 0-13% after an ACL rupture and if a meniscus was damaged, the prevalence was higher and between 21-48% (Øiestad et al. 2009). This seems to demonstrate that an ACL tear alone is not a sufficient condition to start the development of knee OA. In many cases, an ACL injury exposes to a higher risk of meniscal damage because of augmented anterior translation of the tibia (Arner et al. 2016). Knee OA following these injuries may be the result of the summation of inflammatory events that disrupted the joint's homeostasis. Moreover, knee biomechanics change after an ACL injury, with abnormal knee kinematics and higher forces on the joint cartilage (Yoo et al. 2005; Wang et al. 2015). This, together with possible menisci injuries, contributes to joint tissue remodelling that can lead to OA over time.

One of the major risk factors for knee OA is obesity, which has increased considerably in the last ten years and it was suggested that this is another reason why the number of people with knee OA is likely to increase too over time (Kurtz et al. 2007; Flegal et al. 2010). People who are obese, with a Body Mass Index (BMI) ≥ 30, have a 60% risk of developing OA symptoms (Murphy et al. 2008). It is not clear whether obesity plays a role in developing knee OA because of altering knee biomechanics or because of causing changes at a systemic level. A study found that an excessive amount of body fat could help sustain chronic low-grade inflammation, with the result of increasing the risk for OA (Wluka et al. 2013). Another study conducted on a large cohort of people who underwent knee or hip joint replacement found a weak association between obesity and hip or wrist OA (Stürmer et al. 2000). Conversely, there was a strong association between knee OA and obesity. These results led the authors to conclude that obesity does not seem to be a systemic factor for developing OA, but rather a mechanical risk factor for it because of the high BMI placing a considerable overload on the knee joints (Stürmer et al. 2000). Loening et al. (2000) investigated the role of joint overload on cartilage metabolism on bovine articular cartilage and found that chondrocyte apoptosis took place when the tissue was exposed to excessive compressive forces. This study showed how

overloading of the joint cartilage can start a process of degeneration within the joint, and possibly lead to knee OA.

The role of exercise as a risk factor for OA is not clear yet as the evidence surrounding this topic is often contrasting. Results from the Framingham study showed how recreational exercise such as jogging, walking or more intense exercise did not cause a higher or lower risk of developing OA (Felson et al. 2007). However, both low and high levels of physical activity have been identified as risk factors for developing OA. A low level of physical activity can play a role in chronic low-grade inflammation (Handschin and Spiegelman 2008) and it causes loss of muscles mass in the long term, resulting in decreased stabilisation of the knee joints, which is partially granted by muscles (Roos et al. 2011). This can cause an abnormal distribution of forces within the knee joint and result in cartilage damage as discussed previously. Conversely, high levels of physical activity have been associated with an increase in knee osteophytes and as a risk factor for knee OA (Vignon et al. 2006).

Malalignment of the knee joint either in valgus (knock-knee*)* or varus (bowlegs), is not an indicator of the presence of knee OA but it does increase the risk of developing the disease (Sharma et al. 2001). This can be explained by the fact that in a valgus or varus knee, forces are focalised either on the lateral or medial side of the tibial plate, respectively. In normal circumstances, with an optimal knee alignment, forces are distributed on both sides of the tibial plateau, with the medial side bearing around 60% of the body load during walking (Houglum et al. 2012).

2.2 BIOMECHANICAL ALTERATIONS IN THE PRESENCE OF KNEE OSTEOARTHRITIS

Altered joint loading and certain gait characteristics may be involved in the progression of knee OA. Hence, there has been a growing interest in investigating the biomechanical alterations in the presence of signs of knee OA to aid in designing strategies to modify the OA-related biomechanical changes and slow down the advancement of the disease. One of the most reported findings is that people with knee OA walk slower than NPs (Ismailidis et al. 2020), and those with severe knee OA have a slower gait than subjects with moderate OA (Astephen et al. 2008a). In general, gait alterations tend to become more pronounced with the disease progression (Astephen et al. 2008a; Mills et al. 2013). The medial knee compartment

is more often affected by knee OA than the lateral or patellofemoral ones (Hubley-Kozey and Astephen Wilson 2018), thus, the external knee adduction moment, which was correlated to knee pain levels (Thorp et al. 2007) and is a proxy of medial knee joint loading (Favre and Jolles 2016), has been widely studied. It was found that the disease progressed faster in people with a greater knee adduction moment in mid-stance (Chehab et al. 2014). Moreover, the knee adduction moment, while bimodal in NPs, was found to be monophasic in people with knee OA (Astephen et al. 2008a).

Joint moments are influenced by the magnitude of the ground reaction force (GRF) and the moment arm between the GRF vector (pointing towards the body's centre of mass) and the joint centre. A larger moment arm would result in larger joint moments. Given that joint moments are dependent on joint angles and GRFs, these have also been explored. NPs have a biphasic vertical GRF, while patients with symptomatic OA have a monophasic, or a less obvious biphasic waveform (Costello et al. 2021). Nevertheless, when considering the frontal plane angles, there is a higher degree of error when utilising motion capture techniques (Hubley-Kozey and Astephen Wilson 2018). Because of this, the literature looking at frontal plane angles is relatively limited. People with severe OA display a larger knee adduction-abduction ROM (Esrafilian et al. 2013), an increased knee adduction (i.e., varus) at heel strike and weight acceptance (Duffell et al. 2017), compared to NPs. Common findings at the knee are that people with knee OA have reduced knee flexion at heel strike and terminal stance (Favre et al. 2014; Ismailidis et al. 2020), a reduced peak knee flexion during swing (Esrafilian et al. 2013; Turcot et al. 2013; Ismailidis et al. 2020), resulting in a smaller knee flexion-extension ROM (Mills et al. 2013; Turcot et al. 2013), with an overall stiff knee.

Walking requires the combination of movements of all lower limb joints, and movement alterations affecting the knee, consequently, producing compensations at the hip and ankle, the pelvis and the trunk alike. At the lower limb joints, it was found that subjects with knee OA have increased hip flexion at heel strike (Turcot et al. 2013) and smaller hip extension at terminal stance, resulting in decreased sagittal hip ROM (Astephen et al. 2008a; Ro et al. 2018). Several authors described decreased hip adduction during stance (Mills et al. 2013; Duffell et al. 2017) and

increased adduction during swing (Baert et al. 2013). Another common finding is reduced ankle plantarflexion around push-off (Naili et al. 2017a).

Considering that the upper body is the weightiest segment of the body, the movements of the trunk may have an influence on hip-knee-ankle joint moments, by shifting the body's centre of mass and, therefore, the moment arm. Patients with severe knee OA tend to have an increased pelvis anteversion and obliquity (Turcot et al. 2013). Moreover, a recent systematic review found a greater trunk lean towards the stance limb in subjects with varying degrees of knee OA (Iijima et al. 2019). Previous research found that increasing the trunk lean as described above, reduces the knee adduction moment in healthy volunteers (Nüesch et al. 2016). Nevertheless, this strategy had the opposite effect in patients with severe knee OA and a varus malalignment (Iijima et al. 2019), suggesting that it is the knee alignment, more than the trunk lean, playing a major role in determining the knee adduction moments. Therefore, patients with a severe knee varus, may not benefit from leaning towards the ipsilateral limb, neither biomechanically, nor in terms of pain reduction (Hunt et al. 2008; Simic et al. 2012). The very limited research investigating trunk sagittal angles in people with knee OA, reports that the flexion-extension ROM is increased, but only in patients with a knee varus alignment (Turcot et al. 2013).

2.3 TREATMENTS FOR KNEE OA

In the UK, among those older than 45 years, about 8.75 million people (33%) required treatment for OA (Versus Arthritis 2019). As discussed in the previous section, the mechanisms underlying OA initiation are not fully understood and a cure for OA has not been found yet. Therefore, the treatment for OA is focused on slowing down its progression and improving the symptoms such as pain and stiffness of the knee joint (Michael et al. 2010).

2.3.1 Conservative treatments

Conservative treatment is recommended when knee OA is in its initial stages or moderate, and it involves pain management with pharmacological and non-pharmacological options, interventions to modify the OA-related biomechanics, and lifestyle changes. Since the focus of the current thesis is related to total knee

replacement (TKR), these treatment options will only be described in brief. Pharmacological management of knee pain comprises oral analgesics, topical treatments with analgesics and intra-articular injections. Paracetamol is widely utilised for treating joint pain and it gives relief in the short term (Bannuru et al. 2010), however, it should be utilised carefully due to its side effects (Craig et al. 2012). Intra-articular injections of corticosteroids are efficacious in pain relief but only for short periods (McAlindon et al. 2014). In terms of non-pharmacological pain management, while manual therapy may be beneficial in selected cases, electrotherapy or acupuncture are not recommended due to their little efficacy (NICE 2022). Other options are to utilise walking aids to help unload the painful joints, and orthoses like insoles or knee braces to help correct the lower limb biomechanics. A Cochrane Review found that people who wore orthoses had significantly better function and lower levels of pain than controls (Brouwer et al. 2005). Wearing knee braces is also effective in improving stiffness, reducing pain and function (Raja and Dewan 2011).

In the UK, the National Institute for Health and Care Excellence (NICE) suggests that people who suffer from knee OA should be encouraged by a healthcare professional to change those lifestyle aspects that constitute risk factors for OA progression (physical inactivity, being overweight), and be guided on how to self-manage the condition (NICE 2022). Additionally, the NICE (2022) advises physiotherapists to include stretching of lower limb muscles and exercises to improve aerobic fitness and muscle strength. When pain and function are considerably impaired and both pharmacological and non-pharmacological options are not effective in improving symptoms, other more invasive treatment options, such as knee arthroscopy, high tibial osteotomy (HTO) or unicompartmental/TKR, should be considered.

2.3.2 Total knee replacement

Knee replacement surgeries are classed as invasive, major surgical procedures. In the presence of unicompartmental knee OA, unicompartmental knee arthroplasty (UKR) and HTO are generally the favoured surgeries and are usually performed on younger patients (Dettoni et al. 2010). UKR consists in replacing the condyle and the underlying tibial surface on the affected side only with an artificial

joint surface. HTO involves opening a wedge in the proximal tibia to realign the lower limb and re-distribute the knee loads, with the aim to reduce the stress in the knee compartment affected by degenerative processes and delay a potential total knee replacement (Dettoni et al. 2010).

The presence of end-stage knee OA, where several compartments of the knee are affected, and symptoms interfere severely with an individual's life, usually warrant referral for TKR surgery (Price et al. 2018). TKR is performed with the aims of reducing pain, increasing the quality of life, and improving function re-establishing the correct lower limb alignment with a balanced tension of the ligaments during the knee ROM (Cyteval 2016). During a TKR surgery, the osteophytes are removed, and the distal femur, proximal tibia and, when necessary, the articular surface of the patella are resected to accommodate the new artificial joint surfaces. The ACL is always removed, while the posterior cruciate ligament (PCL) may be spared, depending on the clinical presentation and the implant design the surgeon deemed most appropriate for the patient (Vaienti et al. 2017). A metal femoral and tibial component (a patellar polyethylene articular surface too, if necessary) are fixated to the bones; finally, a polyethylene insert is positioned between the femoral and tibial metal components, to provide stability and mimic the function of the menisci and the PCL, if this was removed (Innocenti 2022). In the past, a TKR prosthesis survivorship was around 10 years, but given the advancements in implant designs and techniques, a modern prosthesis usually lasts about 25 years in 82% of cases, as reported by a recent systematic review with meta-analysis (Evans et al. 2019).

Most prostheses in the UK are cemented (85% of surgeries in 2017) and the use of non-cemented TKR decreased from 6% to 3% between 2009 and 2017 (NJR 2018b). The choice of mainly utilising cemented implants comes from the evidence that non-cemented prostheses have a worse survival rate in the long term, and in some cases the implant may migrate from the initial position (Gandhi et al. 2009a), resulting in revision surgery.

There are various TKR implant designs which accommodate the choice of sacrificing or retaining the PCL and are called cruciate-sacrificing (posteriorly stabilised or cruciate substituting) and cruciate-retaining designs, respectively (Innocenti 2022). The National Joint Registry (NJR) reported that 75% of TKR prostheses in the UK were unconstrained (Brittain et al. 2022), meaning that the PCL

was spared. The unconstrained implant tends to allow for more physiological movements in the presence of a preoperatively stable knee and the knee joint is stabilised by collateral ligaments and the PCL (Cyteval 2016). In the cruciate-sacrificing implant, the PCL is sacrificed during the surgery and the shape of the polyethene insert stops the tibia from translating posteriorly during knee flexion, offering additional stability to the knee (Cyteval 2016). The polyethylene insert can also have different characteristics and be fixed-bearing, mobile-bearing or offer a medial pivot. In the fixed-bearing implant, the tibial component and the polyethene insert are fixed together; in the mobile-bearing design, the polyethene component can rotate by a few degrees in the transverse plane, mimicking the physiologic movements of menisci, hence, allowing for greater knee mobility (Innocenti 2022). The number of mobile and fixed-bearing prostheses utilised by surgeons was similar in the UK, the NJR reported that 44% of implants were fixed-bearing and 56% were mobile-bearing (NJR 2018a). The choice to use one or the other may have been due to the similar outcomes that these implants have (Jiang et al. 2016). In fact, a Cochrane Review demonstrated that the use of fixed or mobile bearing implants seems to produce similar results in terms of revision and mortality rates, complications, quality of life, functional and clinical outcomes, however, the evidence was of moderate to low quality and more research is needed (Hofstede et al. 2015).

The NJR reported that the number of knee replacement surgeries in the UK was 237 924 between January 2019 and December 2021, 86.1% of them were TKRs, and of these, only 1% of patients received a bilateral TKR (Brittain et al. 2022). In 2021, TKR was performed due to the presence of knee OA in 97% of cases, only 1% of procedures were bilateral, and 52% of patients had their patella resurfaced (Brittain et al. 2022). British patients who underwent TKR in 2021 had a mean BMI of 30.7 ± 5.7 kg/m^2 (meaning that they were obese, according to the World Health Organization (WHO 2010)), had a mean age of 70 (interquartile range 13), and a slightly higher proportion were females (56.2%) (Brittain et al. 2022). Most patients were affected by mild disease (American Society of Anaesthesiologists physical status (ASA) 2) (72%), followed by those who had a debilitating systemic disease (ASA 3) (21%) and only a minority of people were fit and healthy (ASA 1) (6%) or severely compromised in terms of comorbidities (<1%) (ASA > 3) (Brittain et al. 2022).

Except for the period between 2020 and 2021, where the volume of surgeries reduced drastically due to the global pandemic, the number of TKR procedures has been increasing over the years in the UK (Culliford et al. 2015) as well as in other countries (Nemes et al. 2015; Leitner et al. 2018; Sloan et al. 2018; Daugberg et al. 2021). It was estimated that the number of procedures will grow by 401% by 2040 (Singh et al. 2019), and the reasons for the increase are complex to define. Some authors suggested this was due to the population ageing (Kurtz et al. 2007), others found that the increase in procedures could not be attributed to an increase in population or risk factors such as a high BMI. Nevertheless, it was proposed that another cause could be a decrease in health inequalities and easier access to healthcare (Price et al. 2018).

TKR is major surgery and patients often expect an improvement in pain levels and knee function therefore, it is important to have adequate outcome measures to quantify the change in these variables before and after the surgery.

2.4 ASSESSMENT OF TKR OUTCOMES

One of the aims of TKR surgery is to improve physical function, which is the ability to carry out everyday life tasks (Terwee et al. 2006). Various outcome measures are utilised to evaluate the effectiveness of TKR in improving pain and function, and they can be subjective, therefore based on the patient's perception, or objective, based on what can be observed and measured by a clinician or researcher with validated tools and technologies. In the past, the most used outcome measures were surgeon-reported such as clinical presentation and knee ROM, implant survivorship and assessment of radiographs after the surgery (Wylde et al. 2009). Nowadays, there is a focus on the patient's perspective and therefore, satisfaction, and Patient-Reported Outcome Measures (PROMs), investigating pain, physical function and how knee OA may affect participation, are commonly utilised to establish the success of the surgery (NHS Digital 2019).

2.4.1 Subjective outcomes

Satisfaction is a parameter utilised to determine the success of TKR surgery, but it does not give a specific indication of the perceived levels of pain and function. The level of satisfaction is usually lower for TKR if compared to total hip

replacement, where 91% of people are satisfied with the outcome (Hamilton et al. 2012a). It is not clear why the levels of satisfaction for TKR are lower than those for hip replacement (Baker et al. 2007). A systematic review and meta-analysis (Khatib et al. 2015) including nineteen studies found that patients' satisfaction ranged between 71.7-92.5%. One of the issues in measuring satisfaction is that it is not usually assessed with validated questionnaires (Klem et al. 2020). For instance, the NHS in England measures satisfaction by asking the following two questions and offering a multiple-choice answer:

"How would you describe the results of your operation?
[Excellent] [Very good] [Good] [Fair] [Poor]
Overall, how are the problems now in the hip [or knee] on which you had surgery, compared to before your operation?
[Much better] [A little better] [About the same] [A little worse] [Much worse]" (NHS Digital 2018b).

Other studies (Giesinger et al. 2015) assessed satisfaction with a different set of questions, investigating the overall satisfaction with the surgery, if the surgery improved pain and function and if the patient was willing to have the surgery again under similar circumstances. The fact that there is no systematic way to measure satisfaction may produce different outcomes depending on the questions utilised and this could constitute an obstacle in comparing results from different studies. Additionally, satisfaction was demonstrated to be influenced by a variety of factors. Scott et al. (2010) analysed satisfaction after TKR surgery in 1217 patients and found that after one year, 18,6% of them were unsure or unsatisfied with the outcome. The authors (Scott et al. 2010) observed that dissatisfaction was correlated to minimal improvement in the pain scores, the presence of pain in other joints, depression and initial expectations. Barrack et al. (2014) explored satisfaction after one to four years after TKR in 661 people and roughly 10% of them were dissatisfied; it was found that dissatisfaction was associated with socioeconomic factors as those who earned less than USD 25,000 per year were more likely to be dissatisfied. Hamilton et al. (2013) examined satisfaction in 4709 patients after TKR and found that the level of satisfaction could be predicted depending on the hospital experience the patient had, patient's initial expectations and pain improvement.

Since April 2009 the National Health System (NHS) started to collect PROMs data on a variety of joint replacement procedures in England, including TKR, to measure how effective the treatment provided by the NHS was, and to gain the perceptions of patients in the care received (NHS Digital 2017). The use of PROMs can be cost-effective (Giesinger et al. 2014) since the questionnaires only take a few minutes to be completed and they are self-administered. Certain PROMs are condition-specific, and allow to measure the knee-related symptoms, the perceived function, and how these are affected by knee OA and TKR, including but not limited to:

- The Oxford Knee Score (OKS) investigates the function of the knee in activities of daily living, also exploring how much the pain interferes with sleeping and work (Dawson et al. 1998). The OKS ranges from 0 (i.e., maximal difficulties or pain) to 48 (i.e., no issues, optimal function)
- The Knee injury and Osteoarthritis Outcome Score (KOOS) includes five sub-scales assessing symptoms, pain, activities of daily living, and the perceived physical performance during sport-related activities such as running, squatting, jumping, etc. during the previous week (Roos et al. 1998). However, a person who is not particularly active may find it difficult to answer these questions and end up leaving them blank. Each sub-score ranges from 0 (i.e., maximal difficulties or pain) to 100 (i.e., no issues)

TKR surgery succeeds in reducing pain and enhancing the patient's perceived function post-TKR. Data from NHS England showed that 81% of people had an improvement in their general health following TKR, and 93.5% of patients had an improvement in pain and knee-related function measured via the OKS (NHS Digital 2018a). Knee-specific subjective outcomes improve significantly early post-TKR (Mizner et al. 2011) and continue to improve six and twelve months after the surgery (Hamilton et al. 2012b).

One of the issues of assessing function with PROMs is that some of them tend to have a floor or ceiling effect post-TKR, meaning that a large proportion of patients reaches the lowest, or highest possible score after surgery, respectively (or the best achievable health state, function and no pain), limiting the measuring abilities of some questionnaires. Eckhard et al. (2021) examined three hundred and eighty patients one-year post-TKR and found that the KOOS Activities of Daily Living

(ADL) and Pain sub-scales had a ceiling effect in over 15% of participants, the threshold utilised to determine the presence of a ceiling effect (Terwee et al. 2007). There does not seem to be evidence of ceiling effects with the OKS (Clement et al. 2020; Eckhard et al. 2021), but the EQ-5D was found to have a ceiling effect in 84.4% of patients, while the WOMAC stiffness had a floor effect in 64.6% of ninety-eight patients one-year post-TKR (Giesinger et al. 2014).

Another aspect to be aware of when utilising PROMs is that they are significantly, and negatively influenced by factors that are not related to function, such as living in disadvantaged areas (Sanchez-Santos et al. 2018), depression and anxiety (Baker et al. 2012; Sanchez-Santos et al. 2018). Moreover, several studies showed that PROMs do not correlate well with objective function (Stratford et al. 2010; Mizner et al. 2011; Onodera et al. 2020). Often, patients who received a TKR have a significant improvement in PROMs scores, despite a decrease in knee ROM and quadriceps strength (Mizner et al. 2011), and a decline in performance-based tests (PBTs) (Mizner et al. 2011; Stevens-Lapsley et al. 2011a), measuring the physical function objectively. These findings showed that patients tend to overestimate the outcome of the TKR when PROMs are utilised as an outcome measure.

Basing the evaluation of TKR success solely on subjective outcomes does not seem to give a clear picture of the objective function and therefore, several researchers have recommended reporting PROMs scores together with objective measures of physical function such as knee ROM, knee muscles' strength, and PBTs (Dobson et al. 2013; Singh et al. 2017; Capin et al. 2022) (Figure 3).

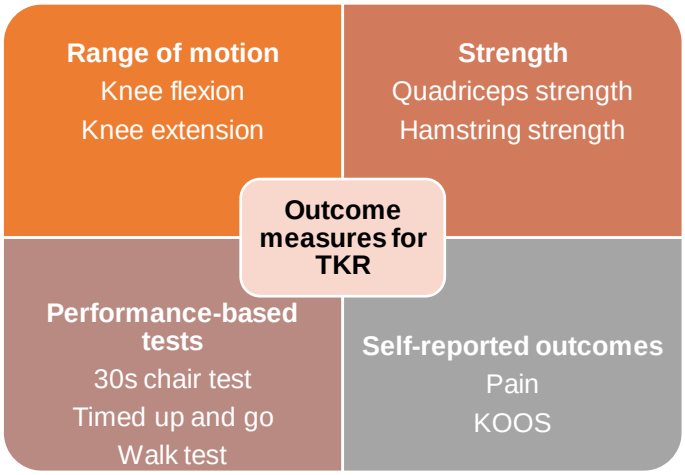


Figure 3 Minimum set of outcomes to be reported following TKR. Modified from Capin et al. (2022).

2.4.2 Objective outcomes: measures of physical function used in the clinic

A recent systematic review of the literature found that the most frequently reported measure of function in randomised-controlled trials post-TKR was knee ROM (Vajapey et al. 2020). This metric is measured by surgeons pre and post-procedure in the operating theatre and can be easily assessed in any clinical setting with goniometry to evaluate the presence of limitations in extension or flexion that can be addressed by rehabilitation. Earlier studies found that full knee extension may not be achieved early after surgery, but it is reached about six months post-TKR (Schache et al. 2019). Compared to pre-surgery, knee flexion was found to be reduced in the early stages of recovery post-TKR (Hamilton et al. 2012b; Giesinger et al. 2014), and to improve to 115-120 degrees approximately six months post-surgery (Giesinger et al. 2014; Schache et al. 2019), and reach roughly 120 degrees one-year post-TKR (Giesinger et al. 2014; McClelland et al. 2017), which is sufficient to carry out most daily activities in the Western culture, needing about 110° of knee flexion (Rowe et al. 2000). Previous research showed that patients post-TKR utilise only a portion of their available full knee ROM during functional activities such as lunging and squatting, probably due to impairments in neuromuscular movement strategies (McClelland et al. 2017). While knee ROM limitations have been proven to be associated with the speed of stair negotiation and gait (Christensen et al. 2022), it is not a direct measure of overall physical function, whereby PBTs, measuring the speed or number of repetitions of a task during everyday life activities, may be more suitable for this purpose.

PBTs are not commonly utilised to evaluate the outcomes post-TKR in clinical practice yet, partially because of time or equipment constraints (Christensen et al. 2022), but possibly also because their use was considered optional by the 1997 OMERACT (Outcome Measures in Rheumatology) recommendations (Bellamy et al. 1997). Therefore, only 14.7% of the randomised controlled trials assessing treatment outcomes in knee and hip OA adopted PBTs between 1997 and 2017 (Smith et al. 2019). It was only in 2013, following a systematic review of the literature and expert opinion, that the Osteoarthritis Research International (OARSI) suggested utilising PBTs, alongside PROMs (Dobson et al. 2013), to measure physical function objectively in people with OA or after joint replacement surgery. There is a wide variety of PBTs to assess walking, balance, and muscle strength. Moreover, there is

a variation in how a task may be assessed, for instance, a sit-to-stand task may be evaluated via the 30s chair test, or by the five-times sit-to-stand (Coleman et al. 2020). In order to standardise the recommended assessment of objective function, OARSI suggested utilising five tests, where three core ones should be utilised as a minimum (Dobson et al. 2013), as shown in Figure 4, including:

- Timed up-and-go (TUG): the time needed to stand up from a chair, walk to a line three meters away, turn around, go back to the chair and sit down; initially developed to test the risk of falls in elderly people and people with Parkinson's disease
- Six-minute walk test: the maximum distance covered in six minutes; more commonly utilised to test the physical function in respiratory and cardiological dysfunctions alongside the perceived exertion during the test
- 30s chair test (30sCT): the maximum achievable number of sit-to-stand repetitions in thirty seconds; a test related to quadriceps muscle strength
- 40m fast-paced walk test: the maximum walking speed achieved to cover a distance of 40 m
- Stair climb test: the time needed to negotiate stairs as fast as possible.

Figure 4 Recommended Set of PBTs for knee osteoarthritis and total knee replacement (modified from Dobson et al. 2013).

The objective outcome assessed via PBTs showed that patients' scores tend to decline one-month post-TKR, but then improved significantly three- and six-months post-surgery, although the recovery was not clinically relevant (Stevens-Lapsley et al. 2011a). Some researchers found that there was no significant improvement in the TUG and six-minute walk test pre to three months post-TKR (Stratford et al. 2010). Additionally, it was found that PBTs scores improved significantly one-year post-TKR, but only by small amounts (Naili et al. 2017b; Tolk et al. 2017; Christensen et al. 2022), and were not comparable to non-pathological subjects (Naili et al. 2017b).

A growing body of evidence has been produced on the measuring properties of the PBTs, and while it was suggested that most of these measures are valid, reliable, and responsive (Dobson et al. 2012; Kroman et al. 2014; Dobson et al. 2017; Coleman et al. 2020), the evidence is still developing for evaluating people with OA and post-TKR (Kroman et al. 2014; Alghadir et al. 2015; Huber et al. 2016; Sharma et al. 2023). Some researchers contested the use of PBTs on the basis that they do not measure physical function and are not responsive (Tolk et al. 2017). Tolk et al. (2017) found that none of the three core tests suggested by OARSI had good construct validity and suggested not utilising the PBTs proposed by OARSI in clinical practice. However, the hypothesis testing for validity was mainly based on comparisons with the KOOS, OKS, and Physical Function-Short Form which are all subjective measures whose scores are mostly influenced by pain levels rather than what can be achieved objectively (Boonstra et al. 2008; Mizner et al. 2011). It is well established in the literature that the relationship between subjective outcome measures and PBTs is often poor both pre- and post-TKR (Gandhi et al. 2009b; Stratford et al. 2010; Mizner et al. 2011; Hamilton et al. 2012b; Graff et al. 2016). With the suggestion that PROMs and PBTs may measure different constructs of function, OARSI recommended that subjective and objective outcome measures in severe knee OA and post-TKR should be utilised together (Dobson et al. 2013).

It may be argued that the measurement properties of PBTs should be tested against other measures of objective function, rather than PROMs, however, this has proven difficult since no gold standard has been identified to measure function objectively (Dobson et al. 2013; Graff et al. 2016). Limited research was carried out comparing objective outcomes and certain PBTs. It was found that there was an

association between a slower stair climb test and a more restricted knee ROM and between a slower 40 m fast-paced walk test and a reduced knee ROM and quadriceps strength in people with severe knee OA (Christensen et al. 2022). Knee ROM and quadriceps strength measure limited aspects of physical function. To date, there is only one research group who investigated the relationship between gait biomechanics and PBTs. Naili et al. (2017) compared forty people with knee OA on the waiting list for TKR to twenty-five healthy controls. Participants were asked to perform the single-leg mini squat, TUG and five-time sit-to-stand, and all volunteers were involved in gait analysis. The gait deviation index (GDI), evaluating gait kinematics of the lower limb, and GDI-kinetic, only observing the kinetic variables in the lower limb, were utilised as objective outcome measures for evaluation against the PBTs. The GDI is a summary measure of the gait kinematics collected from the joints of the lower limbs during the whole gait cycle. The GDI is a score that quantifies how much a gait pattern deviates from that of a reference group of healthy subjects. The authors did not find any correlations between the PBTs examined and GDI (i.e., gait kinematics) but observed a moderate association between PBTs and GDI-kinetic, showing that the further the gait kinetics deviated from the reference group, the longer it took to complete the TUG and the five-time sit-to-stand (Naili et al. 2017a).

OARSI clearly stated that the PBTs do not completely fulfil all the principles for measuring knee and hip function and further research is needed to enhance the evidence on their clinical measurement abilities in people with OA (Dobson et al. 2013). Given the age range of people undergoing TKR, there may be the presence of comorbidities and other musculoskeletal dysfunctions (Gandhi et al. 2009b) that may affect the outcomes of the PBTs. Moreover, the tests are performed one single time and they may be affected by temporary factors such as motivation while doing the test or other conditions present at the moment of testing (tiredness, pain, etc.) (Myers et al. 1993).

An aspect to be aware of is that PBTs measure the time required to accomplish a task and they do not provide data regarding the joint function. It could be argued that carrying out an activity faster after TKR, does not inform on the quality of movement. People who underwent TKR may walk faster after the recovery but may also still be affected by limping, pain and instability and this means that their

function would still be impaired (Bolink et al. 2015a). PBTs would not be able to inform on these aspects.

2.4.3 Measures of objective function in research

Three-dimensional (3D) gait analysis through motion capture technologies (i.e., stereophotogrammetry) has been an informative tool aiding in the understanding of outcomes post-TKR to inform clinicians on what level of recovery can be achieved. Although TKR improves gait biomechanics, several functional deficits persist, and often gait parameters are not comparable to NPs (Hubley-Kozey and Astephen Wilson 2018). A systematic review on gait analysis of patients who received a TKR found that, much like pre-surgery, subjects tend to walk slower at self-selected speed (0.8-1.1 m/s) (McClelland et al. 2007b), have a reduced knee ROM and reduced flexion during stance in comparison to healthy controls (McClelland et al. 2007b; Yoshida et al. 2008; McGinnis et al. 2013). Moreover, only between 20-36% of patients tend to walk showing a biphasic knee flexion-extension moment pattern and this means that the majority of patients have either a quadriceps overuse behaviour (flexion moment during single-leg support) or quadriceps avoidance behaviour (extension moment during stance) (McClelland et al. 2007b). Several studies found that compensations are still present at the hip and ankle, with reduced hip external adduction moments (Yoshida et al. 2008), limited hip extension (Saari et al. 2005b), and increased ankle dorsiflexion (Levinger et al. 2013).

Most of the existing literature looked at differences in maxima and minima values (discrete values) or joint ROM between people with knee OA and post-TKR and NPs. However, this limits the analysis to a reduced amount of information, potentially discarding important features. To overcome this limitation, several researchers, including our research group, aimed to examine gait features throughout the whole gait waveform utilising principal component analysis (PCA) (Astephen et al. 2008b; Linley et al. 2010; Federolf et al. 2013; Astephen Wilson et al. 2015; Hatfield et al. 2015; Brenneman and Maly 2018; Costello et al. 2021). This technique, described in section 3.7, allows identifying differences in magnitude, ROM and phase shifts between subjects, without having to choose arbitrary values or discarding possibly important information (Robertson et al. 2013). Within our research group, PCA has been combined with the Cardiff classifier (Jones 2004).

The method is based on the Dempster-Shafer theory of evidence (Jones et al. 2006), and it allows discriminating the gait biomechanics of severe knee OA against a reference group of NPs, based on 3D gait data that includes joint kinematics, joint kinetics and ground reaction forces (GRF) across the whole gait cycle and displays the classification into a simplex plot that can be easily interpreted.

Previous studies showed that the Cardiff Classifier can successfully discriminate between NP and people with knee OA with high accuracy (Jones et al. 2008; Worsley 2011; Metcalfe 2014; Biggs et al. 2019a; Whatling et al. 2022) and identify a change in gait biomechanics before and after total hip replacement surgery (Biggs et al. 2021) and pre to post-TKR, showing that gait function does not return to normal (Whatling 2009; Watling 2013; Worsley et al. 2015; Biggs 2016; Metcalfe et al. 2017; Biggs et al. 2019a; Whatling et al. 2022). Interestingly, Biggs et al. (2019b), found that one-year post-TKR, only a third of the gait features analysed via PCA improved, and did not entail the knee joint, but the hip adduction moment and internal-external rotations, ankle dorsiflexion moment and the three components of the GRF.

The Cardiff classifier seems to be a promising measure of objective function, the categorization of gait data proves valuable for both diagnosis from 3D motion data, and monitoring of changes pre- to post-intervention. Moreover, a recent study showed that the change in overall gait biomechanics measured via the Cardiff classifier correlated strongly with the change in the OKS one-year post-TKR (Biggs et al. 2019a). This finding was in contrast with previous literature observing a poor correlation between PROMs and objective function measured by PBTs and mentioned earlier (Gandhi et al. 2009b; Stratford et al. 2010; Mizner et al. 2011; Hamilton et al. 2012b; Graff et al. 2016). There is limited evidence comparing gait biomechanics to PROMs, that could aid corroborating the results of Biggs et al. (2019a), and most studies focussed on early stages of recovery post-TKR. Liebensteiner et al. (2008) found strong correlations between pelvic obliquity, temporospatial parameters and PROMs three months post-TKR. More recently, Bonnefoy-Mazure et al. (2017) found only a poor relationship between PROMs and cadence three months post-TKR and PROMs and knee ROM during gait one year after the surgery. Similarly, Senden et al. (2011) observed a weak relationship between gait spatiotemporal parameters and PROMs three months after TKR.

It was suggested that gait biomechanics measured with a summary measure of gait quality, as opposed to timed activities, may be more closely related to what patients perceive (Biggs et al. 2019a). Nevertheless, Naili et al. (2017) only found a weak relationship between the GDI-kinetic and the KOOS ADL and Quality of Life and no association between GDI and PROMs. The GDI (Schwartz and Rozumalski 2008) and GDI-kinetic (Rozumalski and Schwartz 2011) combine 3D gait data and were initially developed to give an estimation of the gait quality in children with cerebral palsy, establishing their deviation from a reference group of NP children. The GDI and GDI-kinetic utilise nine joint kinematic variables and nine joint kinetic variables across the gait cycle, respectively, to calculate an index. Previous research found that the GDI was able to identify significant gait deviations in people with hip OA (Rosenlund et al. 2016b), lower limb rheumatoid arthritis (Esbjörnsson et al. 2014), and knee OA (Naili et al. 2017a; Kobsar et al. 2019; Naili et al. 2019a). Additionally, the GDI was utilised to measure gait quality change pre- to post-hip replacement (Jensen et al. 2015a; Naili et al. 2019b) and pre-to post-TKR (Naili et al. 2017b). The GDI-kinetic was adopted to estimate gait kinetics deviations in people with knee OA (Naili et al. 2019a) and to assess the gait improvement pre- to post-total hip replacement (Naili et al. 2019b) and pre to post-TKR (Naili et al. 2017b). One of these studies (Naili et al. 2017b), found that at one-year post-TKR, patients' GDI was comparable to those of NPs, while joint kinetics evaluated via the GDI-kinetic did not go back to normal. This contrasts with what was observed when the Cardiff Classifier was used as a measure of function pre to post-TKR, showing that most patients do not show lower limb biomechanics comparable to NPs after the operation (Worsley et al. 2015; Biggs et al. 2019a; Biggs et al. 2019b). The relationship between GDI or GDI-kinetic and the Cardiff classifier has never been explored. Having a better understanding of the measuring properties of outcome measures of function and which is the most adequate to measure the degree of improvement achieved through surgery, may aid in setting realistic expectations for the patient.

2.5 FACTORS RELATED TO TKR OUTCOMES

Individuating the factors that may lead to a poor outcome, may help address them in early stages to improve the levels of pain, function and satisfaction after the

surgery. Various approaches could be adopted to understand what may influence the TKR outcome. Some of the existing literature developed regression models, other researchers looked at associations between certain factors and function, while some reports analysed differences in TKR outcomes by sex, implant design, etc.

A growing emphasis has been placed on the patient's view on the quality of life and function following TKR and therefore, PROMs are now routinely utilised within the UK NHS before and after TKR to determine its outcome. Hence, most of the existing literature seems to focus on predicting PROMs rather than objective outcomes. Because of this, the main existing literature on TKR outcome prediction will be discussed below, regardless of whether the outcome explored was objective or subjective, to highlight what may be the factors affecting the outcome and aid in defining the ones to be explored in the current thesis.

Irrespective of the outcome measure utilised, the recurring finding of previous research seems to be that preoperative and post-TKR function are correlated, where a lower baseline score is predictive of a worse postoperative function. Baseline WOMAC values were predictive of scores one week (Kennedy et al. 2006), six months (Fortin et al. 1999), and two years (Fortin et al. 2002) post-TKR. The times needed to complete the TUG were correlated pre to twelve weeks post-surgery (Gandhi et al. 2009b). (Kennedy et al. 2006) found a correlation between the distances covered during the 6MWT pre to post-TKR. In line with the existing literature, (Watling 2013) observed a trend between poor outcomes post-TKR and high preoperative Belief of OA (BOA) in a small group of patients. Additionally, (Bade et al. 2014) and (Pua et al. 2019) found that pre-operative knee flexion and extension, measured with a goniometer, were predictive of knee flexion-extension six months post-TKR. Knee ROM is a relevant measurement that could be linked indirectly to knee function, as ROM limitations may make it difficult to achieve daily tasks. When looking at the amount of improvement pre to post-TKR, the literature utilising the Cardiff classifier seems conflicting, where (Worsley 2011) found that the better the gait function pre-TKR, the smaller the change in function six months post knee arthroplasty, while (Biggs 2016) observed that the functional status pre-surgery was not related to the amount of improvement seen pre to 9+ months post-TKR.

In addition to the above, several patients' non-modifiable characteristics were found to correlate with or predict TKR outcomes. The literature looking at age is

contrasting, where some research found no association between PROMs post-TKR and age (Kennedy et al. 2006; Williams et al. 2013), or GRF and age (Kramers-de Quervain et al. 2012). However, age seems to be predictive of gait temporospatial parameters two years post-TKR (Kramers-de Quervain et al. 2012) and there are opposing findings on age as a predictor of PBTs' scores post-TKR (Kennedy et al. 2006; Bade et al. 2014).

Reports suggest that being female is associated with worse postoperative outcomes, whether subjectively or objectively measured. Being female was a predictor of poorer performance-based test scores (Kennedy et al. 2006) and smaller knee ROM (Pua et al. 2019) post-TKR. Moreover, while females were found to have worse preoperative PROMs than males, they had a larger improvement in PROMs pre to post-TKR (Liebs et al. 2011; Mandzuk et al. 2015). However, there are mixed findings regarding differences in post-operative PROMs between sexes, with some authors finding no differences between males and females at different time points post-TKR (Liebs et al. 2011), and some others observing a higher quality of life and function in males (Mandzuk et al. 2015). Looking at biomechanics, males and females seem to have different biomechanics impairments following TKR (McKean et al. 2007; Astephen Wilson et al. 2015; Paterson et al. 2018; Paterson et al. 2020), and males may tend to have a larger improvement in gait six months post-knee arthroplasty (Worsley 2011).

There has been an ongoing debate on whether a posterior-sacrificing or cruciate-retaining knee implant design may produce different outcomes. (Beach et al. 2019) found no significant differences in knee kinematics during walking between medial pivot, cruciate-retaining and cruciate-sacrificing knees. (Watling 2013) observed that patients with a posterior sacrificing implant seemed to have a better gait function one-year post-TKR than those with a cruciate-retaining knee. Some authors identified a larger knee ROM in posterior-sacrificing knees, although not clinically relevant (2 degrees) (Verra et al. 2015) and others did not (Serna-Berna et al. 2018). Moreover, earlier reports found that patient-perceived pain and function were not different among patients who received a posterior sacrificing or retaining prosthesis (Serna-Berna et al. 2018). A systematic review with meta-analysis found that patella resurfacing produced better outcomes in terms of pain and function than not resurfacing the patella (Longo et al. 2018). However, a recent NICE guideline

highlighted that there was no difference in PROMs between resurfacing the patella versus not resurfacing it (NICE 2020b).

A previous study found that having several painful joints was a predictor of slower walking speed 2 years post-TKR (Kramers-de Quervain et al. 2012). Moreover, a higher number of comorbidities were predictive of gait parameters, influencing negatively walking speed and peak vertical GRF (Kramers-de Quervain et al. 2012). In their validation study, (Riddle et al. 2022) found that an increase in the comorbidities score (modified Charlson) increased the chance of having a bad OKS outcome.

Looking at modifiable factors, the role of BMI in influencing TKR outcomes is not clear. (Stickles et al. 2001) found that obese patients had a higher risk of having difficulties negotiating stairs one year after TKR, however, obese patients had a similar change in WOMAC pre to one-year post-TKR compared to patients who were not obese. (Xu et al. 2018) found that patients with a higher BMI had TKR surgery earlier in life and had poorer self-perceived function (OKS) ten years after the surgery than patients with a BMI < 30, even when controlling for sex, age and comorbidities. (Giesinger et al. 2018) found that there was no difference in the OKS score change pre to one-year post-TKR across BMI classes.

2.6 SUMMARY

Knee OA is a disease that affects an increasing number of people, having a considerable impact on their everyday function and activities. Several biomechanical alterations can be observed in subjects suffering from knee OA. In the presence of end-stage knee OA, TKR surgery is performed to reduce pain and improve function, and the number of TKRs is expected to grow in the next decades. As with every health condition and intervention related to it, it is key to evaluate the outcome, to understand the degree of success of the intervention and to identify factors influencing the function, to address them.

Currently, the main method to assess knee outcomes in the NHS pre- and post-TKR is via questionnaires, which are subjective and explore what patients perceive they can do, rather than what they can do objectively. It is believed that subjective and objective measures of function measure different constructs, as previous research showed that they do not seem to correlate well. Therefore, there is

a growing interest in coupling subjective and objective measures of function, an approach which has been also recommended by OARSI, an international organisation bringing together healthcare professionals and researchers in the field of OA.

There are several measures of objective function for patients pre- and post-TKR; The use of PBTs may be more feasible in clinical settings, while 3D gait analysis is mainly utilised in research. The interpretation of the outputs from gait analysis can be complex and, therefore, techniques like PCA can be utilised to reduce the amount of data while retaining the most important information in it. Results from PCA can be utilised within summative measures of gait function, such as the Cardiff Classifier, developed in our research group, which returns information on the overall gait function, similar to the GDI and GDI-kinetic.

Selecting the optimal tool for assessing the objective function is crucial for advancing research and gaining insights into the expected level of improvement for patients post-TKR. Throughout this work, the Cardiff Classifier was employed to understand how objective gait function may relate to other measures of gait biomechanics (Chapter 4), to measures of objective physical function utilised in research and clinically (Chapter 6), and to the patient's perception of their functional capabilities pre-and post-TKR (Chapter 6 and 7). The ability to objectively measure function pre- and post-TKR can aid in identifying factors influencing postoperative outcomes, which was explored in Chapter 5. This knowledge can explain potential challenges that patients may encounter after having had a TKR, and aid advancements in objective outcome assessment, enhancing patient care.

Chapter 3: Assessment of function at Cardiff University

3.1 INTRODUCTION

The studies contained in Chapters 4, 5, and 6 of this Thesis are part of wider research carried out within the Biomechanics and Bioengineering Research Centre Versus Arthritis that was approved by Wales Research Ethics Committee 3 (10/MRE09/28) and Cardiff and Vale University Health Board Research and Development Committee. Data for the TKR project are part of Work Package 3 of the BBRCVA research and were collected since November 2009; Work Package 3 will be identified as the TKR study from now onwards. Sections 4.2.1-4.2.2, 5.2.1-5.2.2, 6.2.2 of Chapters 4, 5, and 6, respectively, clarify what was the primary data collected and processed by the author. Throughout the years, there has been a gradual evolution of the TKR study. This resulted in the introduction of new assessments and outcome measures, an adaptation of the protocol and methods (following the approval of the relevant ethics) and a modernisation of the equipment utilised to collect three-dimensional gait data. In this chapter, the methodology elements that are common to the studies in Chapters 4, 5, and 6 will be discussed. The methods that are specific to one chapter only will be discussed within that chapter.

Chapter 4 additionally included participants from another research Institute (Karolinska Institutet, Stockholm, Sweden), while Chapter 7 was developed solely on data that were collected at a different Institution (Australian National University, Canberra, Australia). These data were based on a different ethical approval, study protocol and equipment. Therefore, it was deemed more appropriate to discuss the participants' recruitment, data collection and processing within the relevant chapters, to avoid confusion. However, the summative measure for assessing objective function, the Cardiff classifier, is in common with all the studies presented in this thesis and is described in the current chapter. Therefore, the data collection and processing presented followingly is relative exclusively to the participants employed in chapters 4, 5, and 6 (i.e., the Cardiff University cohort). Table 1 shows for which chapters the methods have been described in the current chapter.

Table 1 Overview of the methods within the thesis

Chapter number	Method to assess joint biomechanics	Institution of data collection	Description of data collection in current chapter?	Description of data processing in the current chapter?	Summative measure to assess objective function
4	3D motion analysis	Cardiff University	Yes	Yes	Cardiff classifier
		Karolinska Institutet	No	Only the biomechanical model	
5	3D motion analysis	Cardiff University	Yes	Yes	Cardiff classifier
6	3D motion analysis	Cardiff University	Yes	Yes	Cardiff classifier
7	Computed tomography and fluoroscopy	Australian National University	No	No	Cardiff classifier

3.2 DATA MANAGEMENT AND SECURITY, CONFIDENTIALITY

All data collected for the TKR study were stored and managed in agreement with General Data Protection Regulations (European Parliament and Council of the European Union 2018) and will be kept for a minimum of 15 years, following the Data Protection Act (2018) and Cardiff University Regulations (Cardiff University 2020). A unique study number was created on the BBRCVA central Access database (Microsoft, Redmond, WA, US) and assigned to each participant for data pseudonymization. Personal identifiable data were recorded into the BBRCVA Database which is password-protected and could be accessed only by the Centre's researchers having a research contract with the relevant health board. Additionally, personal identifiable data, data produced during and after the assessment were recorded into an additional, project-specific TKR Access Database (Microsoft, Redmond, WA, US), implemented by the current author, which is password-protected. The TKR Database and digital copies of the documents with identifiable data were stored in a password-protected external drive (NAS Drive) accessible only to the researcher and staff involved in the research. The NAS Drive was located within a locked room at the MSKBRF whose access was granted through a key fob only to authorised research staff.

Paper documents were stored within a lockable cupboard whose key was in a safe, with a code known only to relevant research staff. The safe and the lockable cupboard were located within a locked room at the Musculoskeletal Biomechanics

Research Facility (MSKBRF), School of Engineering, whose access was granted through a key fob only to authorised research staff.
Motion analysis data and videos were stored in:

- NAS Drive
- The main researcher's laptop and external hard drive which were encrypted, password-protected and kept in a secure lockable place
- The MSKBRF Clinical laboratory's main computer which was password-protected. Access to the clinical laboratory was limited to the members of the research team that have a key fob to open the locked door

3.3 DATA COLLECTION

3.3.1 Participants' recruitment

All potential participants were selected through consecutive sampling and were given ample time to consider their involvement in the research. All participants recruited in the current thesis gave written consent before data collection.

NHS patients who were on a waiting list to receive a primary TKR at the Cardiff and Vale Orthopaedic Centre (CAVOC) (University Hospital Llandough, Cardiff, UK) were approached by a research officer and a brief explanation of the study protocol and aims was provided. If they were interested, they signed a Permission to Contact, allowing the researcher to contact them, and were given a Patient Information Sheet. After a few days, the main researcher or research officer based at Cardiff School of Engineering contacted the patients via phone or email. If the patients agreed to enrol in the study, they were screened for inclusion and exclusion criteria.

Patients had to be over the age of 18 and no older than 80 years old at the moment of recruitment, and to be on a list to receive a primary TKR. Patients were excluded if they had conditions affecting their movements (visual, vestibular, neurological conditions – e.g., previous cerebrovascular accident with sequelae, Alzheimer's disease, Parkinson's disease, multiple sclerosis, etc.) if they were unable to walk unaided for at least 10 meters (the length of the walkway within the laboratory) or to provide written consent. Potential participants who had had previous injuries to the lower limb that would have made the participant unsuitable for the study in the researcher's and/or clinician's opinion were excluded (e.g., fixed

fractures of the ankle). However, patients who already had a TKR in the contralateral knee, a high tibial osteotomy (HTO), or a hip replacement were included, to make the studied sample representative of the general population with knee OA. About a third of people who undergo TKR, suffer from OA on the contralateral knee (Malahias et al. 2019), and according to the National Joint Registry (National Joint Registry 2019), about one out of five patients (22,6%) suffering from knee OA, receives a TKR in both knees. Moreover, between 21% (Primeau et al. 2021) and about 30%, (Niinimäki et al. 2012; W-Dahl et al. 2012) of patients undergo TKR ten years after a high tibial osteotomy, and previous studies showed that the presence of knee OA was associated with a higher risk of hip OA and vice versa (Prieto-Alhambra et al. 2013). Patients were recruited between November 2009 and December 2022.

A convenience sample of non-pathological volunteers (NPs) was recruited from the public in the Cardiff area (Wales, UK) through a poster located within leisure centres and Cardiff University premises, word of mouth, and, following a protocol amendment (which received ethical approval), via social media posts (Twitter, Yammer), and emails to local societies and groups. Upon contact, a brief explanation of the study protocol and aims was provided. If potential participants were interested, they were given or sent a Healthy Volunteer Information Sheet via post or email. A few days later subjects were screened for inclusion-exclusion criteria once they confirmed their interest in participating. NPs had to be between 18 and 80 years of age and to be able to provide written consent. Exclusion criteria were a history of pathologies or instability of the lower limb joints, a previous knee injury, the presence of musculoskeletal pathologies and/or neurological, visual, or vestibular conditions that may have affected the way they moved. NPs were recruited between February 2007 and December 2022.

3.3.2 Three-dimensional gait analysis equipment

All data for the TKR study were collected within Cardiff University School of Engineering (Cardiff, UK) in two different laboratories, the Cardiff University Motion Analysis Laboratory (between February 2007 and mid-May 2017) and the MSKBRF Clinical laboratory (between mid-May 2017 and December 2022).

Motion capture cameras and force plates

Joint kinematic, kinetic and ground reaction force (GRF) data were collected via a motion capture system whose hardware was upgraded several times over the years, due to the evolving technology. The Cardiff Motion Analysis Laboratory had an 8 m walkway; until April 2012 the Cardiff Motion Analysis Laboratory was equipped with nine Qualisys ProReflex 1000 cameras (Qualisys, Sweden), sampling at 60 Hz, and two Bertec force plates (Bertec Corporation, Ohio, USA) (sampling at 1080 Hz) as described in the Thesis by (Whatling 2009). From about June 2012 until October 2015 the Cardiff Motion Analysis Laboratory was upgraded with eight Oqus 300+ and one Oqus 110 motion capture cameras (Qualisys, Sweden), sampling at 60 Hz, and four Bertec force plates (Bertec Corporation, Ohio, USA), sampling at 1080 Hz. Additionally, Two Sony HDR-CX130/160 video cameras (Sony, Japan) were synchronised with the motion analysis system. These were positioned in the sagittal and frontal planes of the volunteer motion and used for data verification. More details can be found in the Theses by (Watling 2013; Biggs 2016).

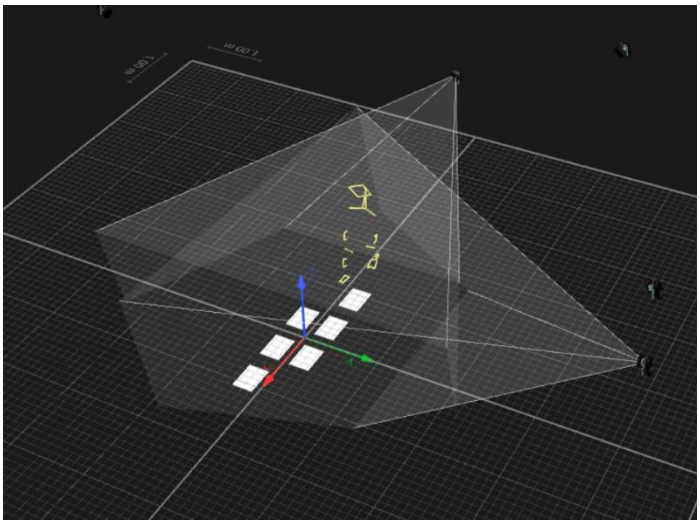


Figure 5 Representation of the clinical lab in Qualisys Track Manager. In the figure: lab coordinate system, video cameras field of view, and a participant walking on the walkway.

From mid-May 2017, 3D gait analysis was moved into the MSKBRF Clinical laboratory, which had a 10 m walkway, and it was surrounded by twelve Oqus 700+ motion capture cameras (Qualisys, Sweden) sampling at a frequency of 200 Hz. Additionally, 2 Oqus 210c high-speed video cameras (Qualisys, Sweden), were used to capture videos (24 frames per second) and were synchronised with the motion capture cameras. One video camera was mounted laterally to the walkway (subject's sagittal view) and the other one was positioned at the start of the walkway (subject's frontal view) to allow seeing the subject from two angles (perpendicular to each other, Figure 5). The walkway was equipped with six floor-embedded force plates

(600 mm x 400 mm) (Bertec Corporation, Ohio, USA) (Figure 6), and recorded data at 2000 Hz (10 samples per camera frame).

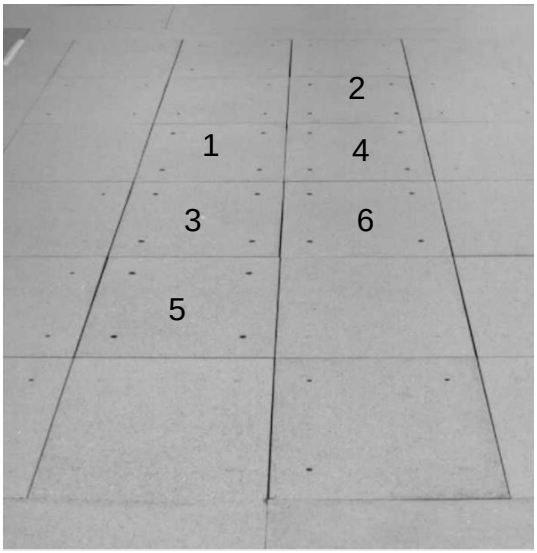


Figure 6 Force plates' location in the MSKBRF Clinical Lab and identification numbers.

Within both laboratories, the force plates had strain-gauged load transducers that allowed measuring the three components of the GRF, the moments around the x, y and z axes of the force plate and the centre of pressure (CoP) (where the force is applied, this is calculated with equations utilising forces and moments data collected by the force plate in question).

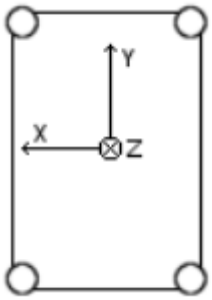


Figure 7 Internal coordinate system of a force plate.

The coordinate system of the force plate was in the centre of the surface of the force plate, the y-axis points forward, the x-axis to the left (looking in the y-axis directions) and the z-axis pointing down towards the floor (Figure 7). The calibration matrix for the force plate was stored within each plate. All the force plate amplifiers were connected to a trigger button that allowed resetting the load to zero on all the force plates at once.

Marker set

The retro-reflective markers (14 mm diameter) were positioned onto the relevant body landmarks using a medical-grade, double-sided, hypoallergenic tape, to reduce the risk of skin irritation. Marker clusters were secured in place with reusable elastic bands (pre-COVID-19) or with self-adhesive, single-use elastic

bands (Peha Haft®, Hartmann, Heidenheim an der Brenz, Germany) (following COVID-19 risk assessment amendments to reduce the spread of infections) and single-use elasticated tubular support bandages (Tubigrip™, Mölnlycke Healthcare, Gothenburg, Sweden) that were cut to size, to prevent the marker clusters from migrating during the assessment.

The marker set for the assessment of the lower limb evolved and all versions were based on the Calibrated Anatomical System Technique (CAST) marker set (Cappozzo et al. 1996a) which uses twenty markers and four rigid marker clusters (four markers each) on the lateral aspects of the thighs and legs and is commonly used in motion analysis. Originally, this marker set included two additional markers, one on each greater trochanter, which allowed to create the thigh segment (as described in (Whatling 2009; Watling 2013; Biggs 2016). Since mid-May 2017, the greater trochanter markers were removed and a marker was attached on each iliac crest (total markers n = 22) to have availability of additional markers to track the movement of the pelvis, when in certain instances (e.g., arm over the pelvis during walking), the ASIS markers may have been occluded from the view of the camera. This method was utilised in a previous study (McClelland et al. 2017) and had been formerly validated (McClelland et al. 2007a). Figure 8 shows the most recent marker set adopted since December 2019.

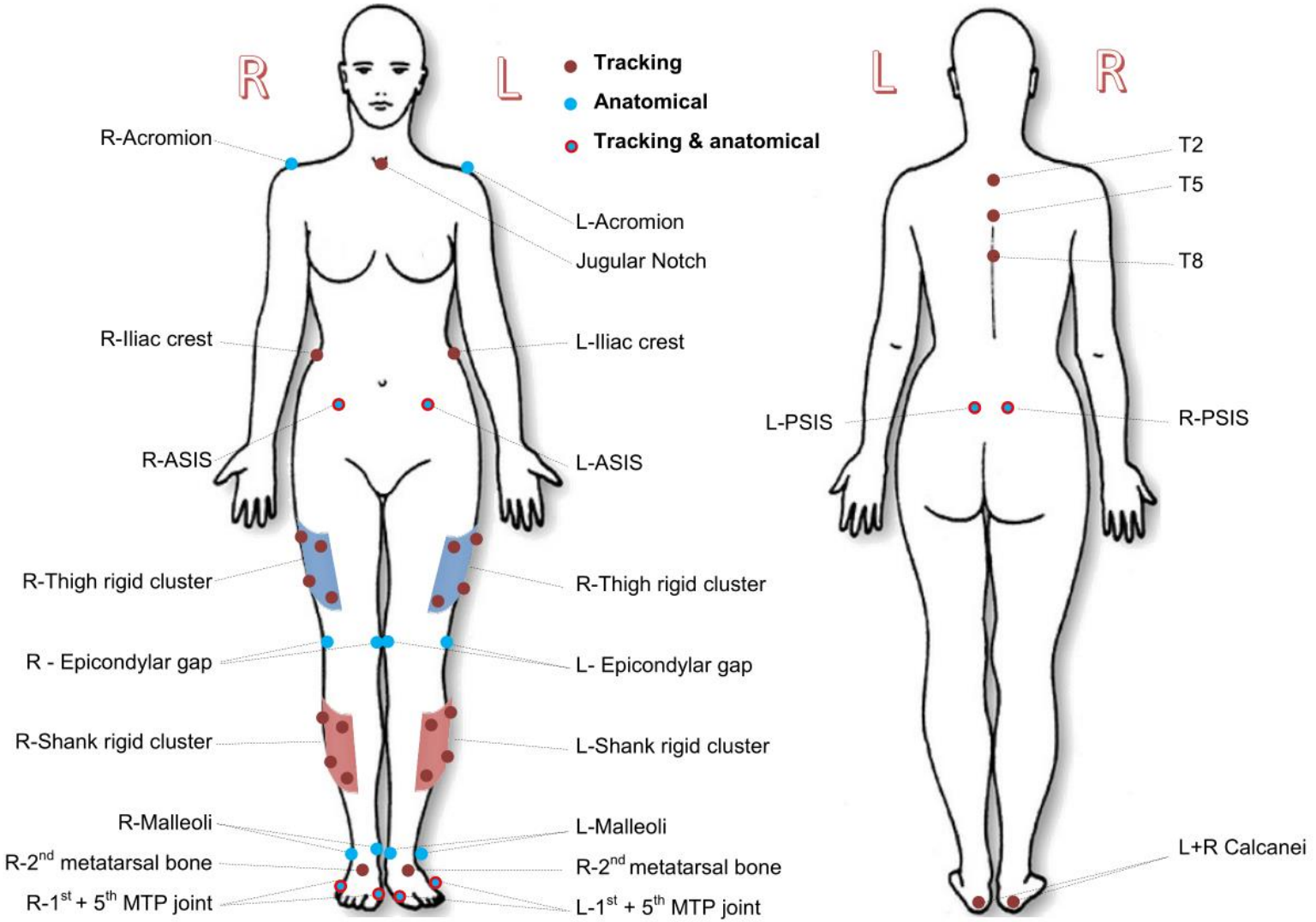


Figure 8 Marker set. R: right; L; left; MTP: metatarsophalangeal; T: thoracic vertebra; ASIS: anterior superior iliac spine; PSIS: posterior superior iliac spine.

3.3.3 Data collection protocol

NPs attended a single session, while TKR participants attended a maximum of four sessions, between a few days to a month before the knee surgery, at three, six and twelve months after TKR. All sessions lasted a maximum of three hours and data were usually collected between 9:00 am and 5:00 pm. Most of the data were collected in the morning.

On the data collection date, the laboratory heating was turned on to allow recording data at a comfortable room temperature (at least 20˚ C). The equipment and force plate amplifiers were switched on to warm up, following the manufacturer's recommendation. Qualisys Track Manager (QTM) (Qualisys, Sweden) was the software interface utilised for collecting motion capture and force plate data. The data collection protocol was earlier illustrated by Watling (2013) and Biggs (2016) and will be described followingly.

Laboratory calibration

The Cardiff University Motion Analysis Laboratory calibration was previously described by Watling (2013). The focus of the following description is on the MSKBRF Clinical laboratory. An L-shaped frame equipped with four markers (14 mm) was placed around force plate 1 (Figure 9).

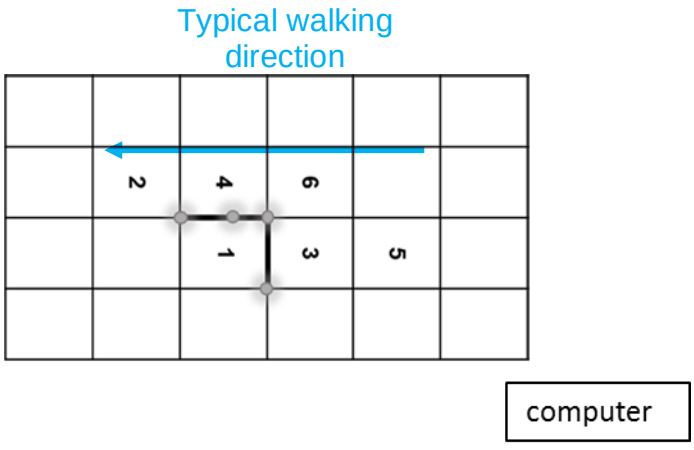


Figure 9 Representation of the MSKBRF Clinical lab, the force plates configuration and the L-shaped frame placed on force plate 1 for the lab calibration.

The long arm of the L-frame pointed in the direction of walking and the short arm pointed to the left in the MSKBRF Clinical laboratory. The calibration happened in two steps:

1. The camera system was calibrated using a dynamic calibration
2. The position of each force plate relative to the laboratory global coordinate system (GCS) was determined.

The laboratory calibration, lasting 60 seconds, was carried out with a carbon fibre wand that had a T shape and was equipped with two markers at a known distance between them (601.4 mm in the one utilised in the MSKBRF Clinical laboratory). The wand was waved in correspondence with the force plates and within the volume where the participant was recorded for the assessment. The dynamic calibration allowed to:

1. Define the laboratory GCS' origin (centre of the marker at the corner of the L-frame) and orientation. The long arm corresponded to the x-axis of the GCS, the short arm to the y-axis and the z-axis was virtually created by QTM as the cross product of the y and x-axes, pointing towards the ceiling
2. Establish the position and direction of each camera relative to the GCS
3. Guarantee data accuracy within the volume of interest: the level of disagreement between the 2D marker positions detected by the camera in question and all the others had to be less than 1 mm for the calibration to be accepted

During the dynamic calibration, each camera recorded the position of the retroreflective markers in 2D. The combined data from each camera allowed establishing of the position of the markers in a 3D space and their positions relative to the GCS. The calibrated volume was examined to detect potential un-calibrated areas and if these were present, the calibration was repeated to avoid higher errors of measurements within the un-calibrated volumes. Figure 10 represents an example of the calibrated volume in the laboratory, with no un-calibrated volumes and with a representation of the GCS axes.

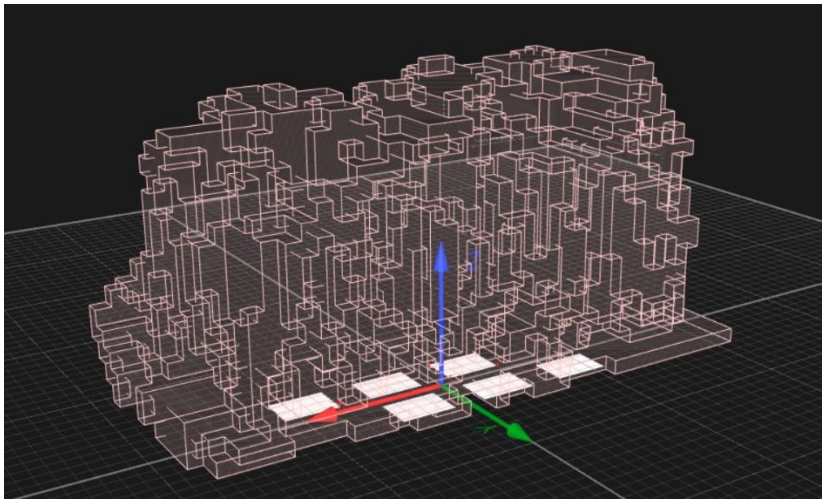


Figure 10 Calibrated volume example. The MSKBRF Clinical lab GCS is represented (red arrow: x-axis, green arrow: y-axis, blue arrow: z-axis), and the white tiles represent the force plates.

The calibration of the force plates was carried out to establish the position and orientation of the force plates relative to the GCS. In the MSKBRF Clinical laboratory, four markers were placed on one force plate in fixed locations and a 1-2 second recording was taken. Then, the orientation of the force plate was determined manually in QTM, and the software performed a transformation to determine where the force plate was located within the GCS. This was then replicated for all the force plates.

Participants' data collection

On the assessment day, the participant was welcomed into the facility and a short introduction to the laboratory and the people involved in the session was made. Health and safety measures were explained to the participant. Then, a brief explanation of the TKR study was followed by the participant's potential questions. After ensuring that the content of the Information Sheets had been read and understood, the main researcher explained the Consent Form. Informed Consent was carried out following BBRCVA standard operating procedures and under the principles of Good Clinical Practice. The participant was given plenty of time to read the Consent Form that was initialled, signed, and dated by the participant, countersigned and dated by the researcher.

Followingly, the relevant questionnaires were filled out (section 3.5) and then double-checked by the main researcher to ensure all questions were answered. In addition, past and present medical history that was relevant to the study were discussed with the participant and noted on the Participant's Datasheet.

The participant then changed into comfortable clothes (shorts and a vest top). Subsequently, height (SECA height measure, Germany) and weight (SECA scale, Germany) were measured. Moreover, knee width, depth and girth were quantified as described by Whatling (2009); Watling (2013); Biggs (2016) but these data were not used in this thesis and therefore the methods shall not be illustrated. The marker set was applied to the relevant landmarks that were identified via manual palpation by the main researcher. During the sessions, electromyographic data on lower limb muscles were collected too but since these data will not be analysed in this thesis, the methods and protocol for utilising this equipment shall not be described.

At the beginning of the 3D gait analysis assessment, several static calibrations of the duration of approximately 1-3 seconds were recorded with the participant standing still on the force plates, with their arms crossed on the chest or by the sides of the body. These were utilised during processing to define the body segments and the relative positions of the body segments and tracking markers.

To homogenize the data collection, the participant was barefooted during all trials; footwear may have different characteristics between patients, and it has been demonstrated that shoes may affect the loading of the lower limb joints (Shakoor and Block 2006). All participants walked at self-selected walking speed over the walkway equipped with force plates. The presence of "tiles" (force plates) on the walkway could be seen and the participant was encouraged to look forward rather than on the floor to avoid aiming at the tiles. Participants were not made aware of the presence of the force plates on the floor. A clean force plate hit was achieved when the heel strike and toe-off for the foot in question were comprised within the limits of one single force plate and when the opposite foot did touch said force plate. The starting point was adjusted each time to make it easier for the person to unintentionally achieve a clean force plate hit and make data collection more efficient and shorter for the participant. Lines of different colours and at different distances from the force plates were located at each end of the walkway for this purpose.

The locomotor system is highly complex, and movements are a result of the combination of motion in multiple joints with several degrees of freedom (Donà et al. 2009). Therefore, variability between trials was expected as a result of adaptations to external and internal factors. To account for this, several walking trials were recorded. At least 6 gait cycles (or clean force plate hits) were recorded for both left and right foot, this was deemed to be a reliable number for gait analysis, given previous findings (Diss 2001; Maynard et al. 2003). At the end of each recording, the participant waited a few seconds before the next trial started.

When the participant showed evident walking difficulties, a chair was placed at each end of the walkway to allow resting between recordings, and the participant was invited to walk both in the usual laboratory's walking direction (Figure 11, blue arrow) and in the opposite direction from which they came (Figure 11, green arrow), to reduce the overall distance covered during the assessment and avoid exacerbating joint pain and fatigue.

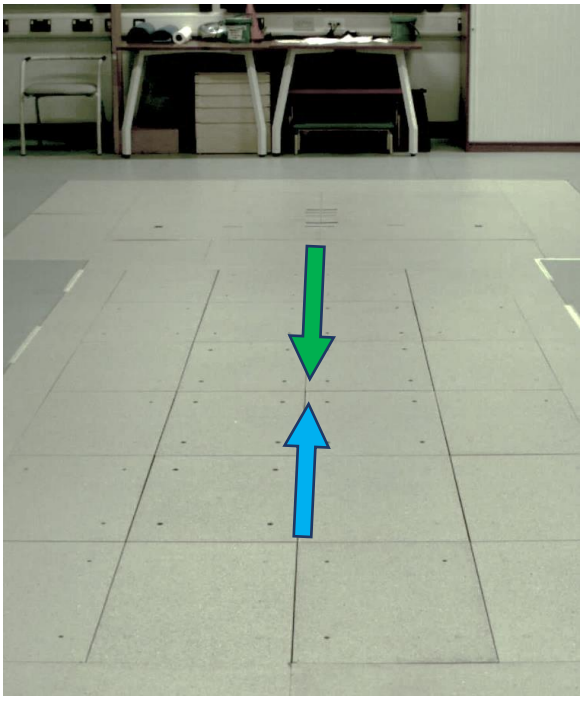


Figure 11 MSKBRF lab's walkway. Usual walking direction (blue arrow), additional walking direction for subjects with walking difficulties (green arrow).

Between tasks that required a change of the laboratory setting, the participant was invited to take a seat while the researcher and assistant(s) prepared the equipment. Sit-to-stand and stair negotiation were assessed during the data

collection. However, these data were not presented in the current thesis and the methods shall not be described.

3.4 SOURCES OF ERROR IN 3D MOTION ANALYSIS

Excluding errors in calibrating the laboratory, the main sources of error in motion analysis derive from to the misplacement of anatomical markers; this is due to the presence of tissue over the anatomical landmark, to the fact that landmarks are not always clear, sharp prominences but may have a larger area, to the operator's inexperience or technique utilised to identify the anatomical landmarks (Della Croce et al. 2005). To limit this, the researchers collecting data were trained on locating the correct anatomical landmarks and proceeded to data collection only once signed off by the most experienced researchers. The assumption with 3D motion analysis is that the movement of the bones corresponds to that of the markers placed on the participant's skin. However, considering that the skin stretches and moves during joint rotations, that muscles contract and relax producing a tissue displacement and that together with fat tissue they are subject to inertia, the result is that markers will move independently from the bone, producing soft tissue artifacts (STA) that are inevitable and have been a known source of error in motion analysis (Cappozzo et al. 1996b). Research has been carried out to estimate the STA utilising invasive methods such as retroreflective markers fixated to the bones, showing that the largest errors are at the thigh (Leardini et al. 2005). Less invasive methods to accurately measure the dynamic motion of a joint in vivo are magnetic resonance imaging, or fluoroscopy, which employs X-rays, but both are time-intensive and limited to examining one joint at a time (Wang et al. 2023).

A solution to limit STA in motion analysis is placing markers on bony landmarks to identify the coordinate system for that segment, and then track the movements of that segment via the markers on the marker clusters that are placed on the lateral aspects of the thighs and legs, where the STA is more modest. Errors may originate from this approach, where the assumption is that the relative distance between the markers on the bony landmarks and the marker clusters is fixed. This may not be the case and lead to errors in joint kinematics and kinetics. To limit this, marker clusters were strapped securely on the participant's limb and in some cases, double-sided hypoallergenic tape was utilised in addition to the elastic strap. Earlier

research showed that 3D gait analysis has good reliability in the sagittal plane kinematics, good, but inferior reliability in the frontal plane, conversely, the transverse plane has moderate reliability (McGinley et al. 2009). Errors in the sagittal plane are less than 4°, and in the frontal plane are usually around 2°, which is considered to be acceptable (McGinley et al. 2009).

3.5 PATIENT-REPORTED OUTCOMES

3.5.1 Knee Osteoarthritis and Injury Outcome Score (KOOS)

The KOOS (Appendix A) is a subjective outcome measure including 42 questions that are widely used in the clinic and research (Appendix A), with a recent systematic review finding that it is the second most reported outcome in randomised controlled trials (Vajapey et al. 2020). The KOOS takes approximately ten minutes to be completed, and it evaluates five domains, including Pain (nine questions), Symptoms (seven questions), Activities of Daily Living (ADL) (seventeen questions), Sports and Recreational activities (five questions) and knee-related quality of life (QoL) (four questions) (Roos et al. 1998). Each answer has a five-point Likert scale, ranging from 0 to 4. Each KOOS subscale raw score is then converted, and ranges from 0, meaning extreme knee issues, to 100, no knee issues. The KOOS is a reliable questionnaire, as demonstrated by the study of Collins et al. (2011) including subjects with knee OA which found an interclass correlation coefficient (ICC) for pain ranging between 0.80-0.97, Symptoms from 0.74 to 0.94, ADL between 0.84-0.94, sports and recreational activities from 0.65 to 0.92, QoL between 0.60-0.91. The KOOS was responsive to change in patients who underwent TKR surgery, with the QoL sub-score being more responsive than the ADL and Pain sub-scores (Collins and Roos 2012). The responsiveness of the KOOS was found to remain consistent at different time points post-TKR (three, six-, and twelve months post-surgery) (Collins et al. 2016). The minimal detectable change in people with knee OA was 15.5 points for Symptoms, 13.4 for Pain, 15.4 for ADL, 21.1 points for QoL, and 19.6 for Sports and Recreational Activities (Collins et al. 2011).

3.5.2 Oxford Knee Score (OKS)

The OKS (Appendix B) is a valid and responsive questionnaire developed to assess the outcomes of TKR surgery (Dawson et al. 1998) and is routinely utilised in

the NHS in the UK to evaluate the patient's perception regarding total knee arthroplasty (NHS Digital 2017). It is a questionnaire including 12 items, five of which evaluate the patient-perceived function, and the remaining ones assess pain in the past four weeks (Appendix B). Each question has a five-item Likert scale, from 0, severe pain and difficulties, to 4, no issues. The overall score is the sum of the scores from each question, ranging from 0 to 48 (most severe symptoms). The questionnaire does not seem to suffer from floor or ceiling effects pre- or post-TKR (Harris et al. 2017). The OKS is responsive (Ko et al. 2013), reliable in measuring patient-perceived pain and function (Xie et al. 2011), the overall scale (ICC 0.93), function (ICC 0.92) and pain (ICC 0.91) have good test-retest reliability in individuals with knee OA (Harris et al. 2013). A recent study on 694 487 patients who received a primary hip or knee replacement, or revision found that the minimal important change after primary TKR was 10.5 points (Sabah et al. 2022).

3.6 DATA PROCESSING

3.6.1 Labelling trials

QTM (Qualisys, Sweden) was utilised to identify and label marker trajectories for all static calibrations and trials. Initially, an "Automated Identification of Markers" (AIM) model had been developed by Biggs (2016) using dynamic trials, and perfected by the author during the processing of data that were collected from the start of the PhD project, or that had not been processed yet. Additionally, a new AIM model was developed by the author to include an additional body segment to process the data described in Chapter 6. More precisely, each walking trial was manually labelled and then utilised to create and train an AIM model, which allowed to identify and label markers' trajectories automatically, thus, reducing the processing time. The latest labels' names used to identify the markers' trajectories in QTM are displayed in Figure 12.

Gaps in marker trajectories could occur when a marker was occluded for a certain amount of time or when a trajectory had not been correctly identified by the AIM model. Each marker trajectory graph was visually inspected to identify labelling errors and gaps. All the trial files where the markers trajectories had gaps not related to an incorrect application of the AIM model, were processed with the feature "polynomial gap-filling". The polynomial gap-filling option uses a procedure

(interpolation) that estimates the missing trajectories by connecting the x, y, and z trajectories before and after the gap. This implies that there must have been data on both sides of the gap. If a trajectory gap was larger than 10% of the sampling rate, the trajectory was left unfilled (i.e., if the camera's sampling rate was 200 frames per second, the maximum width of a gap that could be filled was 20 frames). This was to limit the fabrication of data. Subsequently, each trial was visually inspected to verify that all markers had been properly identified and that the gaps had been filled correctly. When the trajectory with large gaps was that of a tracking marker, a note was made to avoid using this marker in the next stage of the data processing. Each of the correctly labelled files was exported in a .c3d format to process data in Visual 3D (V3D) (C-Motion Inc., Maryland, USA).

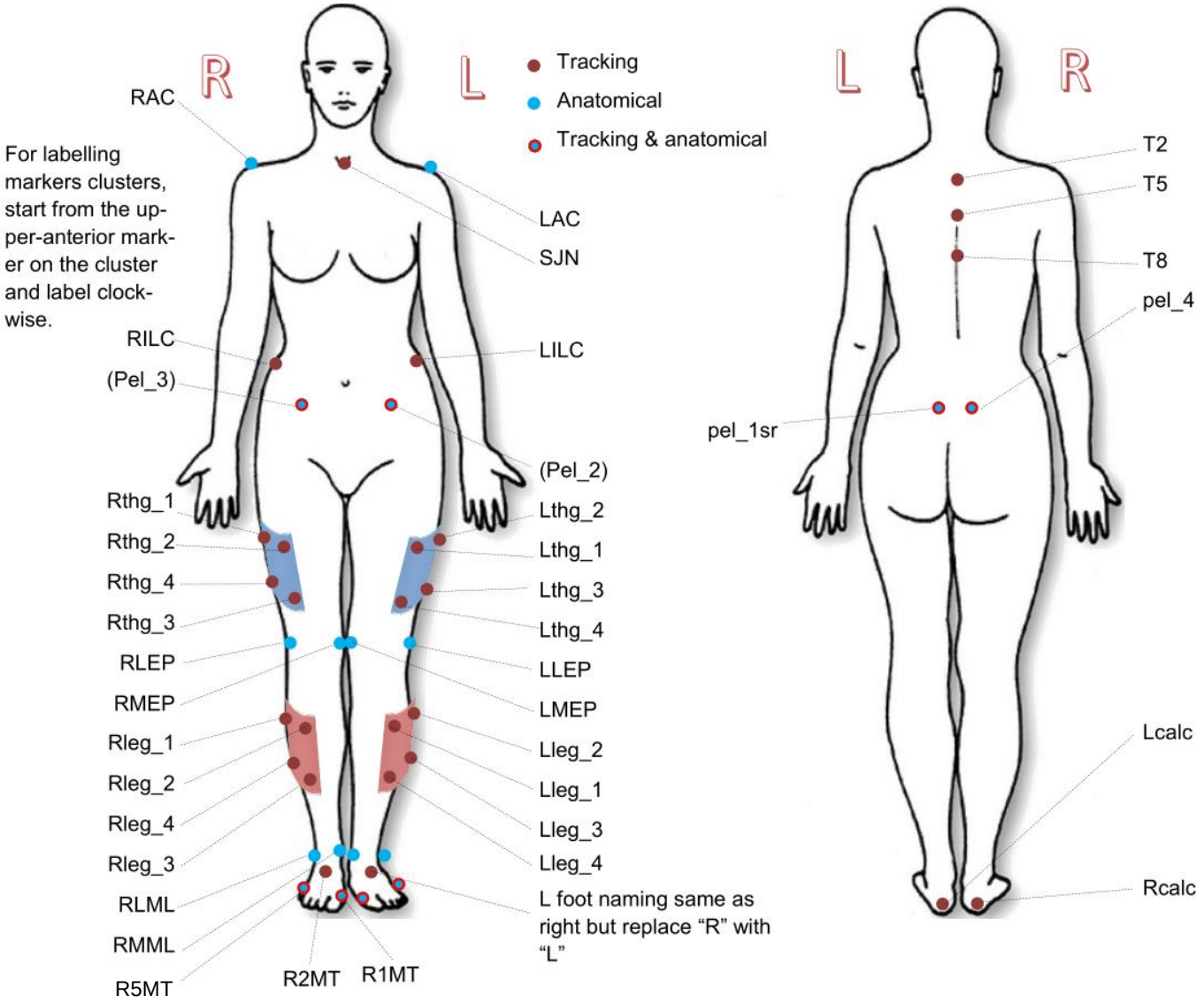


Figure 12 Markers labels used in QTM for body markers and Pointer markers. R: right; L; left; ASIS: anterior superior iliac spine.

3.6.2 Calculation of biomechanics data in gait trials

V3D (version 6, C-Motion Inc., Maryland, USA) was used to filter the raw marker and force plate data, build a biomechanical model of the participant, and calculate the spatiotemporal parameters, the GRF and joint biomechanics. All the previous steps were approached in a standardised manner, via a bespoke pipeline that was initially developed by Whatling (2009), refined by Watling (2013), then Biggs (2016), and lastly, by the author. Specifically, the model template and pipeline adopted for the data utilised in Chapters 4 and 5 corresponded to those utilised and described by Biggs (2016), with the addition of the foot progression angle calculation (added by the author to the existing pipeline). The author added the trunk segment and the calculation of joint angles for this (described in section 6.2.3) to the pre-existing model template and pipeline, to allow for processing of the data described in chapter 6.

Biomechanical model

Each segment was modelled as a rigid body and followed the International Society of Biomechanics (ISB) recommendations for the lower limb and the trunk (Wu et al. 2002). All the segments' coordinate systems were oriented so that the x-axis was mediolateral and pointing to the right, the y-axis was anterior-posterior pointing forward, and the z-axis was axial pointing upwards. The origin of each coordinate system was proximal to the segment. The biomechanical model had six degrees of freedom, allowing for three joint rotations and three joint translations and, therefore, the movement of the joint centre of rotation. Anatomical markers and virtual landmarks were utilised to determine each segment coordinate system orientation, while tracking markers (those on the marker clusters), and tracking-and-anatomical markers (those on the pelvis and trunk), were utilised to track the movement of the segments.

Modelling the Pelvis. In our research group (Watling 2014; Biggs 2016), two pelvis segments were used for biomechanical calculations, a Visual 3D Pelvis (used for kinematic calculations only), and the CODA pelvis (used for kinetic calculations only). The Visual 3D Pelvis segment was defined using PSIS markers, and virtual markers on the iliac crests. The main feature of the Visual 3D Pelvis segment was that its tilt in the sagittal plane was zero. However, the pelvis tilt relative to an

imaginary line parallel to the floor, most commonly defined as anteversion or retroversion angle (i.e., anteriorly or posteriorly tilted, respectively) is regularly accounted for in clinical assessments. Not considering the naturally occurring pelvis anteversion, would have resulted in calculating smaller hip joint flexion angles. Therefore, in the current thesis, the pelvis segment was developed using the PSIS and ASIS markers to create the CODA pelvis, commonly utilised in gait assessments. This segment identifies a pelvis that has sagittal tilt and hip joint angles were calculated relative to this. The origin of the coordinate system was the midpoint between the ASISs, the anterior-posterior axis x lie between the ASIS and the mid-point of the PSIS. A mediolateral axis pointed in the direction of the right ASIS, perpendicularly to the x-mediolateral plane, was the z-axis, and the cross product of the x and z axes gave the y-axis. This segment was used for both kinetic and kinematic calculations.

Modelling the thigh. The origin was the hip joint centre, and the distal reference points were the medial and lateral knee markers. Determining the hip joint centre location is challenging due to its anatomical position deep into the pelvis which cannot be palpated, and it has to be created artificially to build the thigh segment. Several equations based on MRI studies have been proposed to determine the hip joint centre utilising the pelvic markers as a reference (Bell et al. 1990; Davis et al. 1991; Harrington et al. 2007). The hip joint centre was created using Harrington 1 regression equations (Harrington et al. 2007), in line with previous work of Biggs (2016). Equations 1-2-3 calculate the right hip joint centre. Pelvis depth in Eq. 2 was the distance between the ASIS mid-point and the PSIS midpoint.

$$x\ (mediolateral\ axis) = 0.33 * ASIS\ distance + 0.0073 \quad \text{Eq. 1}$$

$$y\ (antero - posterior\ axis) = -0.24 * Pelvis\ depth - 0.0099 \quad \text{Eq. 2}$$

$$z\ (axial\ axis) = -0.30 * ASIS\ distance - 0.0109 \quad \text{Eq. 3}$$

To calculate the left hip joint centre the result for the x-axis had to be multiplied by -1. Looking at these equations, the importance of locating the ASIS correctly (both in the pelvis' coronal and sagittal plane) is evident, given that the hip joint centres position strictly depends on these measures. The only difference from the ISB recommendations (Wu et al. 2002) was that the thigh coordinate system's z-

axis was the line connecting the midpoint between the epicondylar gaps (rather than the epicondyles) and the hip joint centre. The thigh marker cluster was then used to track the movements of this segment during dynamic trials.

The choice of placing the knee markers on the epicondylar gaps rather than the epicondyles follows the work previously done at Cardiff University (Biggs 2016). Patients with end-stage knee OA often have a high BMI, coupled with the swelling of the knee joint due to inflammation; these factors make it extremely challenging to palpate the epicondyles, and positioning the markers at this level could result in marker placement errors, also due to the "wobble" of the soft, swollen tissues around the epicondyles. The role of the knee markers is to determine the orientation of the femur and the tibia in the three planes. For instance, placing the markers too anteriorly on the knee could result in constructing a model with a flexed knee. If one of the knee markers is placed too anteriorly, and the other too posteriorly, this could result in an erroneous internal or external rotation of the hip and could affect the calculations of knee adduction and abduction angles. In fact, there could be what was defined as "crosstalk" between anatomical planes, where the observed knee adduction/abduction could be a result of knee flexion.

Modelling the shank (leg). The only difference from the ISB recommendations (Wu et al. 2002) was that the z-axis was the line connecting the midpoint between the malleoli and the epicondylar gaps (rather than the epicondyles). The origin of this segment was the mid-distance between the left and right epicondylar gaps. The orientation of the shank's coordinate system was determined by the knee and ankle markers, the shank marker cluster was then used to track the movements of this segment during dynamic trials.

Modelling the foot. Although the foot is made of numerous bones and could be modelled as multiple segments, it was modelled as a single segment given that studying the movement between the foot sections was not of interest within this thesis. The origin of the foot coordinate system was the midpoint between the malleoli, with the anterior-posterior axis running from the coordinate system origin to the midpoint between the first and fifth metatarsal heads. The coordinate system for this segment followed the ISB recommendations (Wu et al. 2002), which causes the foot to be in plantarflexion. However, this does not represent the genuine position of the foot, which normally lays flat on the floor, unless there are issues at the ankle

joint or spasticity of the gastrocnemius. The video assessment of the patients participating in the current project revealed that all of them stood with their feet flat on the floor. Therefore, to rectify the foot plantarflexion, a virtual foot (used only for kinematic calculations) was created to be parallel to the floor, flowing Biggs' (2016) previous work, where the malleoli, first and fifth metatarsal head markers were projected onto the floor and used as a reference to define the origin of the coordinate system and the anterior-posterior axis orientation, respectively.

Virtual laboratory. A virtual laboratory coordinate system was created to follow the orientation of the participant's pelvis' coordinate system, meaning that the Y-axis was always pointing forward, in the walking direction. This allowed to easily calculate joint biomechanics when the participant walked in both directions within the laboratory.

V3D Pipeline and calculation of biomechanical variables

A static calibration was loaded, and the biomechanical model described above was applied to the participant, scaled to the specific height and weight to allow for kinetic calculations. Additionally, with this step, the relative position of the markers on the clusters and the segment coordinate system was determined, assumed to be fixed and was later applied to each walking trial. The walking trials files were associated with the model built in the previous stages.

Both analogue data from the force plates and the markers' trajectories were filtered with a Butterworth bidirectional, fourth-order, low-pass filter at a cut-off frequency of 6 Hz, which was the procedure previously followed at Cardiff University (Watling 2013; Biggs 2016) and habitually utilised in gait analysis (Kaufman et al. 2016a; Langley et al. 2019). Walking trials were visually inspected for quality check purposes and errors were investigated and rectified. A 20 N force plate threshold (vertical GRF) was utilised to determine heel strike and toe-off, a technique previously followed at Cardiff University (Watling 2013; Biggs 2016), and common in previous studies (Yoshida et al. 2012). When the foot hit the floor outside the force plates at the end of the walkway, an automatic gait event detection algorithm was utilised to determine the last heel strike [based on the foot segment kinematics patterns observed during previous force plate contacts earlier in the trial (Stanhope et al. 1990)], and the data within the first force-plate-determined heel strike and the

automatic detection of the second heel strike were only used for kinematic calculations. Unclean force plate hits were discarded from further analysis.

Pelvis, hip, knee, and ankle joint angles were calculated in all three anatomical planes for each frame recorded, and then time normalised to 0-100% of the gait cycle (i.e., 101 time points). Pelvis angles were calculated relative to the virtual laboratory coordinate system. The Cardan-Euler sequence for the pelvis segment was Z-Y-X (rotation, obliquity, tilt), following the technical note of (Baker 2001), to avoid imprecisions in calculating pelvis obliquity.

For all the other segments, joint angles were calculated with the proximal segment as a reference and therefore the distant segment moved relative to the proximal, where the first rotation was about the mediolateral axis of the proximal segment, the second rotation was about a floating axis perpendicular to the mediolateral axis of the proximal segment and the proximal-distal axis of the distal segment, the last rotation was about the proximal-distal axis of the distal segment (Wu and Cavanagh 1995; Wu et al. 2002); therefore, the Cardan-Euler sequence was X-Y-Z (flexion-extension, adduction-abduction, rotation). The foot progression angle was calculated as the angle between the longitudinal axis of the foot in the transverse plane, and the anterior-posterior axis of the virtual laboratory's coordinate system. Flexion, pelvis anteversion, ankle dorsiflexion, adduction, and internal rotation angles were positive (+).

The three components of the GRF were calculated relative to the force plate's coordinate system, expressed in Newtons in the raw data, then scaled by dividing the signals by the subject's mass times 9.81 m/s^2 (i.e., expressed in times Body Weight) to allow for data comparison across subjects, and normalised to 0-100% of stance.

The GRF produces external moments around the hip, knee and ankle joints, which are counteracted by the rotational forces generated by the soft tissues (i.e., internal moments). External moments were calculated with inverse dynamics for all three anatomical planes for the stance phase of the gait cycle and normalised to % Bodyweight * Height. More precisely, joint moments were dependent on the GRF magnitude, the moment arm between the GRF vector and the centre of rotation of the joint, and the segment inertial properties. Segment mass and its inertial properties were automatically calculated by V3D and were determined using

Dempster (1955) and Hanavan (1964) work as references, respectively. There is no obvious convention on the best coordinate system to utilise to calculate joint moments, with some authors computing them as proximal moments, distal moments, joint coordinate system, plane of progression (Robertson et al. 2013). Joint moments in this thesis were calculated relative to the distal segment of the joint in question (i.e., distal joint moments). Joint powers for the hip, knee and ankle were calculated with inverse dynamics as a product of the joint moment by the distal segment angular velocity, normalised by body weight (W/kg) and % of the gait cycle for each of the anatomical planes. Gait speed was calculated as the stride distance divided by the stride time of the leg of interest and averaged across trials.

3.7 DATA REDUCTION AND CLASSIFICATION OF GAIT FUNCTION

Motion analysis produces a considerably large amount of data that can be difficult to interpret. Furthermore, considering that for each variable 101 data points are available, the complexities of analysing such data with conventional statistical methods comparing differences between or within groups become apparent. One of the strategies extensively adopted in biomechanics is extracting discrete data at the maximum or minimum peaks of the waveform, or in the case of knee adduction moments, the area under the curve (i.e., impulse) also called parametrisation. While this method allows to easily interpret the data, the limitation is that potentially important information is disregarded, and it does not allow to account for patterns in the overall waveform. Furthermore, it is not always possible to identify obvious peaks (as in the knee adduction moment in the presence of severe knee OA), or peak values may occur at different points of the gait cycle between groups of subjects, making it difficult to interpret the findings. Principal component analysis (PCA) allows to consider the whole waveform of a set of data and to reduce the amount of information to key features that can be interpreted. The use of PCA in gait analysis has become increasingly common, mainly driven by the work of Deluzio and colleagues on individuals with knee OA (Deluzio et al. 1997; Deluzio et al. 1999; Deluzio and Astephen 2007).

PCA has been utilised within our research group in combination with the Cardiff classifier, based on the Dempster-Shafer theory of evidence and described in more detail in section 3.7.3. The Cardiff classifier has been employed at Cardiff University

to investigate biomechanical differences between individuals with medial knee OA and NPs (Bowd 2022), severe knee OA and NPs, and to evaluate the change in gait biomechanics post-TKR (Watling 2013; Metcalfe 2014; Biggs 2016; Biggs et al. 2019a; Biggs et al. 2019b; Whatling et al. 2022), post-HTO (Whatling et al. 2020), post-hip replacement (Biggs et al. 2021). The Cardiff classifier utilised within this thesis is based on the research conducted by Dr Biggs (Biggs 2016), who expanded upon the work of Jones (2004) to develop the classifier in its current form.

3.7.1 Principal component analysis

PCA is a statistical procedure utilised for dimensionality reduction and data analysis. It is an orthogonal transformation of the original data into a new set of variables, called principal components, that are independent from one another and that describe *features of variance* of a waveform, based on the variation within the data. Followingly, the explanation of PCA is based on the example of 3D gait data analysis. The gait waveform for a variable expressed from 0 to 100% of the gait cycle has 101 data points. PCA transforms each of the data points into a new set of uncorrelated variables across all participants, resulting in 101 principal components (PCs). Each PC describes different patterns of the original waveform, for instance, phase shifts, differences in waveform magnitude, etc. Some of the advantages of using PCA with gait data are the fact that only a small number of PCs is needed to describe the original data with sufficient detail, each of the PCs has associated scores that can be utilised for statistical analysis (i.e., differences in features pre- to post-surgery), the principal components are a new set of independent variables (Robertson et al. 2013).

PCA was carried out with eigen decomposition of the correlation matrix, with a code initially developed in MATLAB (MathWorks Inc., Natick, Massachusetts, US) by Jones (2004), and that was later refined by Biggs (2016). Followingly, the description of the procedure to perform PCA within the code. The original time series data is organised within a *n x p* matrix X where each row represents a participant (n), and each column (p, a total of 101 columns) is the value of the variable of interest (e.g., knee flexion-extension during the gait cycle) at each time point. PCA is performed separately on each gait variable.

1) Data standardisation: the original data mean is removed from each data point, and the result is divided by the original data standard deviation (allowing to obtain a z-score), which scales each of the original values, so they have zero mean and a standard deviation of 1 (Chau 2001). Portions of the original waveform with large variation would have had a greater effect in determining the PC direction, had the values not been scaled
2) Correlation matrix: to understand the variation in the data between subjects and over time, a covariance matrix of the columns of X is calculated. Given that the data had been standardised, the covariance matrix $C = X^TX/(n-1)$ corresponds to the correlation matrix (Chau 2001). The elements off the diagonal of the new correlation matrix C represent the correlation coefficient between all possible pairs of the standardised values
3) Eigen decomposition: the eigen decomposition of matrix $C = PDP^T$ returns a 101 x 101 matrix of eigenvectors (*P;* each of the columns is an eigenvector, a unit vector) (i.e., the PCs) and likewise, a 101 x 101 matrix of eigenvalues (diagonal matrix *D*) (i.e., the percentage of variance explained by each PC). Matrix P transforms the original data points into a new coordinate system, where the axes of the coordinate system are the PCs (Chau 2001). PC1 explains the largest amount of variance (i.e., the axis along which the data is more “spread out”, or the axis with the minimal perpendicular distance from each point), PC2 represents the second highest amount of variance in the data, but orthogonally to PC1, etc.
4) Retention of PCs: PCA is performed to reduce data dimensionality, and therefore, only a few PCs are retained to describe the original data. There are various approaches to aid in determining the number of PCs that should be retained for further analysis. For instance, Kaiser’s rule (PCs with eigenvalues > 1 are retained) (Kaiser 1960) tends to retain a large number of PCs that may not be informative. Another approach is retaining the number of PCs whose amount of variance explained is larger than 95% of the variance of the original data (Robertson et al. 2013). The approach adopted in the current thesis follows the work of Biggs (2016), where the first three PCs for each gait variable were retained for further analysis within the Cardiff classifier, as it

was demonstrated that a sufficient and large amount of variance of the original data was explained by implementing this approach

5) PC scores: these represent the standardised data projected onto the retained PCs axes. The matrix X containing the original data had n x p dimensions (n = number of participants; p = number of features). The retained PCs (or eigenvectors) (k) were within a PC matrix of k x p dimensions, and the PC scores were calculated as S = X * PC (* = matrix multiplication), where the new matrix S had n x k dimensions, each row n (i.e., participant) contained the PC scores across the PCs retained. This means that each participant had a PC score for each of the gait features (i.e., PCs) retained. The PC score of a gait feature represented how far the shape of the waveform of a subject was from the mean of the entire dataset for that particular feature. The advantage of obtaining PC scores is that they can be utilised for statistical testing (Robertson et al. 2013)
6) Calculation of factor loadings: factor loadings (or loading vectors) represent the correlation between the PCs and the original data, also representing the variance depicted by the PC in question at each time point of the gait cycle or stance. Multiplying the eigenvectors by the square root of eigenvalues returned the factor loadings.
7) PC reconstruction: this was performed to aid the interpretation of the biomechanical meaning of the retained PCs. The obtained data (i.e., PC scores), now in a new axis system, had to be reverted into the original axis system, this was done by multiplying the PC score by the associated eigenvector. The resultant vector was then multiplied by the standard deviation of the original data (i.e., the standard deviation of the data at each point of the gait or stance cycle), to which the mean of the original data was added (i.e., the mean of the data at each point of the gait or stance cycle). Of note, if the PC scores of all 101 PCs were utilised to carry out PC reconstruction, the original waveform for that variable would be obtained

3.7.2 Interpretation of principal components

Interpretation of the biomechanical meaning of each PC retained was achieved following two methods, a visual interpretation of the r*epresentative*

extremes and of the *single-component reconstruction*. The interpretation of a PC allows an understanding of the features that were depicted from the raw waveform, and it aids the explanation of the differences in PC scores between patients with severe knee OA and NP. This task was achieved by analysing the graphs produced by the custom-written MATLAB code (MathWorks Inc., Natick, Massachusetts, US) performing the PCA (Figure 13), previously developed in our research group. The combination of the two abovementioned methods was recommended by Brandon et al. (2013), who found that utilising solely representative extremes to interpret the biomechanical meaning of a PC may lead to a biased reading as the raw data may represent information including the variance from another PC. The interpretation of representative extremes involves comparing the raw waveforms of the participants with a low (5^{th} percentile) and high (95^{th} percentile) PC score in question (relative to the mean waveform for that feature) (Figure 13a) and retaining for interpretation only the section where the squared loading vector waveform had the greatest magnitude (Figure 13b) (McKean et al. 2007).

The squared loading vector is a waveform representing the proportion of variance depicted by the PC in question at each time point of the gait cycle or stance. A “magnitude” feature was identified when the squared loading vector waveform had high values throughout the whole gait cycle and a visible distance could be seen on the raw waveforms graph’s ordinate (raw waveforms in Figure 13a). A “difference” feature was found when the squared loading vector had a high magnitude(s) in correspondence with the raw waveforms’ peak(s). Finally, a “phase shift” feature was identified when, in correspondence with a peak on the raw waveforms, the squared loading vector was zero, preceded and followed by a high magnitude. Single-component reconstruction was obtained by multiplying the PC scores of the 5^{th} (low) and 95^{th} percentile (high) subjects in Figure 13a by their eigenvector which can be seen in Figure 13c. These were the same two participants that were chosen as representative extremes (Figure 13a), as described above. The interpretation was based on a visual assessment of the low and high PC reconstruction.

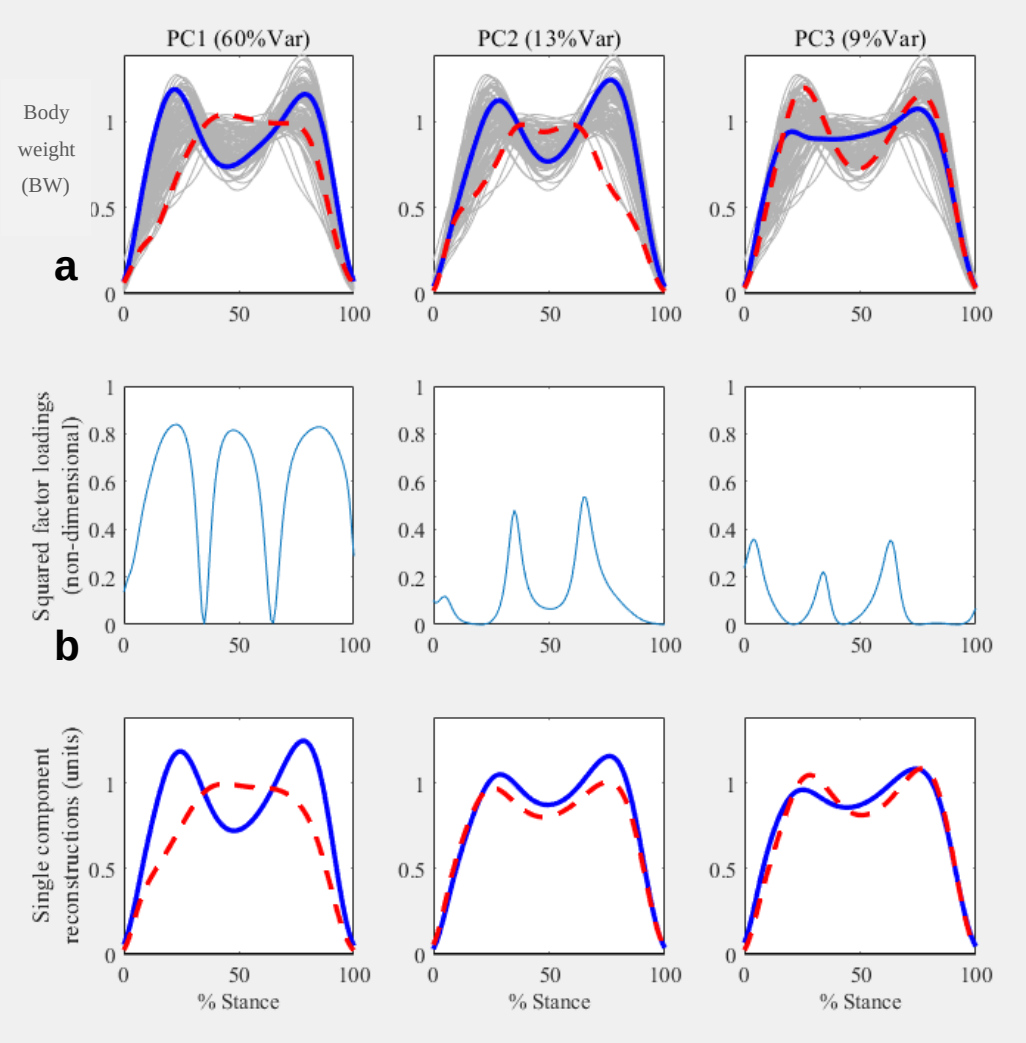


Figure 13 **a)** Example of the vertical ground reaction force raw waveforms for all the patients and non-pathological volunteers (grey waveforms) where the participants with low principal component score (5th percentile; red, dashed line) and high principal component score (95th percentile; blue, solid line) were highlighted. PC1 represents 60% of the variance of the original data. **b)** The waveform displays the squared factor loadings, which is the ratio of variance represented by the PC in question at each stage of the gait cycle (1 on the ordinates = 100% of the variance). **c)** Single component reconstruction of the participant with a low principal component score (red, dashed line) and high principal component score (blue, solid line) seen in **a**.

3.7.3 The Cardiff classifier

The Cardiff classifier is based on the Dempster-Shafer theory of evidence, also called the belief theory, which allows to mathematically combine information that has a degree of uncertainty (i.e., conflicting and agreeing evidence) for decision-making (Dempster 1968; Shafer 1976). It is called belief theory because it is centred on the concept of belief functions, the belief or uncertainty about the truth of a proposition. In the case of the cohorts examined in this thesis, the belief is centred around having osteoarthritic or non-pathological biomechanics. The main idea of the theory is to combine evidence from multiple sources to arrive at a decision, and to combine and update the belief when new evidence is added.

In the case of utilising 3D gait input data, the classifier assigns three Belief values to each individual entered in the classification, a Belief of NP (BNP, the level of belief that the subject has an non-pathological gait), a Belief of OA (BOA, the level of belief that the subject has an osteoarthritic gait), and an Uncertainty level (the level of belief that the subject has neither a non-pathological or an osteoarthritic

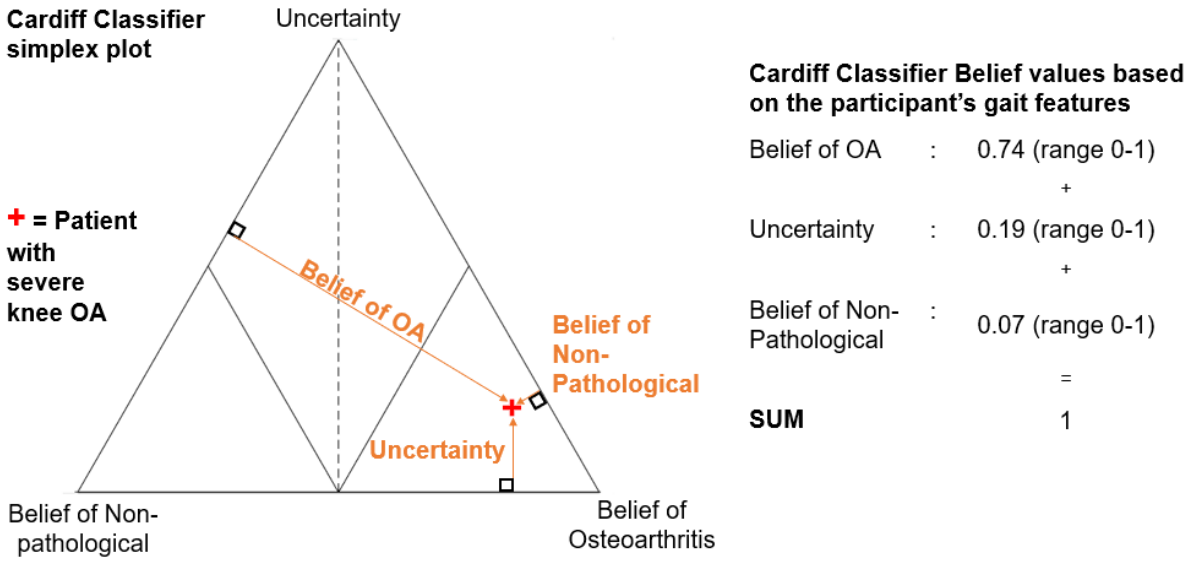


Figure 14 Cardiff classifier simplex plot, example of classification of a patient with severe knee OA (the other patients and NPs were removed, for ease of understanding). The black, dashed line represents the boundary between BOA and BNP. The black solid line represents BOA or BNP = 0.5. The orange arrows represent the shortest distances between the plotted point and the opposite side of the triangle and are a visual representation of how the position of the patient is determined on the simplex plot, based on the patient's Belief values.

gait), based on the evidence from their gait biomechanics (i.e., PC scores for each PC chosen in the analysis). Each of the Belief values ranges from 0 (e.g., null BOA) to 1 (e.g., maximum BOA), and their sum is always 1. Once the three Belief values are calculated, the position of the subject in a simplex plot (an equilateral triangle where each side is of unit length) is determined and can aid in understanding their gait function (Figure 14).

The Cardiff classifier algorithm was developed in MATLAB (MathWorks Inc., Natick, Massachusetts, US) by Jones (2004), and was later refined by Whatling (2009), and Biggs (2016). Following PCA described in the previous section, three PCs were retained for each variable of interest. For instance, if the researcher was interested in assessing ten gait variables, there would be thirty PC scores for each subject. The PC scores were utilised within the classifier for further analysis. In the Excel spreadsheet (Microsoft, Redmond, WA, US) containing the first three PC scores for variable and each participant, the groups were identified in classes, 0 for NPs, 1 for patients pre-TKR and 2 for subjects post-TKR. The classifier technique can be summarised in three main steps:

1) Each input feature (i.e., PC score) was transformed into a confidence factor using a sigmoid function (Figure 15.1), called Confidence Function (Eq. 4), which can range from 0 to 1. This is the level of confidence that the PC score for a certain gait variable supports the hypothesis that the individual has, for instance, an OA gait (e.g., *cf(x)* = 1 if there is certainty in OA)

$$cf(x) = \frac{1}{1 + e^{-k(x-\theta)}} \quad \text{Eq. 4}$$

Where k defines the steepness of the sigmoid function (Eq. 5), *x* is the value of the feature (i.e., the PC score of a PC), θ is the value at which *cf (x)* = 0.5 (i.e., the value of *x* where the evidence supporting their gait being non-pathological is equal to that supporting their gait being osteoarthritic) (Jones 2004), which is calculated with Eq. 7.

$$k = \frac{I \times \rho(x, y)}{\sigma_x} \quad \text{Eq. 5}$$

Where *l* is a constant (Eq. 6), ρ is the Pearson correlation coefficient between the variable *x* and the classes *y* (i.e., NP = 0 and OA = 1), σ_x is the standard deviation of *x* for both groups.

$$l = \frac{n}{\sum_{i=1}^{m} |\rho_i|} \quad \text{Eq. 6}$$

Where *n* is the number of participants, *m* is the number of input features used within the classifier, ρ_i the Pearson correlation coefficient between all the input features and the class labels (see Eq. 5).

$$\theta = \mu_{NP} + \mu_{OA}\left(\frac{\sigma_{NP}}{\sigma_{NP} + \sigma_{OA}}\right) \quad \text{Eq. 7}$$

Where μ represents the mean of the group and σ the standard deviation of the feature (*v*) for the group. With this equation (Eq. 7), the heterogeneity of variance between groups is accounted for (Biggs 2016).

2) The Body of Evidence (BOE) was created by converting each confidence factor in m(OA), m(NP) and m(θ) functions (Figure 15.2); Eq. 8, Eq. 9 and Eq. 10 show how the m(OA), m(NP) and m(θ) were calculated, respectively (Safranek et al. 1990). These were dependent on the control variables A and B, where A denotes the dependence of the m(OA) on the confidence factor, B represents the maximal support which can be assigned to either m(OA) or m(NP).

$$m(OA) = \frac{B}{1-A}cf(x) - \frac{AB}{1-A} \quad \text{Eq. 8}$$

$$m(NP) = \frac{-B}{1-A}cf(x) + B \quad \text{Eq. 9}$$

$$m(\theta) = 1 - m(OA) - m(NP) \quad \text{Eq. 10}$$

A and B were defined as:

$$A = \frac{\Theta_U - \Theta_L}{1 + \Theta_U - 2\Theta_L} \quad \text{Eq. 11}$$

$$B = 1 - \Theta_L \quad \text{Eq. 12}$$

The upper and lower boundaries of uncertainty θ_U and θ_L, were defined in the work of (Jones 2004) and chosen after investigating the best combination of values to obtain an accurate classification of gait function between severe knee OA and NPs. The values were $\theta_U = 1$ and $\theta_L = 0.8$, and were the same for each input feature, as in Jones (2004), meaning that every supporting evidence could contribute a maximum of 0.2 for generating the Belief value

3) Dempster's rule of combination was utilised to combine the BOE for each feature; the combined BOE (CBOE) for all the features is the probability that a subject belongs to the OA, NP class or Uncertainty. In the following expressions, two BOE, m_i and m_j were combined to calculate the BOA (Eq. 13), the BNP (Eq. 14) and Uncertainty (Eq. 15)

$$m_i \oplus m_j(OA) = \frac{m_i(OA)m_j(OA) + m_j(OA)m_i(\theta) + m_i(OA)m_j(\theta)}{1 - C} \quad \text{Eq. 13}$$

$$m_i \oplus m_j(NP) = \frac{m_i(NP)m_j(NP) + m_j(NP)m_i(\theta) + m_i(NP)m_j(\theta)}{1 - C} \quad \text{Eq. 14}$$

$$m_i \oplus m_j(\theta) = \frac{m_i(\theta)m_j(\theta)}{1 - C} \quad \text{Eq. 15}$$

$$C = m_i(OA)m_j(NP) + m_j(OA)m_i(NP) \quad \text{Eq. 16}$$

Where C is the normalisation factor for conflicting probability. The order in which the BOE are combined do not affect the final Belief value, since Dempster's rule is associative and commutative (Gerig et al. 2000).

Following the CBOE, if a subject's BOA > BNP, the participant was classed as having osteoarthritic function (Figure 15.3), if BNP > BOA, the participant was classed as having non-pathological function (Safranek et al. 1990).

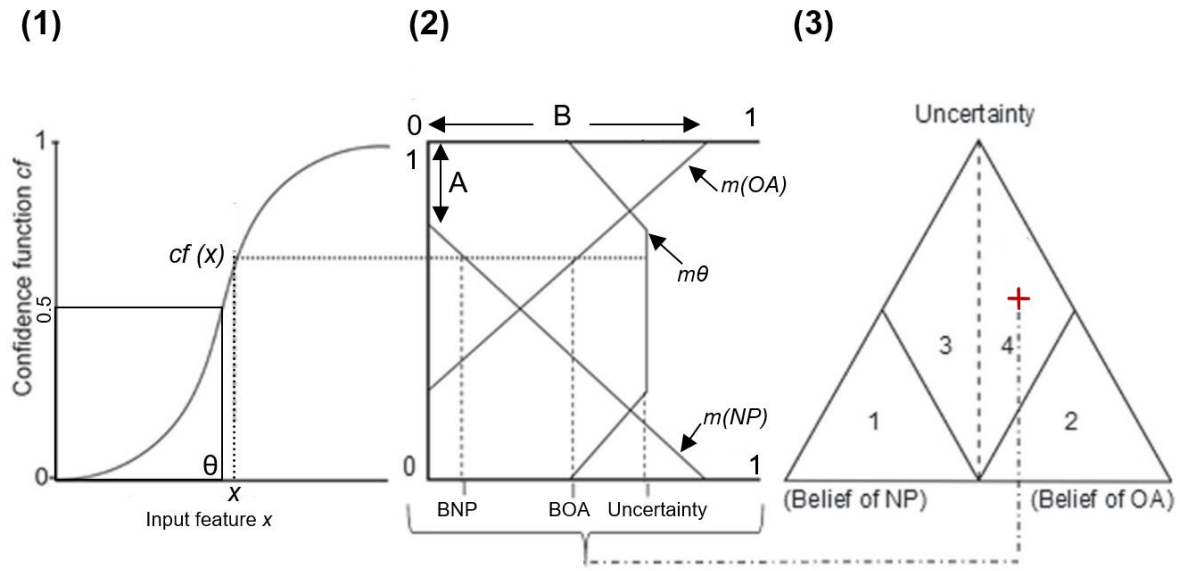


Figure 15 The three steps of the Cardiff classifier, example for an input variable. 1) Conversion of input variable into a confidence factor; 2) conversion of the confidence factor into a Body of Evidence; 3) conversion of the Body of Evidence in its simplex plot coordinates, with an example participant with severe knee OA (red cross). The simplex plot has four regions, region 1 is the dominant NP classification region where BNP ≥ 0.5, region 2 is the dominant OA classification region where BOA ≥ 0.5, region 3 is the non-dominant NP region where 0.5 > BNP > BOA, region 4 is the non-dominant OA region where 0.5 > BOA > BNP. The black, dashed line represents the boundary between BOA and BNP classification.

Model validation and method for ranking of the input features

Within the Cardiff classifier MATLAB algorithm, model validation was performed to assess the error (or misclassification) rate, and, therefore, the robustness of the classification method. The process of calculating the classification error would be biased if the estimation of the error was based on the same cases utilised to determine the classification parameters (Henery 1994). A solution to this is to utilise two independent data sets, the first to the train the classifier, the second to test it (Henery 1994). This process will be described followingly, and for ease of understanding, an example including n = 20 subjects (10 NPs and 10 patients pre-TKR), and 30 gait features (i.e., PCs) for each subject will be used.

To perform model validation (and, therefore, to evaluate the *classifier classification accuracy*), a leave-one-out cross-validation approach was utilised (i.e., out-of-sample classification), where the classifier was tested on different subjects

(i.e., out-of-sample cases) to the ones utilised to define the control parameters (i.e., in-sample cases). More precisely, in the training phase, the classifier was trained utilising (*n* - 1) subjects (in the example, 19 subjects) and all 30 input features. The control parameters (k, θ, A, B) were defined for each of the 30 input features, based on the in-sample 19 subjects. In the testing phase, the control parameters were utilised to calculate a BOE for the one out-of-sample subject for each input feature, in other words, the classifier was tested on the subject that had been left out. For the out-of-sample subject in question, BOEs of each input feature were then combined with Dempster-Shafer theory of evidence, to calculate a BOA, BNP and Uncertainty level, which allowed to position the subject in the Cardiff classifier simplex plot (Figure 16). This process was repeated *n* times (i.e., 20 times, or loops, in the example), namely, until all 20 subjects had been left out.

At the end of this process, a simplex plot was produced showing the out-of-sample classification, therefore, the out-of-sample *classifier classification accuracy* could be determined. As mentioned earlier, the participants classes were defined a priori, where NPs were assigned a class of 0 and patients with severe knee OA a class of 1 (i.e., supervised classifier). A classification error (or misclassification) was determined when, for instance, a subject with severe knee OA had a BNP > BOA (i.e., the classification did not match the class label, and the subject was located on the incorrect side of the classifier simplex plot). The *classifier classification accuracy* was calculated as 100 minus the classification error (or as the proportion of all subjects who were correctly classified using the combination of all input variables out of the total NPs and patients).

During each loop of the leave-one-out cross-validation, the classifier ranked the input features which most accurately discriminated between subjects with severe knee OA and NPs (Figure 16). The *feature accuracy* of each input feature was calculated during the training phase from the in-sample subjects (in the example, 19 subjects) at each loop. For an input feature, the BOE for each in-sample subject was defined [i.e., m(OA), m(NP) and m(θ)]. The *feature accuracy* of an input feature corresponded to the proportion of the subjects whose class corresponded to the correct BOE, out of the 19 in-sample cases within that loop [e.g., patients with severe knee OA were expected to have a m(OA) > m(NP)],

At each loop, input features were ranked based on their *feature accuracy*, from the one with the highest accuracy, to the one with the lowest. It is worth noting that the ranking of features may have differed across loops of the leave-one-out cross-validation method. In this thesis, the eighteen features which had the highest *feature accuracy* in discriminating between patients with severe knee OA and NPs were utilised in the testing phase of the classifier (and, therefore, for testing the one subject left out during the leave-one-out cross-validation), unless otherwise specified. Previous research found that retaining 15 to 18 features allowed for an accurate classification (Biggs 2016; Bowd 2022).

It was also demonstrated that the classifier had the ability to discriminate between patients and NPs with a 90% accuracy even with as little as five subjects within each group (Biggs 2016).

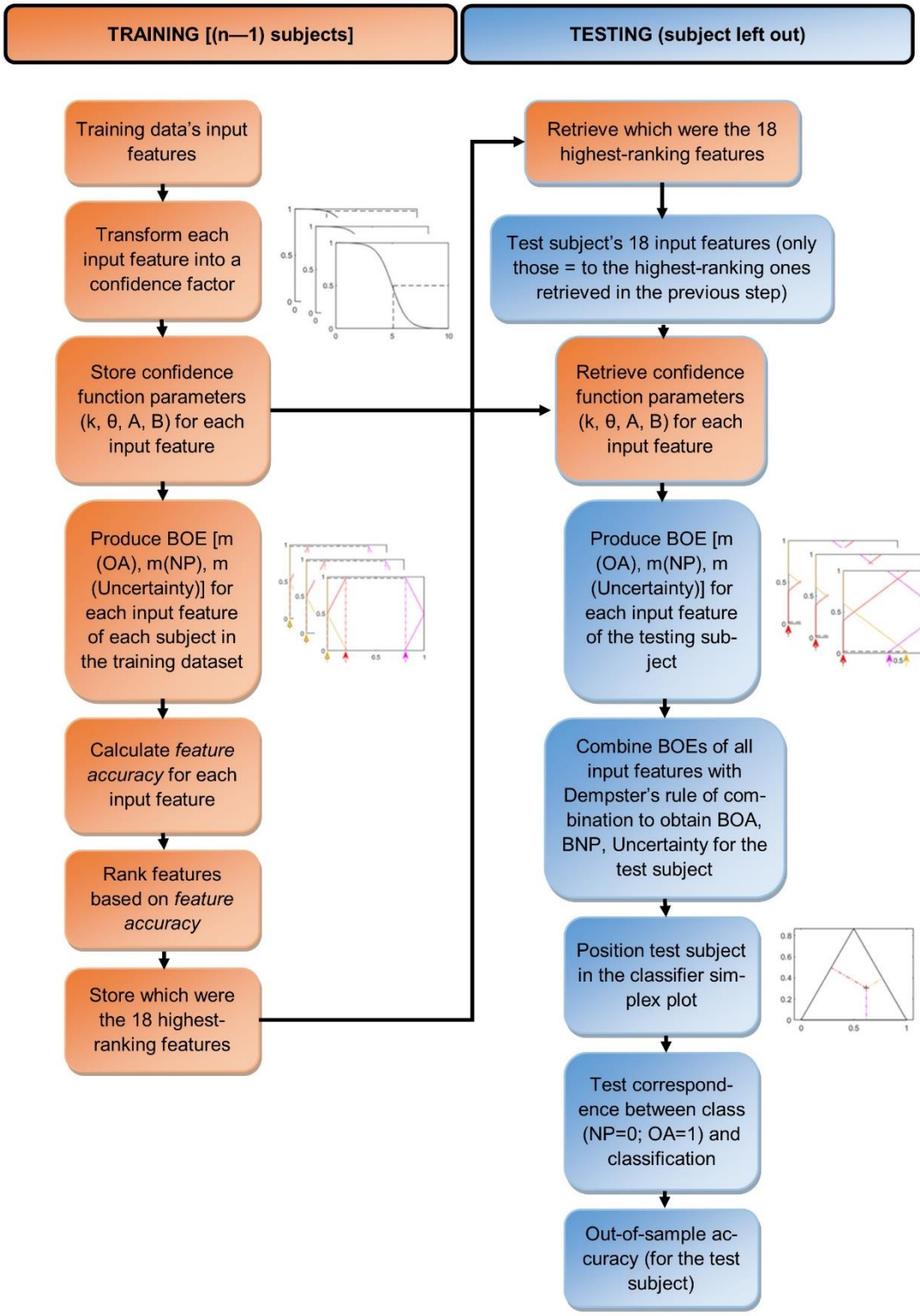


Figure 16 Graphical representation of the training and testing phases of the classifier for a loop of the leave-one-out cross-validation. n: number; BOE: Body of Evidence; OA: osteoarthritis; NP: non-pathological; BOA: Belief of OA.

Classification of objective physical function based on the full dataset

Once model validation had been performed, a final classification of the subjects was carried out, based on the full dataset, as explained below. Therefore, the classifier was trained utilising *n* subjects (in the example, 20 subjects) and all 30 input features. This means that the control parameters for the Confidence Function (k, θ, A, B) were calculated for each of the 30 input features, based on the 20 subjects; the eighteen highest-ranking variables discriminating patients pre-TKR from NPs were determined based on the whole sample (according to the *feature accuracy*, as described above).

Followingly, the classifier was tested on the 20 subjects (*n* subjects), meaning that the control parameters defined in the training phase were utilised to calculate, for each of the *n* subjects, the BOE for each of the retained eighteen highest-ranking input features defined from the training dataset. For each subject, the BOEs of each of the input features were then combined with Dempster-Shafer theory of evidence, to return a BOA, BNP and Uncertainty level, which were the classifier outputs utilised to perform all statistical testing. The control parameters for the Confidence Function, alongside the eighteen highest-ranking variables discriminating patients pre-TKR from NPs determined in the training of the classifier, were used to classify the participants post-TKR.

Chapter 4: Comparative assessment of gait biomechanics via the Cardiff classifier, GDI and GDI-kinetics

4.1 INTRODUCTION

3D gait analysis generates large volumes of data describing joint biomechanics in all three planes of movement that are very complex to interpret. Therefore, several research groups, including ours, developed summative measures of gait function, the GDI – GDI-kinetic and the Cardiff classifier, where the index, or the BOA value, respectively, are representative of the objective function or the gait quality of people pre- and post-TKR.

For any outcome measure utilised in longitudinal studies, such as those looking at functional recovery pre to post-TKR, it is important to be able to depict changes and therefore, to be responsive. The internal responsiveness of a measure is its ability to quantify change, and while there is no gold standard to measure responsiveness (Husted et al. 2000) it is usually measured via the effect size (ES) of a change score pre- to post-intervention (Portney and Watkins 2014). Only two studies reported a statistically significant difference in the patients' BOA values pre- to post-surgery using the Cardiff classifier (Biggs et al. 2021; Biggs et al. 2019). Of these studies, only one (Biggs et al. 2019a) reported the responsiveness of the BOA pre- to 9+ months post-TKR (large ES), in a relatively small group of patients (n = 22). Given the limited research in this area, it is highly relevant to explore the Cardiff classifier's internal responsiveness to change in various and larger cohorts of patients undergoing TKR, and to compare it to the responsiveness of other similar measures of objective function.

The Cardiff classifier's abilities to measure function have previously been compared to patients' self-reported function (measured via PROMs). An earlier study on twenty-two patients found a moderate to strong correlation between the BOA and two subjective measures of gait function (Knee Outcome Survey – KOS, and OKS) pre-TKR and at one-year follow-up, a strong correlation in the BOA and KOS change scores, and between the BOA and OKS values change (Biggs et al. 2019a). These findings showed that a poorer patient's perceived function was associated with more compromised gait biomechanics and that an improvement in gait function was related to their perception of functional improvement (Biggs et al. 2019a).

Although PROMs may be considered the gold standard in measuring patients' perceived function, they are subjective in nature and, as previously shown in Chapter 2 (Literature Review) most reports found a discrepancy between objective and subjective measures of function (Stratford et al. 2010; Mizner et al. 2011; Stevens-Lapsley et al. 2011b). There are no studies, to date, comparing the Cardiff classifier's outputs to other methods to quantify objective gait function in people with knee OA and post-TKR. This information would help in understanding how the classifier performs compared to measures that evaluate a similar construct. Within the research community, there is no gold standard "tool" summarising gait variables and returning a value indicating the gait "quality" of a subject compared to that of a non-pathological group. GDI and GDI-kinetic have been increasingly utilised to objectively quantify the gait quality and its change pre to post-TKR in several previous reports as discussed in the Chapter 2 (Literature Review) (Naili et al. 2017a; Naili et al. 2017b; Kobsar et al. 2019; Naili et al. 2019a). Therefore, it was deemed appropriate to compare the Cardiff Classifier outcomes to that of the GDI and GDI-kinetic, because all three measures can objectively quantify gait function against normative data. It is unknown if the objective gait function pre-surgery and the outcome after TKR may be different when measured via the Cardiff Classifier, GDI and GDI-kinetic. Choosing the best tool to evaluate the objective gait function is fundamental for the advancement of the research and to understand what level of improvement patients have post-TKR.

The work contained in this chapter develops from a collaboration with Dr Josefine Eriksson Naili based at Karolinska University Hospital and Karolinska Institutet (Stockholm, Sweden), where the GDI and GDI-kinetic have been used to evaluate the gait "quality" in patients pre- and one-year post-TKR. Dr Naili and Professor Anders Holsgaard-Larsen have previously collaborated with Professor Cathy Holt, the main supervisor of this thesis, and Dr Paul Biggs, whose thesis work formed the basis for the investigations within the current thesis. Their work utilised the Cardiff Classifier to assess gait function in patients undergoing total hip replacement (Biggs et al. 2021). The author of the current thesis was interested in continuing the previous collaboration that Dr Biggs established between our research groups with the intention to expand the research on the measuring properties of the Cardiff classifier. Moreover, Dr Naili and Professor Holsgaard-Larsen had an interest

in exploring the potential of the Cardiff classifier, compared to the GDI. Further collaborations were discussed between the two research groups, leading to the development of aims 3 and 4 (described below). Dr Naili contributed their NPs and patients' pre- and post-TKR biomechanical data to the author who performed the statistical analysis, data interpretation and writing of the findings. The preliminary findings for Aim 3 were presented in the form of a poster at the Osteoarthritis Research World Congress in 2022 (De Vecchis et al. 2022). The remaining aims for the current chapter were developed by the author and the outputs were reviewed by Dr Naili.

It must be noted that the participants from the Karolinska cohort were involved in two previous studies (Naili et al. 2017c; Naili et al. 2017b). One of the studies (Naili et al. 2017c) looked at discrete knee kinetics and kinematics in the sagittal and frontal planes and compared the differences in these variables between patients with a good or poor outcome as defined by a patient-reported questionnaire (KOOS). The second study (Naili et al. 2017b) investigated the GDI and GDI-kinetic change pre to one-year post-TKR on the operated and contralateral lower limb; it was found that post-TKR there was a moderate improvement in GDI and GDI-kinetic scores (ES = 0.4 and 0.5, respectively). The reference groups utilised in the above-mentioned studies (Naili et al. 2017c; Naili et al. 2017b) were different from the ones adopted in the present investigation. The patients from the Cardiff and Karolinska cohorts had similar baseline characteristics and it was deemed appropriate to investigate the responsiveness to change of the Cardiff Classifier in both cohorts of patients to enhance the strength of the findings. Regarding the Karolinska cohort, the novelty of the current work lies in comparing the GDI and GDI-kinetic internal responsiveness and outputs to the ones obtained via the Cardiff classifier, in investigating gait changes in this cohort through the classifier, and in comparing the findings from the classifier to the Cardiff cohort.

Aim 1: Investigate the internal responsiveness of the classifier and compare it to the responsiveness of the GDI and GDI-kinetic in the Cardiff and Karolinska patient groups.

Hypothesis 1.1: the classifier’s BOA has a large responsiveness to change in two groups of patients pre- to post-TKR.

Hypothesis 1.2: the classifier's BOA is a more responsive measure of gait change pre to post-TKR compared to the GDI and GDI-kinetic in both patient groups, given previous results (Naili et al. 2017b; Biggs et al. 2019a).

Aim 2: Explore the relationship between the BOA and the GDI and the BOA and GDI-kinetic in two groups of patients pre- and post-TKR.
Considering that the gait variables used in the classifier were almost comparable to those utilised to calculate the GDI and GDI-kinetic for the two cohorts, the following hypotheses were formulated:
Hypothesis 2.1: there is a significant, negative, moderate correlation between the BOA and GDI pre-TKR in both patient groups.
Hypothesis 2.2: there is a significant, negative, moderate relationship between the BOA and GDI-kinetic pre-surgery in both patient groups.
Hypothesis 2.3: there is a significant, negative, weak correlation between the BOA and GDI post-TKR in both patient groups (due to the findings from Naili et al. (2017b) indicating a return to normal kinematics as measured by the GDI).
Hypothesis 2.4: there is a significant, negative, moderate relationship between the BOA and GDI-kinetic post-surgery in both patient groups.

Aim 3: Explore the association between BOA change-GDI change, and BOA change-GDI-kinetic change pre- to post-surgery in two groups of patients pre- to post-TKR.
Hypothesis 3.1: there is a significant, moderate correlation between the BOA change and GDI change pre- to one-year post-surgery in both patient groups.
Hypothesis 3.2: there is a significant, moderate relationship between the BOA change and GDI-kinetic change pre to one-year post-TKR in both patient groups.

To aid the interpretation of the findings from the above aims, aims 4 and 5 were developed. Moreover, it is not clear whether the highest-ranking features discriminating patients from NPs found in previous reports (Biggs 2016; Biggs et al. 2019b) were unique to the Cardiff cohort, or if they could be considered recurring and key features aiding in discriminating patients pre-TKR from NPs. Only two other studies utilised the Cardiff classifier in patient groups other than the Cardiff cohort

(Worsley 2011; Metcalfe 2014), Worsley (2011) analysed a mixed group of patients undergoing TKR and UKR, while Metcalfe (2014) failed to report an interpretation of the biomechanical meaning of the gait features utilised in their classifier. For these reasons, it is not clear how the findings of Worsley (2011) and Metcalfe (2014) may relate to those of Biggs (2016) and Biggs et al. (2019b). The availability of two comparable datasets of patients offered the opportunity to explore the answer to the following question: *are there similarities in the eighteen highest-ranking features discriminating patients from NPs in two separate patient groups?*

This information would advance the understanding of which are the recurring, and therefore, key gait features discriminating patients with severe knee OA from NPs.

Aim 4: Exploring the gait features discriminating OA from NPs within Cardiff and Karolinska cohorts and determining if there were similarities in the gait features best-discriminating patients pre-TKR from NPs within the Cardiff and Karolinska cohorts.

Aim 5: Exploring the gait biomechanics changes post-TKR in the two patient groups, which of the eighteen highest-ranking gait features improved and which were comparable to NPs.

Taking into account the findings from previous relevant literature (Biggs et al. 2019b):

Hypothesis 4: several gait features improve pre- to post-TKR but patients do not show a gait pattern equivalent to that of NPs at one-year follow-up in either group.

4.2 METHODS

4.2.1 Participants

In this retrospective analysis of prospective, longitudinal studies, participants' data were gathered from two different datasets described below. All participants gave written informed consent before data collection.

Karolinska cohort

Stockholm's Regional Ethical Review Board granted ethical approval for the research. The dataset for these participants was utilised for previous studies (Naili et al. 2017c; Naili et al. 2017b; Naili et al. 2019a). Twenty-nine patients matching the inclusion criteria of suffering from severe primary knee OA and awaiting to receive a TKR were recruited between 2010 and 2013 at Karolinska University Hospital and Ortho Center Löwenströmska Hospital (Stockholm, Sweden). Patients' exclusion criteria were a BMI > 40, neurological disorders, rheumatoid arthritis, previous major surgeries in the lower limbs (such as joint replacements), inability to walk several times unaided for 10 meters, conditions compromising gait (e.g., neurological), and inability to understand written and oral Swedish. Patients were evaluated pre- and one-year post-TKR. Data were analysed on patients who had all biomechanical data needed for this study available, therefore, the final number of patients included was twenty-seven (joint moments and power data could not be calculated for two patients) and only the affected lower limb was analysed. These participants were part of a cohort of twenty-eight patients whose GDI and GDI-kinetic data pre- and twelve months post-TKR have been reported (as mentioned in section 4.1) (Naili et al. 2017b).

Moreover, data were collected on a convenience sample of twenty-five NPs between 2013 and 2015 in the Stockholm area. They were matched to the patients on sex and age strata. NPs were excluded if they had pain in the lower limbs, neurological and/or musculoskeletal conditions. For this Thesis, the relevant data was shared between Dr Josefine Eriksson Naili and the author and comprised pseudonymised data contained in Microsoft Excel sheets (Microsoft Corporation, Redmond, WA, US) including anthropometric, demographic information, questionnaires scores and the extracted time series of joint kinetics, kinematics and GRF relevant for this study. Biomechanical data was extracted for both lower limbs,

for a total of fifty cases. Data sharing was permitted by an amendment to the original ethical approval.

Cardiff cohort

Details on ethical approval, participants' recruitment, inclusion and exclusion criteria were outlined in chapter 3 (sections 3.1 and 3.3). Participants included in this study were recruited between November 2009 and November 2021. As opposed to the Karolinska cohort, nine Cardiff patients (fifteen knees, 34.1% of the total) had already had previous major surgeries in the lower limb and a preliminary analysis showed that their BOA change (median -0.11 ± 0.35 interquartile range - IQR), the main outcome assessed for the responsiveness to change analysis, was statistically no different ($p = 0.379$, non-parametric test) from those who had not had previous major surgeries in the group (median -0.20 ± 0.20 IQR), justifying their inclusion in the current study. Thirty-eight patients were excluded from the study at various stages for the reasons specified in Figure 17. Therefore, data were analysed for thirty-nine patients who had gait data available for both the pre- and one-year post-TKR follow-up and who had good quality 3D gait analysis data. Five patients had a bilateral TKR (three at different times, two during the same surgery) and therefore, data were extracted for both lower limbs in these cases. One participant had a TKR in both knees but was analysed for one knee only as had not attended the 12-month appointment for the second knee. Consequently, forty-four lower limbs were considered both for the pre- and post-TKR analysis. The choice of implant design was at the surgeon's discretion and dependent on the participant's clinical presentation.

Additionally, a convenience sample of thirty-four NPs (not age, nor gender-matched) was recruited between February 2007 and December 2021. NP data were extracted for both lower limbs, except for four participants who did not have data for the contralateral limb due to technical issues, for a total of sixty-four lower limbs.

Part of the data included in this work was collected during the PhD research of Dr Gemma M. Whatling (Whatling 2009), Dr Daniel Watling (Watling 2013) and Dr Paul R. Biggs (Biggs 2016) (Table 2). The data included in this thesis is part of the same longitudinal data collected during, between or after the abovementioned research. From the research of Dr Whatling, nine NPs were included; additionally,

from the work of Dr Watling, nine patients pre-TKR, seven participants one-year post-surgery and nineteen NPs satisfied the current inclusion criteria and were included in the present work. From the work for the thesis of Dr Biggs, the gait biomechanics of six patients pre-TKR and eight participants post-TKR were included in the current research. Furthermore, data on twenty-one participants pre-TKR and ten patients post-TKR were collected between the conclusion of Dr Biggs' PhD and the start of the current research and were included in the present investigation. Also, data on three participants pre-TKR, fourteen patients post-TKR and six NPs were collected by the author and included in this study. The author quality-checked all data already processed for previous projects and re-processed some data to correct minor technical issues.

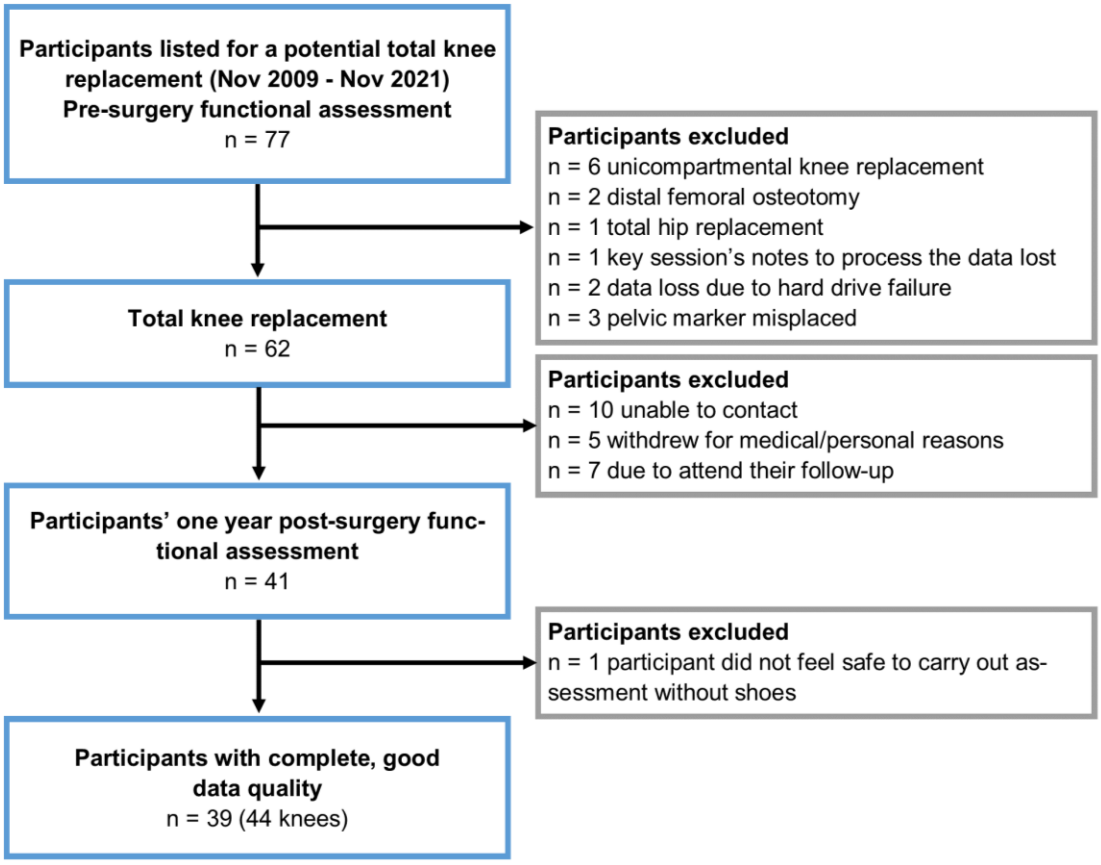


Figure 17 Flowchart of the Cardiff patients included and excluded from the study.

Table 2 Number of participants included in the current PhD research from earlier projects and data collected by the present author

Research data source	Whatling (2009)	Watling (2013)	Biggs (2016)	Current PhD research
Non-pathological volunteers, n	9	23 newly collected	9 from Whatling (2009) + 22 from Watling (2013) = **31**	9 from Whatling (2009) + 19 from Watling (2013) + 6 collected by the current author = **34**
Patients pre-TKR, n	-	25 newly collected	25 from Watling (2013) + 15 newly collected = **40**	9 from Watling (2013) + 6 from Biggs (2016) + 21 from post-Biggs PhD (2016) + 3 collected by the current author = **39**
Patients 9+ months post-TKR, n	-	12 newly collected	10 from Watling (2013) + 12 newly collected = **22**	7 from Watling (2013) + 8 from Biggs + 10 from post-Biggs PhD (2016) + 14 collected by the current author = **39**

n: number; TKR: total knee replacement

4.2.2 Data collection and processing

In both cohorts, all participants walked at their selected speed, barefoot, over a 10 m walkway. Moreover, NPs were evaluated in one session. In both cohorts, technical issues were encountered for some participants' walking trials; therefore, the gait data were averaged from a minimum of three gait cycles, which is a common approach previously utilised by several researchers (McGibbon and Krebs 2002; Mündermann et al. 2005; Biggs et al. 2019b). Biomechanical data were normalised to 101 points, for kinematic data and joint powers expressed as a percentage of the gait cycle and for kinetic and GRF as a percentage of stride (0-100%). Joint moments were expressed as external moments.

Karolinska cohort

Three-dimensional (3D) gait analysis data were collected at the Motion Analysis Laboratory at Karolinska University Hospital (Solna, Sweden), equipped with eight cameras with a sampling frequency of 100 Hz (Vicon Motion Systems Ltd, Oxford, UK) and two force plates (Kistler, Winterthur, Switzerland) embedded in the walkway. The marker set included 35 retro-reflective markers applied to the subjects following the Plug-In-Gait full-body model (Davis et al. 1991). Patients were assessed one-month pre-surgery and one-year after TKR.

Data were processed in Vicon Nexus 1.8.5 (Vicon Motion Systems Ltd, Oxford, UK) and a Woltring filter with a mean squared error of 15 was utilised on markers trajectories. Joint angles were calculated with a YXZ sequence (where Y was the flexion-extension axis) (Cardan-Euler angles), with the proximal segment being fixed and the distal segment moving. Joint kinetics were calculated relative to the distal reaction frame and normalised to the subject's body mass. Forces were exported in a c3d file as % body weight, external moments in N.mm/kg and powers in W/kg. Moment data were originally expressed in Nmm/kg in the Karolinska cohort, the author wrote and run a custom-written code in MATLAB (MathWorks Inc., Natick, Massachusetts, US) to transform the data in Nm/kg for ease of data interpretation. The GRF for the Karolinska cohort was expressed in % body weight and it was transformed in N/kg, using a custom-written code developed by the author, for comparative purposes with the Cardiff cohort.

Cardiff cohort

3D gait analysis took place at the School of Engineering, Cardiff University (Wales, UK). A lower limb modified CAST marker set was employed (therefore, not including the upper quadrant of the body), with the methods described in section 3.3. The author quality-checked all data that had been already processed for previous projects and re-processed some data (n = 7 NPs; n = 12 pre-TKR; n = 11 post-TKR) to correct minor technical issues. Additionally, data on thirteen participants pre-TKR, thirteen patients post-TKR and six NPs were fully processed by the author.

4.2.3 Patient-reported outcomes

Patients in both cohorts completed the Knee Injury and Osteoarthritis Outcome Score (KOOS) which has been previously described in section 3.5.1. In brief, the KOOS is a subjective outcome measure evaluating five knee OA outcomes, including pain, symptoms, activities of daily living (ADL), sports and recreational activities, and knee-related quality of life (QoL) (Roos et al. 1998). Each KOOS subscale normalised score ranges from 0, meaning extreme knee issues, to 100, no knee issues.

4.2.4 Summative measures of gait function

The outcome measures for this study were the patients' Cardiff Classifier classification, GDI and GDI-kinetic scores, described in more detail below. Each group of patients were compared against their respective reference group (i.e., Cardiff patients VS. Cardiff NPs; Karolinska patients VS. Karolinska NPs). The choice of comparing patients to the references collected in the same setting was made because previous studies showed that protocol and marker placement differences may generate slightly diverse results. Kaufman et al. (2016) found that when data were collected on ten healthy participants at three laboratories with diverse motion capture systems and marker placements, there was inter-site variability, mainly involving sagittal hip and knee joint moments. A report comparing the joint kinematics of a single healthy subject collected across twelve sites with different motion capture systems and protocols and found that marker placement differences accounted for more than 75% of gait kinematics variability (Gorton et al. 2009). The authors recommended being cautious when merging data from different laboratories that do not have a common standardised protocol and marker placement, especially in the presence of people with pathological gait and lower limb structural alterations (Gorton et al. 2009).

Following the findings and recommendations from previous research, the analysis of the two cohorts was kept separate for this study. This choice followed much consideration and was made because the different data collection methods utilised could have produced imprecision in the results and affected the interpretation of the data.

For each of the cohorts, both lower limbs of the NPs, where data was available, were utilised to create the reference group, an approach utilised in several previous studies (Schwartz and Rozumalski 2008; Rozumalski and Schwartz 2011; Turcot et al. 2013). Only the affected limb was considered in the patients' group.

GDI and GDI-kinetic

The GDI and GDI-kinetic are summative measures of gait function that compare a subject's gait kinematics or kinetics, respectively, to that of a reference group with a non-pathological gait and calculate how much the gait deviates from the controls. An index ≥ 100 implies normal gait, while 10 points below 100 signify one standard deviation away from the norm and identify the presence of a gait impairment (Schwartz and Rozumalski 2008; Rozumalski and Schwartz 2011). The GDI is based on joint kinematic data, while the GDI-kinetic, as the name suggests, is on joint kinetic data. Both measures were developed to evaluate the gait of children with cerebral palsy and therefore, the gait variables selected to develop these methods, are the most relevant for this population. GDI and GDI-kinetic have the advantage of utilising a gait waveform rather than discrete variables such as peaks and troughs, or joint range of movement (ROM). For each variable time series, data is sampled every 2% so only 51 time points of the original curve are retained (from 0 to 100% gait cycle in 2% increments). The GDI utilises nine gait waveforms, including the pelvis and hip sagittal, frontal and transverse angles, knee and ankle sagittal angles and the foot progression angle (the orientation of the foot in the transverse plane, relative to the lab) (Table **4**).

For each participant in the control group, the nine gait variables are arranged consecutively in one column, in other words, a *gait vector* = 459 x 1 gait vectors (nine gait variables * 51 timepoints = 459 rows) as shown in Eq. 17.

$$\text{Gait vector} = \begin{bmatrix} pelvic\ tilt\ 1-51 \\ pelvic\ obliquity\ 52-102 \\ pelvis\ rotation\ 103-153 \\ hip\ flex-ext\ 154-204 \\ hip\ add-abd\ 205-255 \\ hip\ rotation\ 256-306 \\ knee\ flex-ext\ 307-357 \\ ankle\ dorsi-plan\ 358-408 \\ foot\ prog.angle\ 409-459 \end{bmatrix} \quad \text{Eq. 17}$$

The *gait vector* of each NPs is then combined in matrix **A** with m rows and n columns (where m = 459, from the gait vector, and n is the number of healthy volunteers, Eq. 18).

$$A = [(gait\ vector\ NP1)(gait\ vector\ NP2)\ldots] \quad \text{Eq. 18}$$

The GDI was developed so that all control's gait characteristics (no matter how many controls) could be summarised in a certain number of "feature components", to which a single patient's gait can be compared.
The authors (Schwartz and Rozumalski 2008; Rozumalski and Schwartz 2011) determined the best values to calculate the feature components of each control subject, as explained below. To do this, they performed singular value decomposition on the gait data from 6702 strides, from each side of healthy children's gait. The data were initially arranged as in Eq. 17 and Eq. 18.

Singular value decomposition is a method that reduces data dimensionality by determining which data is most informative and which is redundant and can be disregarded. Singular value decomposition was performed as in Eq. 19.

$$A = U\Sigma V^T \quad \text{Eq. 19}$$

Where:
U is a square m x m matrix of vectors **f** (orthonormal: where each vector is a unit vector and is orthogonal to each other); these are called singular vectors (the left singular vectors are in each column), also defined gait features by (Schwartz and Rozumalski 2008): a 459 x 459 matrix
Σ is a diagonal m x n matrix containing singular values **λ**
V^T is a square n x n matrix, containing the right singular vectors on each row

The feature components (c) were calculated with Eq. 20

$$c = \text{gait vector} \cdot f \quad \text{Eq. 20}$$

The authors found that to maintain enough information from the original data, the first 15 feature components (and therefore the first 15 singular vectors f) were sufficient (representing 98% of the variation in the original data). The authors provided the singular vector values in an addendum available online, organised in a 459 x 15 matrix as in Eq. 21.

$$\text{Features} = \begin{bmatrix} f_1^1 & \cdots & f_1^{15} \\ \vdots & \ddots & \vdots \\ f_{459}^1 & \cdots & f_{459}^{15} \end{bmatrix} \quad \text{Eq. 21}$$

Therefore, given a group of n controls, being each NP gait vector organised as in Eq. 18, fifteen feature components can be calculated with Eq. 20 for each NP. The mean value of each of the fifteen NP's feature components is calculated.

$$NP\ mean\ feature\ components = \begin{bmatrix} c_1^{NP} \\ c_2^{NP} \\ c_3^{NP} \\ \vdots \\ c_{15}^{NP} \end{bmatrix} \quad \text{Eq. 22}$$

To calculate the GDI for a pathological subject, called OA in this example, the gait data is organised as in Eq. 17. Following the same procedure for the NP group, fifteen feature components are calculated from the pathological subject's gait vector with Eq. 20.

$$OA\ feature\ components = \begin{bmatrix} c_1^{OA} \\ c_2^{OA} \\ c_3^{OA} \\ \vdots \\ c_{15}^{OA} \end{bmatrix} \quad \text{Eq. 23}$$

The differences between the pathological subject's and the correspondent mean NP feature components are calculated with Eq. 24, returning fifteen values.

$$\text{Feature components } \Delta = \begin{bmatrix} c_1^{OA} \\ c_2^{OA} \\ c_3^{OA} \\ \vdots \\ c_{15}^{OA} \end{bmatrix} - \begin{bmatrix} c_1^{NP} \\ c_2^{NP} \\ c_3^{NP} \\ \vdots \\ c_{15}^{NP} \end{bmatrix} \quad \text{Eq. 24}$$

The OA GDI raw score for a pathological subject is the scaled distance of the subject's 15 feature components from the average reference group's 15 feature components. This is achieved by calculating the natural log of the Euclidean distance

between the subject and controls' feature components. To aid the interpretability of the OA GDI, the raw value can be converted. Firstly, the NP GDI raw score can be calculated for each NP using Eq. 24, where the mean NP feature components would be subtracted from the NP subject feature components. Then, the natural log of the Euclidean distance between the NP subject and the controls' feature components can be calculated. A z-score is then calculated with Eq. 25.

$$z\ score = \frac{OA\ GDI_{raw} - Mean\ (NP\ GDI_{raw})}{standard\ deviation\ (NP\ GDI_{raw})} \quad \text{Eq. 25}$$

The GDI can then be derived with Eq. 26

$$\text{GDI} = 100 - 10 * \text{z score} \quad \text{Eq. 26}$$

The GDI kinetic follows similar principles to the GDI with two differences, the input variables included (Table **4**), and the number of feature components retained, a total of 20, which represent 91% of the variation from the original kinetic data.

Two MATLAB (MathWorks Inc., Natick, Massachusetts, US) codes were utilised to calculate the patients' GDI and GDI-kinetic scores pre- and one-year post-TKR in both cohorts based on the methods provided by (Schwartz and Rozumalski 2008; Rozumalski and Schwartz 2011). The GDI code had been previously created in the Cardiff research group based on the details from the original GDI research articles and the electronic supplement provided by Schwartz and Rozumalski (2008) (a Microsoft Excel spreadsheet). The GDI-kinetic code was created by the current author by modifying the above-mentioned GDI code, to adapt it to the original research paper on the GDI-kinetic (Rozumalski and Schwartz 2011).

Patients' and NPs' relevant gait data were extracted in Excel files (Microsoft Corporation, Redmond, WA, US) and loaded into the GDI and GDI-kinetic codes. For each NP or patient, the original gait waveforms had been averaged across a minimum of three gait cycles to produce a single entry that was inputted in the GDI or GDI-kinetic MATLAB codes (MathWorks Inc., Natick, Massachusetts, US). Data from both lower limbs (averaged separately) were utilised for NPs, when possible, while only data from the affected side were used for the patients. The Cardiff NP group was utilised as a reference group to calculate the GDI and GDI-kinetic of the

Cardiff patients both pre- and post-TKR. All patients' data (either pre or post-TKR) were entered together into the code but each patient's gait was compared independently to the NP reference group to calculate the GDI (or GDI-kinetic, depending on the code utilised) as per (Schwartz and Rozumalski 2008; Rozumalski and Schwartz 2011) instructions. Finally, an Excel spreadsheet (Microsoft Corporation, Redmond, WA, US) was produced, containing a GDI (or GDI-kinetic) score for each patient. The same procedure was followed for the Karolinska cohort, where patients were compared to their respective NP group.

The reference groups for this study were compared against the original dataset of healthy, typically developing children made available within the original GDI research (Schwartz and Rozumalski 2008) via the abovementioned electronic supplement. This approach was previously utilised by other researchers (Esbjörnsson et al. 2014; Jensen et al. 2015b; Rosenlund et al. 2016a) to validate the control group data. Cardiff's and Karolinska's control NP groups had a mean GDI of 95.3 ± 8.5, and 101.0 ± 9.8, respectively. These findings were in line with previous studies where the control group, compared to the original reference group of NP children, had a mean GDI of 94.4 ± 7.3 (mean NPs age 56.9 ± 7.1) (Rosenlund et al. 2016a), 98.7 ± 8.4 (NPs aged 45 to 68 years) (Jensen et al. 2015b), 99.4 ± 8.3 (mean NPs age 56 ± 15) (Esbjörnsson et al. 2014).

Cardiff classifier

For comparative purposes, the variables utilised for the calculation of the GDI and GDI-kinetic were used in the Cardiff Classifier, together with additional gait features that were deemed important for OA gait classification following previous work (Biggs 2016). In the systematic review from McGinley et al. (2009), it was found that across fifteen studies there was consistently low reliability of hip, knee and ankle gait data in the transverse plane (Intraclass Correlation Coefficient or Coefficient of Multiple Correlation below 0.7). Therefore, it was decided not to include transverse plane data in the Cardiff Classifier (excluding the variables in common with the GDI, as previously mentioned) to ensure a high level of accuracy in the data. For both cohorts, 23 gait variables were utilised to train the classifier, as shown in Table **4**. Cardiff's and Karolinska's cohorts were analysed separately. The reader is redirected to section 3.7 for details on the Cardiff Classifier method. The out-of-

sample classification simplex plots were produced and the out-of-sample classification accuracy was calculated as the proportion of out-of-sample subjects out of the total whose classification matched their class. The sensitivity (Eq. 27) and specificity (Eq. 28) of the classification based on the full dataset was reported for each cohort.

$$Sensitivity = \frac{True\ positive}{True\ positive + False\ negative} \quad \text{Eq. 27}$$

$$Specificity = \frac{True\ negative}{True\ negative + False\ positive} \quad \text{Eq. 28}$$

Table 3 shows a summary of the Cardiff Classifier, GDI and GDI-kinetic key aspects. Table 4 shows the gait variables utilised within the Cardiff Classifier, GDI and GDI-kinetic.

Table 3 Comparison between the Cardiff Classifier, GDI and GDI-kinetic

	Cardiff Classifier	GDI	GDI-kinetic
Data reduction	Principal component analysis	Singular value decomposition	
Feature selection	First 3 principal components for each gait variable	15 gait features (the first 15 singular vectors and therefore feature components) of the whole gait vector	20 gait features (the first 20 singular vectors and therefore feature components) of the whole gait vector
Patient's classification	Supervised training based on controls and patients' 18 gait features with the highest accuracy in discriminating patients and controls and the Dempster-Shafer theory of evidence	Natural log of the Euclidean distance between a single patient and controls' feature components	

GDI: Gait Deviation Index

Table 4 Comparison of the gait variables included in the Cardiff Classifier, GDI and GDI-kinetic

Gait variables	Cardiff Classifier	GDI	GDI-kinetic
Pelvis and hip angles in all three planes	✓	✓	✗
Knee and ankle kinematics in the sagittal plane	✓	✓	✗
Knee and ankle kinematics in the frontal plane	✓	✗	✗
Foot progression angle	✓	✓	✗
Hip, knee, ankle sagittal and frontal planes distal moments	✓	✗	✓
Hip, knee, ankle sagittal distal joint powers	✓	✗	✓
Vertical, anterior-posterior, mediolateral ground reaction forces	✓	✗	✗

GDI: Gait Deviation Index

4.2.5 Statistical analysis

All statistical analyses were performed in IBM SPSS Statistics (version 27, IBM, Armonk, NY, US). Normality distribution was determined via the Shapiro Wilk's test (p-value > 0.05: parametric distribution) and inspection of Q-Q plots. For all statistical tests, the significance level was $\alpha = 0.05$, unless otherwise stated. To compare groups' baseline characteristics and differences between patients and NPs, continuous variables were examined with an independent t-test or its non-parametric equivalent (Mann-Whitney U test) and categorical variables (i.e., sex) with Pearson's Chi-Square test.

<u>Aim 1 – Classifier internal responsiveness and comparison with GDI and GDI-kinetic internal responsiveness</u>

The change score for each outcome measure was calculated by subtracting the score obtained at baseline from the one calculated at one-year follow-up. A negative BOA change and a positive BNP, GDI and GDI-kinetic change denoted gait improvement. A positive Uncertainty change denoted an increase in the classifier classification uncertainty level.

To verify if the changes in the BOA, BNP, Uncertainty values, GDI and GDI-kinetic scores pre to post-TKR were significant, a paired samples t-test or its non-parametric equivalent, the Wilcoxon signed-rank test, was run. The internal responsiveness to change was measured via the ES of the change scores pre to one-year post-TKR. The ES was first introduced by (Cohen 1988) and has been widely utilised as a measure of responsiveness for outcome measures in longitudinal

studies looking at patient-reported outcomes (Giesinger et al. 2014; Peter et al. 2019; Shim and Hamilton 2019), performance-based tests (Mizner et al. 2011), and temporospatial variables (Ornetti et al. 2010) following TKR. The ES is a measure of the magnitude of the difference observed and it was calculated using Eq. 29.

$$ES = \frac{\mu 1 - \mu 2}{\sigma 2} \quad \text{Eq. 29}$$

Where μ1 and μ2 are the group means (post-TKR and pre-surgery, respectively), and σ2 is the standard deviation of the baseline measurement. The ES was interpreted using the guidance from Cohen (1988), where an effect size of 0.20 is interpreted as small responsiveness, 0.50 as moderate and 0.80 large.

Aims 2 and 3 – Relationship between the BOA vs. GDI, BOA vs. GDI-kinetic pre- and post-TKR, BOA change vs. GDI change, BOA change vs. GDI-kinetic change

The relationship between the BOA-GDI score, and the BOA-GDI-kinetic score pre-TKR, post-surgery and scores change was verified via Pearson's correlation test (for normally distributed data) or Spearman's correlation test (for non-normally distributed data). Several correlations were conducted on the same dataset when comparing BOA to the GDI and GDI-kinetic; this can result in an increase in Type I error, which is erroneously rejecting the null hypothesis (i.e., finding differences between groups, or false positive). As previously mentioned, the probability of Type I error was set as α = 0.05 (or 5% probability), however, the chance of false positives increases with the number of comparisons made within a dataset, also called familywise error rate (Portney and Watkins 2014). Therefore, a Bonferroni correction was applied to the significance level to account for multiple comparisons (Sedgwick 2014) (BOA-GDI and BOA-GDI-kinetic = 2 comparisons); considering α = 0.05 for each statistical test before the adjustment, the Bonferroni-adjusted α = 0.05/2 = 0.025.

The strength of the correlation, where significant, was interpreted as suggested by (Dancey and Reidy 2011), where r is the correlation coefficient (r_s, Pearson's correlation coefficient, ρ: Spearman's correlation coefficient):

- |r| < 0.3 meant a weak correlation
- 0.4 < |r| < 0.6 was regarded as a moderate correlation
- |r| > 0.7 was interpreted as a strong correlation

The choice of utilising solely the BOA value to represent the Classifier gait function classification of a patient and the BOA change to represent functional improvement has been common practice in our research group and was justified by preliminary tests on the current Cardiff and Karolinska cohorts. Figure 18a and b show that BOA and BNP values have a statistically significant, almost perfect correlation both pre ($\rho = -0.947$, $p < 0.001$) and post-TKR ($\rho = -0.951$, $p < 0.001$). Additionally, BOA change and BNP change pre to post-TKR have also been shown to be strongly correlated with one another ($\rho = -0.809$, $p < 0.001$), as demonstrated in Figure 18c. Therefore, it was concluded that reporting only on the BOA pre or post-TKR and the BOA change would have been sufficient to describe the gait function at one timepoint or the gait change over time.

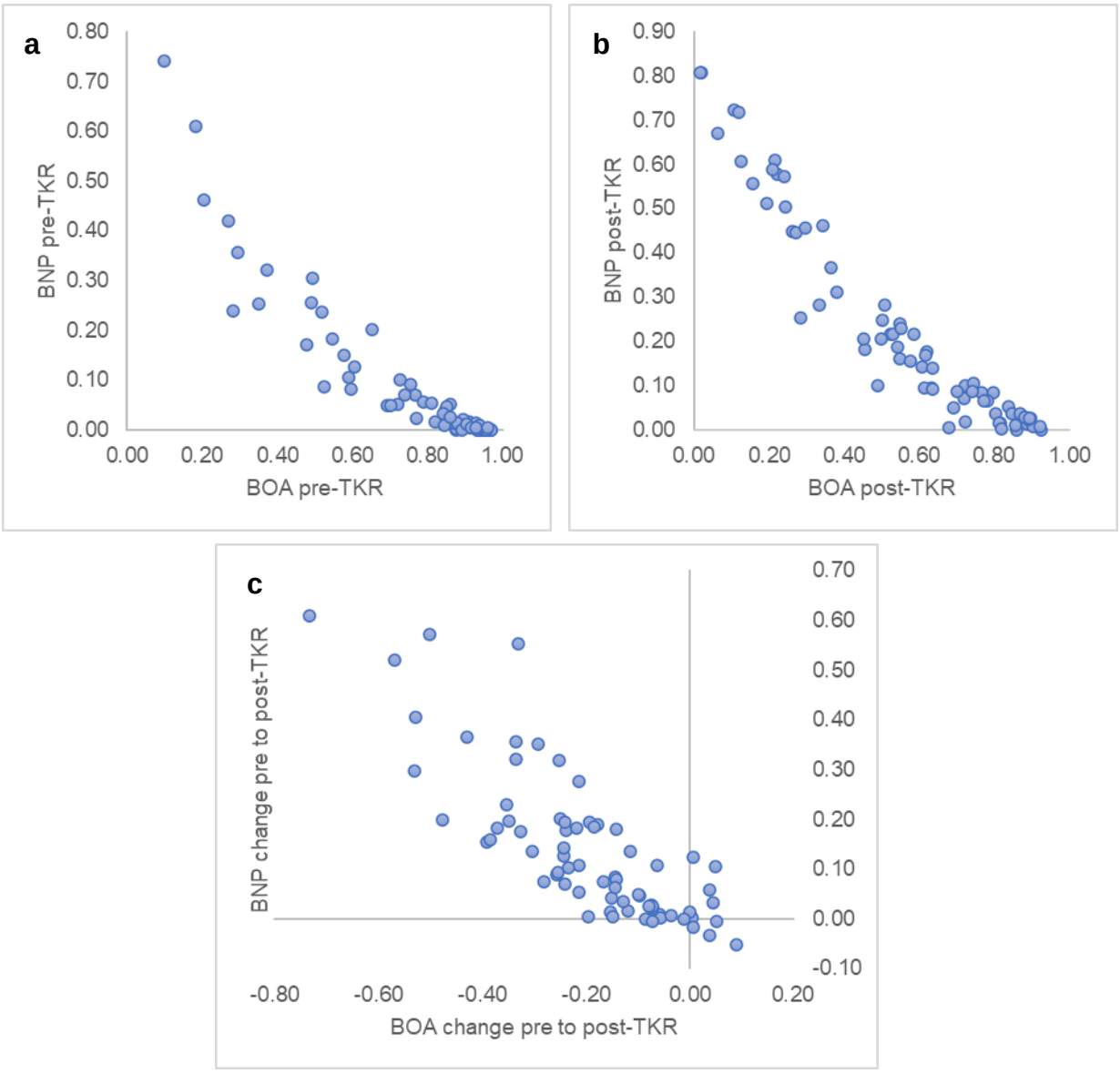


Figure 18 Correlation between the Belief of Osteoarthritis (BOA) and Belief of Non-Pathological (BNP) in the Cardiff and Karolinska patients a) pre-TKR, b) post-TKR; c) Correlation between the change of Belief of Osteoarthritis and the change in Belief of Non-Pathological for Cardiff and Karolinska patients.

Aim 5 – Patients' gait biomechanics changes post-TKR in the two groups

To explore the PC scores differences pre to post-TKR a paired samples t-test or its non-parametric equivalent, the Wilcoxon signed-rank test, was run. To determine whether the gait features returned to a non-pathological level, PC scores differences between patients post-TKR and NPs were assessed via the independent t-test or its non-parametric equivalent (Mann-Whitney U test). A Bonferroni correction

was applied to the significance level to account for multiple comparisons (post-TKR – pre-TKR and post-TKR – NPs = 2 comparisons); considering $\alpha = 0.05$ for each statistical test before the adjustment, the Bonferroni-adjusted $\alpha = 0.05/2 = 0.025$.

4.3 RESULTS

4.3.1 Participants' characteristics

Table 5 shows that pre-surgery, Cardiff patients were significantly heavier, had a significantly higher BMI and could be classed as obese, with a BMI > 30, while the Karolinska patients were overweight (25 < BMI < 29.9), according to the British National Health System (NHS 2019). Cardiff patients' group had half the proportion of women compared to the Karolinska patients and the difference was significant. All Karolinska patients received a cruciate-retaining TKR and so did 68.2% of the Cardiff patients. Surgical notes were not available for 22.7% of the patients from Cardiff (notes were not retrievable for three patients, while six patients, seven knees, had signed a consent form related to a protocol which did not entail accessing the medical records for the purposes of the study) and the remaining 9.1% of participants received a cruciate-sacrificing TKR. In the Cardiff group, nine patients (20.5% of the total) had already had major surgeries in their lower limbs, six in the contralateral limb (three TKR, one UKR, one total hip replacement, one TKR and hip replacement), two patients had had a high tibial osteotomy in both knees, one patient had had a total hip replacement on the same side affected by knee OA.

Table 5 Cardiff and Karolinska patients anthropometrics, demographics

Patients' characteristics	Cardiff patients pre-TKR (n = 44)	Karolinska patients pre-TKR (n = 27)	Significance level
Age (years), median (IQR)	69.0 (8.0)	64 (12)	0.108‡
Height (m), mean (SD)	1.70 (0.10)	1.69 (0.14)	0.882
Weight (kg), median (IQR)	94.0 (30.2)	83.0 (16.0)	0.014‡*
BMI (kg/m^2), median (IQR)	31.4 (10.0)	28.4 (7.4)	0.005‡*
Women, n (% within group)	14 (31.8)	18 (66.7)	0.004*
Implant design			
Cruciate-retaining, n (%)	30 (68.2)	27 (100)	
Cruciate-sacrificing, n (%)	4 (9.1)	0	
Missing details, n (%)	10 (22.7)	0	

TKR: Total Knee Replacement; IQR: interquartile range; SD: standard deviation; m: meters; kg: kilograms; BMI: Body Mass Index; n: number; *: statistically significant at $p < 0.05$; ‡: non-parametric test.

Table 6 shows that Cardiff patients were significantly older, heavier and with a higher BMI than their reference NPs. Moreover, a significantly larger proportion of women was found in the NP group.

Table 6 Cardiff cohort anthropometrics and demographics

	NPs (n = 34) Median (IQR)	Patients pre-TKR (n = 44) Median (IQR)	Significance level
Age (years)	40 (35)	69 (8)	< 0.001*
Height (m)	1.69 (0.09)	1.69 (0.17)	0.596
Weight (kg)	66.5 (15.8)	94.0 (30.2)	< 0.001*
BMI (kg/m^2)	23.44 (5.89)	31.40 (9.95)	< 0.001*
Women, n (%)	22 (64.7)	14 (31.8)	0.006*
Walking speed (m/s)	1.23 (0.12)$^{\times}$	0.82 (0.20)$^{\times}$	< 0.001*

NPs: Non-pathological volunteers; TKR: Total Knee Replacement; IQR: interquartile range; m: meters; kg: kilograms; BMI: Body Mass Index; n: number; $^{\times}$: mean (standard deviation).
*: statistically significant at $p < 0.05$

Table 7 demonstrates that Karolinska patients and their reference NPs had a comparable age. Nevertheless, patients were significantly heavier and had a higher BMI compared to NPs. A similar proportion of women was found between groups.

Table 7 Karolinska cohort anthropometrics and demographics

	NPs (n = 25) Mean (SD)	Patients pre-TKR (n = 27) Mean (SD)	Significance level
Age (years)	65.6 (9.5)	64.4 (7.4)	0.591
Height (m)	1.71 (0.08)	1.69 (0.08)	0.300
Weight (kg)	72.8 (12.2)	72.8 (12.2)	0.002*
BMI (kg/m^2)	24.9 (2.9)	29.7 (4.8)	< 0.001*
Women, n (%)	16 (64.0)	18 (66.7)	0.840
Walking speed (m/s)	1.30 (0.18)	1.09 (0.21)	< 0.001*

NPs: Non-pathological volunteers; TKR: Total Knee Replacement; IQR: interquartile range; m: meters; kg: kilograms; BMI: Body Mass Index; n: number.
*: statistically significant at $p < 0.05$

By visually comparing the information between cohorts, Cardiff NPs were much younger than Karolinska's, but NPs had a similar BMI and there was a similar proportion of women in the NPs from the two cohorts.

Compared to Cardiff NP's gait speed, Cardiff patients walked significantly slower pre-TKR. Similarly, Karolinska's patients pre-TKR were significantly slower than their reference NPs.

The KOOS sub-scores pre-TKR for each subscale could be calculated only for 84.1% of Cardiff patients (n = 37) as the questionnaire was missing pre-surgery (n = 7). Additionally, the KOOS Sports and Recreation sub-score could not be extracted for one Cardiff patient pre-TKR who left too many blank answers in this domain for the score to be calculated. Table 8 shows that both groups of patients had comparable KOOS sub-scores pre-surgery.

Cardiff patients were assessed, on a median, within one month from the surgery (median of 13 days, range 1 to 115 days) and their follow-up was 12.5 months after the surgery, on a median (range: 10 to 30 months post-TKR). For the Karolinska patients, it was not possible to calculate the statistics due to missing data for each participant but, anecdotally, patients were evaluated within one month from the surgery and again twelve months after the operation.

Table 8 Knee Injury and Osteoarthritis Outcome Scores differences between the Cardiff and Karolinska cohorts pre- and post-TKR

Patients' KOOS scores	Cardiff patients pre-TKR (n = 37)	Karolinska patients pre-TKR (n = 27)	Cardiff patients post-TKR (n = 39)	Karolinska patients post-TKR (n = 27)	Significance level (pre-TKR)	Significance level (post-TKR)
KOOS symptoms, median (IQR)	46.4 (25.0)	32.0 (29.0)	85.7 (25.0)	78.6 (32.2)	0.063‡	0.505‡
KOOS pain, mean (SD)	44.8 (20.7)	43.8 (15.6)	84.0 (15.5)	77.1 (20.6)	0.839	0.121
KOOS ADL, mean (SD)	50.1 (20.3)	56.2 (16.1)	82.9 (15.0)	80.7 (19.4)	0.205	0.622
KOOS sports and recreation, median (IQR)	25.0 (30.0)	15.0 (30.0)	55.0 (55.0)	35.0 (50.0)	0.487‡	0.040‡*
KOOS QoL, median (IQR)	25.0 (34.4)	31.0 (12.0)	68.8 (37.5)	62.5 (43.8)	0.652‡	0.633‡

TKR: Total Knee Replacement; IQR: interquartile range; SD: standard deviation; n: number; KOOS: Knee Injury and Osteoarthritis Outcome Survey; ADL: activities of daily living; Knee related QoL: quality of life; *: statistically significant at $p < 0.05$; ‡: non-parametric test.

The KOOS sub-scores post-TKR for each subscale could be calculated only for 88.6% of Cardiff patients (n = 39) as the questionnaire was missing post-surgery (n = 5). Additionally, the KOOS Sports and Recreation sub-score post-TKR was calculated only for thirty-four Cardiff patients as five patients left too many blank answers in this domain for the score to be calculated. In both groups, there was an increase in each of the KOOS sub-scores pre to post-TKR, which was larger than the minimal detectable change in all domains (21.1 for QoL, 19.6 for Sports and Recreational Activities, 15.5 for Symptoms, 15.4 for ADL and 13.4 for Pain) (Collins et al. 2011). Post-TKR, the only difference between the two patient groups was that Cardiff patients had higher scores in the KOOS sports and recreation item.

To avoid confusion between the findings of the Cardiff cohort Cardiff classifier and the Karolinska cohort Cardiff classifier, the Cardiff classifier will be referred to as solely “classifier” in the next sections of this chapter.

4.3.2 Classifier internal responsiveness and comparison with GDI and GDI-kinetic internal responsiveness

Hypothesis 1.1: the classifier’s BOA has a large responsiveness to change in two groups of patients pre- to post-TKR

Table 9 shows that the Classifier was a responsive measure of gait function change in both groups of patients, since BOA, BNP and Uncertainty changed significantly pre- to post-surgery. The BOA change values showed a large ES pre- to one-year post-TKR in both the Cardiff and Karolinska patients. The results confirmed the hypothesis.

For completeness, the BNP and Uncertainty change scores were calculated, the BNP change values showed large ES pre- to one-year post-TKR in both the Cardiff and Karolinska patients. The Uncertainty values increased moderately in both groups after the surgery, denoting a larger uncertainty level in the patients’ gait classification.

Hypothesis 1.2: the classifier's BOA is a more responsive measure of gait change pre to post-TKR compared to the GDI and GDI-kinetic in both patient groups

Table 9 shows that the TKR surgery produced a statistically significant gait improvement in both patient groups one-year post-TKR, as documented by the Classifier and GDI. Additionally, the GDI-kinetic displayed a minor gait improvement after the surgery but it was not statistically significant in either group. The largest responsiveness to change was found in the BOA in both patient groups, while the GDI had a moderate change and the GDI-kinetic had a very small non-significant change pre- to one-year post-surgery. The findings confirmed the hypothesis.

Table 9 Responsiveness to change of the Classifier, GDI and GDI-kinetic

	Measure	Pre-TKR Mean (SD)	Post-TKR Mean (SD)	Difference pre to post-TKR *p*-value	Effect size	Interpretation of the change
Cardiff patients (n = 44)	Classifier					
	BOA	0.77 (0.23)	0.58 (0.27)	< 0.001*	-0.88	Large (decrease)
	BNP	0.02 (0.10)	0.13 (0.39)	< 0.001‡*	1.09	Large (increase)
	Uncertainty	0.15 (0.11)	0.21 (0.09)	< 0.001*	0.55	Moderate (increase)
	GDI	86.6 (8.9)	91.7 (9.2)	0.003*	0.58	Moderate (increase)
	GDI-kinetic	83.8 (9.9)	85.4 (11.8)	0.260	0.16	
Karolinska patients (n = 27)	Classifier					
	BOA	0.73 (0.23)	0.53 (0.25)	< 0.001*	-0.86	Large (decrease)
	BNP	0.03 (0.09)	0.17 (0.28)	< 0.001‡*	0.80	Large (increase)
	Uncertainty	0.17 (0.10)	0.22 (0.08)	0.007*	0.54	Moderate (increase)
	GDI	88.9 (10.5)	95.0 (7.5)	0.008*	0.60	Moderate (increase)
	GDI-kinetic	88.9 (9.6)	92.2 (9.6)	0.153	0.35	

TKR: Total Knee Replacement; SD: standard deviation; n: number; BOA: Belief of Osteoarthritis; BNP: Belief of Non-Pathological; GDI: Gait Deviation Index; ‡: non-parametric test, results are reported in median (interquartile range).
*: statistically significant at $p < 0.05$

4.3.3 Relationship between the BOA vs. GDI and BOA vs. GDI-kinetic pre- and post-TKR

Hypothesis 2.1: there is a significant, negative, moderate correlation between the BOA and GDI pre-TKR in both patient groups

Table 10 shows that there was a significant, negative, relationship between the BOA and GDI values, which was weak for Cardiff patients and moderate in the Karolinska group. The hypothesis was retained for the Karolinska patients but not for the Cardiff ones.

Hypothesis 2.2: there is a significant, negative, moderate relationship between the BOA and GDI-kinetic pre-surgery in both patient groups

A significant, negative, moderate correlation was found between the BOA and GDI-kinetic values pre-surgery equally in the Cardiff and the Karolinska patients (Table 10). The hypothesis was confirmed in both groups of patients.

Hypothesis 2.3: there is a significant, negative, weak correlation between the BOA and GDI post-TKR in both patient groups

Although the GDI and Classifier were linearly related, no significant correlation was found between BOA and GDI either in the Cardiff or Karolinska group post-surgery as shown in Table 10 and, therefore, the hypothesis was disproven.

Hypothesis 2.4: there is a significant, negative, moderate relationship between the BOA and GDI-kinetic post-surgery in both patient groups

A negative relationship was found between BOA and GDI kinetic post-TKR, which was significant and moderate in Cardiff's patients, and not significant in the Karolinska group (Table 10). The hypothesis was confirmed only for the Cardiff patients.

For a visual interpretation of the data, the scatterplots representing the relationship between the BOA vs. GDI and BOA vs. GDI-kinetic pre- and post-TKR described in Table 10 can be found in Appendix C.

Table 10 Correlations between the BOA and GDI, BOA and GDI-kinetic pre- and post-TKR, BOA change and GDI, BOA change and GDI-kinetic change in the Cardiff and Karolinska patients

		Correlation coefficient	Significance level	Correlation strength
Relationship with BOA pre-TKR				
GDI pre-TKR	Cardiff	- 0.372	0.013‡*	Weak
	Karolinska	- 0.553	0.003‡*	Moderate
GDI-kinetic pre-TKR	Cardiff	- 0.623	< 0.001‡*	Moderate
	Karolinska	- 0.523	0.005‡*	Moderate
Relationship with BOA post-TKR				
GDI post-TKR	Cardiff	- 0.121	0.433‡	
	Karolinska	- 0.023	0.911	
GDI-kinetic post-TKR	Cardiff	- 0.624	< 0.001‡*	Moderate
	Karolinska	- 0.313	0.111	
Relationship with BOA change				
GDI change	Cardiff	- 0.152	0.326	
	Karolinska	- 0.199	0.320	
GDI-kinetic change	Cardiff	- 0.465	0.001*	Moderate
	Karolinska	- 0.160	0.425	

BOA: Belief of Osteoarthritis; TKR: Total Knee Replacement; GDI: Gait Deviation Index; ‡: non-parametric test.
*: statistically significant at $p < 0.025$

4.3.4 Relationship between BOA change vs. GDI change, and BOA change vs. GDI-kinetic change

Hypothesis 3.1: there is a significant, moderate correlation between the BOA change and GDI change pre- to one-year post-surgery in both patient groups

Table 10 shows that there was a negative relationship between the BOA change and the GDI score difference pre- to post-surgery, but the correlation was not significant in either patient group. Therefore, the hypothesis was disproven.

Hypothesis 3.2: there is a significant, moderate relationship between the BOA change and GDI-kinetic change pre to one-year post-TKR in both patient groups.

Pearson's correlation revealed that there was a significant, moderate, negative correlation between the BOA change and the GDI-kinetic score change pre- to post-surgery in the Cardiff cohort. However, there was no significant relationship between these variables in the Karolinska patients (Table 10). The hypothesis was confirmed only for the Cardiff patients.

For a visual interpretation of the data, the scatterplots representing the relationship between BOA change vs. GDI change, and BOA change vs. GDI-kinetic change described in Table 10 can be found in Appendix D.

4.3.5 Are there similarities in the eighteen highest-ranking features discriminating patients from NPs in two separate patient groups?

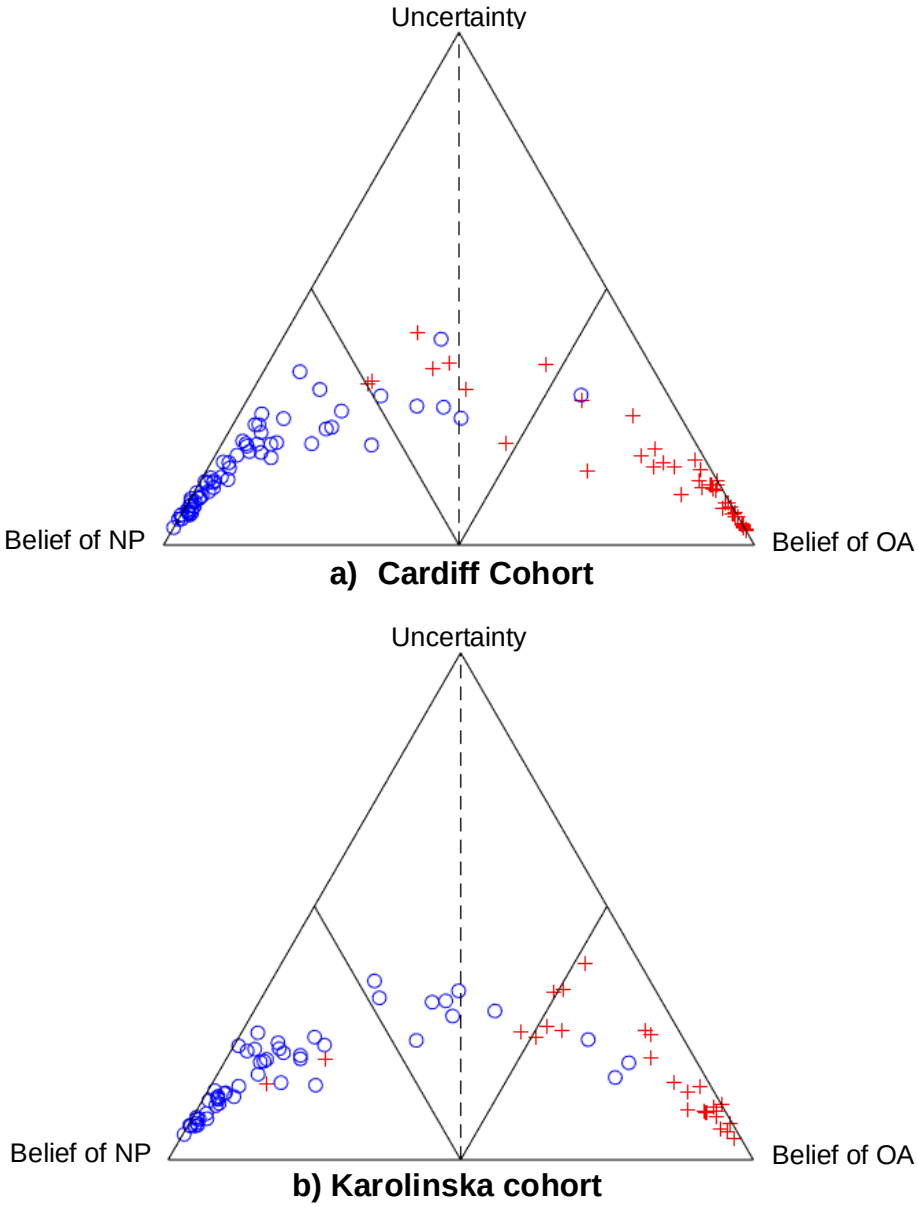


Figure 19 Simplex plots representing the out-of-sample classification accuracy. a) Classifier simplex plot for the Cardiff cohort; b) Classifier simplex plot for the Karolinska cohort. The blue circles represent the non-pathological volunteers, the red crosses the patients pre-surgery. The black, dashed line indicates the boundary between the non-pathological and osteoarthritic area of the plot. The solid lines within the plot, indicate a Belief of osteoarthritis (BOA) and Belief of non-pathological (BNP) of 0.50 (the boundary between dominant BOA and dominant BNP).

Figure 19a shows that the Classifier had a 93.5% out-of-sample classification accuracy in classifying Cardiff participants pre-TKR as OA and NPs (the error was 7 cases out of a total of 108). The classifier accuracy for the Karolinska cohort was similar, showing a 92.2% out-of-sample classification accuracy in classifying participants (6 misclassified lower limbs on a total of 77) (Figure 19b).

In the Cardiff cohort, in the final classification based on the whole dataset, the sensitivity was 93%, and the specificity 98%. It must be noted that no knee radiographs were performed on the NP subjects, and therefore, it was unknown whether the false positives may potentially have had radiographic signs of knee OA. Three patients were classified as having characteristics of a healthy gait (ID 01203, 02575L, 02766) and one NP (ID 02840-left side) was misclassified. NP 02840-left side was in the OA dominant side of the classifier, classed as having an OA gait (this participant had self-reported stiff hips and calves that made it difficult to squat). Table 11 shows that patients ID 01203 and 02575L perceived having less symptoms and pain and had less impact on ADLs due to their knee than the rest of the group. Patients 01203 and 02575L were close to the dominant NP area of the classifier. However, patient ID 02766 had similar scores to the rest of the group, except for perceiving a higher QoL (KOOS QoL).

Table 11 Misclassified Cardiff patients' KOOS scores at baseline

Patients ID	KOOS Symptoms	KOOS Pain	KOOS ADL	KOOS Sports and Recreation	KOOS QoL
01203	57.1	52.8	60.3	15	25
02575 L	71.4	61.1	61.8	50	37.5
02766	46.4	47.2	52.9	25	31.3
Cardiff patients mean (n = 37)	48.4	44.8	50.2	22.9	26.4

KOOS: Knee Injury and Osteoarthritis Outcome Survey; ADL: activities of daily living; Knee-related QoL: quality of life; KOOS = 100 = no knee-related issues, KOOS = 0 = extreme knee-related issues.

In the Karolinska cohort, in the final classification, the sensitivity was 93%, as in the Cardiff cohort, and the specificity was 96%. Two patients were classified as having a non-pathological gait and were in the dominant NP area of the classifier (BNP > 0.50). Additionally, two NPs displayed a pathological gait. Table 12 shows that both patients' symptoms, sports and recreation were much worse than the total patients' mean. Patient K17 had higher pain levels and lower knee-related QoL than

the rest of the group, whereas K34 perceived their ADL and knee-related QoL were less impacted by knee OA than the whole patients' group.

Table 12 Misclassified Karolinska patients' KOOS scores at baseline

Patients ID	KOOS Symptoms	KOOS Pain	KOOS ADL	KOOS Sports and Recreation	KOOS QoL
K17	32.1	33.3	45.6	0.0	18.8
K34	25.0	44.4	79.4	10.0	31.3
Karolinska patients mean (n = 27)	40.1	43.8	56.1	21.1	26.4

KOOS: Knee Injury and Osteoarthritis Outcome Survey; ADL: activities of daily living; QoL: quality of life; KOOS = 100 = no knee-related issues, KOOS = 0 = extreme knee-related issues.

Table 13 and Table 14 show the eighteen highest-ranking PCs discriminating between Cardiff and Karolinska patients and NP pre-TKR, respectively. Most of the highest-ranking gait features differentiating NPs from patients pre-surgery, were related to gait kinetics and GRFs (61.1%, n = 11) in the Cardiff cohort, and in the Karolinska, cohort were related almost equally to kinetic and kinematic aspects, with a slightly higher predominance of gait kinetics (55.6%, n = 10).

The gait features most accurately discriminating people with severe knee OA from reference subjects in the Cardiff group (accuracy > 90%) included vertical GRF PC1 (= patients' loss of the biphasic nature of the waveform with a peak GRF at 50% stance) and hip power PC2 (= patients' earlier power absorption after heel strike, smaller power generation around 60% gait, earlier peak power generation at the end of swing: reduced magnitudes of the waveform during gait).

The gait features most accurately (accuracy > 90%) classifying Karolinska participants were knee flexion-extension angles PC2 (= a larger overall knee ROM in the NP) and, as found in the Cardiff group, hip power PC2 (= NP's larger power generation around 20% gait cycle, earlier peak power generation around 65% gait, delayed peak power absorption around 88% gait cycle).

In both cohorts, excluding GRF, most gait features were related to hip biomechanics (Cardiff: n = 6; Karolinska: n = 5), followed by knee (Cardiff and Karolinska: n = 4) and ankle (Cardiff: n = 3; Karolinska: n = 4) equally in the Karolinska cohort, and pelvis features (Cardiff: n = 2; Karolinska: n = 1). Additionally, excluding joint powers and GRF features, most features were relative to sagittal plane joint angles and moments (six), and four features were related to the frontal

plane in the Cardiff cohort. In the Karolinska cohort, the same number of features were related to the sagittal and frontal planes joints kinetics and kinematics. In both groups, only two features were related to the transverse plane, which is not surprising, given that only pelvis and hip angles in the transverse plane were entered in the classifier.

In both cohorts, most gait features differentiating patients from NPs were PC2 and PC3, depicting variations in gait characteristics (such as delayed peaks and not only magnitude differences), rather than PC1, usually depicting a waveform magnitude difference between groups (Deluzio and Astephen 2007) such as Knee power PC1 seen in Table 15.

For ease of comparison of the eighteen highest-ranking PCs between the two cohorts, Table 15 was produced, showing that for 10 (55.6%) of the 18 highest-ranking gait features the low PC interpretations were the same or similar between the Cardiff and Karolinska groups pre-TKR (shaded in grey); in brackets, it was specified to which group the low PC referred to. Of these, 7 PCs were the same between cohorts and identified the same gait features, including:

- Hip flexion-extension angles PC2 (patients' smaller hip flexion-extension ROM and smaller hip extension around 50% of the gait cycle)
- Hip adduction-abduction angles PC2 (patients' reduced hip adduction-abduction ROM)
- Knee flexion-extension angles PC2 (patients' smaller knee flexion-extension ROM)
- Knee flexion-extension moments PC2 (patients' reduced knee flexion moment around 15-20% stance and reduced extension moment around 65-70% stance)
- Knee power PC1 (describing patients' reduced knee power generation and absorption throughout the gait cycle)
- Vertical GRF PC1 (patients' delayed peak to about 40% and anticipated at 60% of stance - reduced loading rate, higher magnitude around 50% stance: loss of by-phasic nature of the waveform)
- Anterior-posterior GRF PC1 (patients' reduced peak anterior and posterior force)

Three gait features had similar interpretations in the two groups, but were identified by different principal components:

- Ankle dorsi-plantar flexion moments PC1 (Cardiff) and PC2 (Karolinska) (describing patients' larger dorsiflexion moment around 5-20% stance and a smaller peak dorsiflexion moment around 80-85% stance)
- Ankle power PC2 (Karolinska) and PC3 (Cardiff) (patients' increased power absorption in mid-stance and delayed – Karolinska – and reduced – Cardiff – ankle power generation around 55% of the gait cycle)
- Mediolateral GRF PC2 (Cardiff) and PC3 (Karolinska) (patients' reduced loading rate and presence of larger medial force around 50% stance).

Table 13 Cardiff cohort differences in principal component scores between NPs and patients pre-TKR and between patients pre to post-TKR.

Principal components	Variance represented (% of total for the PC)	Ranking based on accuracy (%)	NPs (n=64) Mean (SD)	Pre-TKR (n=44) Mean (SD)	Post-TKR (n=44) Mean (SD)	Difference pre to post-TKR	Effect size	Difference NPs to post-TKR	Effect size
Vertical GRF PC1	60	1 (94)	5.27 (4.19)	-6.61 (7.44)	-2.09 (5.81)	<0.001 ‡*	1.04	<0.001*	-1.88
Hip power PC2	20	2 (91)	2.73 (3.19)	-4.13 (2.67)	-2.85 (2.97)	<0.001*	0.62	<0.001‡*	-1.66
Anterior-posterior GRF PC1	58	3 (87)	4.86 (8.32)	-7.16 (5.33)	-2.04 (5.46)	<0.001*	0.96	<0.001‡*	-1.56
Knee power PC1	29	4 (86)	-3.26 (4.41)	4.74 (2.39)	2.36 (2.13)	<0.001*	-1.00	<0.001*	1.27
Pelvis elevation angle PC2	24	5 (85)	2.90 (2.44)	-4.22 (4.48)	-2.28 (3.38)	<0.001*	0.43	<0.001*	-2.12
Ankle power PC3	15	6 (84)	2.14 (4.27)	-2.81 (2.15)	-1.25 (2.16)	<0.001*	0.72	<0.001‡*	-0.88
Knee flexion-extension moments PC2	24	7 (83)	3.01 (3.32)	-4.37 (3.29)	-2.29 (3.20)	<0.001*	0.63	<0.001*	-1.60
Hip adduction-abduction moments PC2	27	8 (83)	-3.00 (3.04)	4.37 (4.59)	2.93 (4.25)	0.014*	-0.31	<0.001*	1.95
Ankle dorsi-plantar flexion moments PC1	39	9 (83)	3.59 (4.66)	-5.22 (4.44)	-0.20 (3.51)	<0.001*	1.13	<0.001*	-0.81
Knee flexion-extension angles PC2	22	10 (83)	-2.69 (2.72)	3.91 (4.21)	0.68 (3.33)	<0.001*	-0.77	<0.001*	1.24
Hip flexion-extension moments PC2	26	11 (81)	-2.77 (3.36)	4.02 (4.57)	2.83 (4.44)	0.022*	-0.26	<0.001*	1.67
Knee adduction-abduction moments PC2	11	12 (81)	1.66 (2.62)	-2.41 (2.62)	-1.95 (2.48)	0.177	0.18	<0.001*	-1.38
Hip adduction-abduction angles PC2	17	13 (81)	-1.63 (2.67)	3.25 (4.25)	1.77 (2.95)	0.009*	-0.35	<0.001‡*	1.93
Mediolateral GRF PC2	11	14 (79)	1.73 (1.94)	-2.52 (3.39)	-0.75 (2.56)	<0.001*	0.52	<0.001*	-1.28
Ankle dorsi-plantar flexion angles PC1	64	15 (77)	-3.55 (6.77)	4.32 (9.02)	1.62 (5.61)	0.061	-0.30	<0.001‡*	0.81
Hip flexion-extension angles PC2	7	16 (74)	-1.13 (2.69)	2.01 (2.65)	0.34 (2.52)	<0.001*	-0.63	<0.001‡*	1.04
Hip internal-external rotation PC2	5	17 (71)	0.99 (1.90)	-1.44 (2.06)	-0.04 (1.78)	<0.001*	0.68	0.005*	-0.54
Pelvis transverse rotation PC2	24	18 (70)	-1.71 (6.25)	3.00 (4.02)	2.56 (3.59)	0.153	-0.11	<0.001‡*	1.05

GRF: ground reaction force; PC: principal component; NPs: non-pathological volunteers; TKR: total knee replacement; n: number; ‡: non-parametric statistics, median and interquartile range reported. *: statistically significant at $p < 0.025$; in grey the effect sizes of non-significant differences.
The *feature accuracy* of a principal component corresponded to the proportion of subjects who were correctly classified out of the total sample, based on that principal component only. Principal components were ranked from largest to smallest *feature accuracy*.

Table 14 Karolinska cohort differences in principal component scores between NPs and patients pre-TKR and between patients pre to post-TKR.

Principal components	Variance represented (% of total for the PC)	Ranking based on accuracy (%)	NPs (n=50) Mean (SD)	Pre-TKR (n=27) Mean (SD)	Post-TKR (n=27) Mean (SD)	Difference pre to post-TKR	Effect size	Difference NPs-post	Effect size
Knee flexion-extension angles PC2	16	1 (94)	-2.19 (2.87)	4.05 (2.47)	1.99 (2.80)	< 0.001*	-0.83	< 0.001*	1.46
Hip power PC2	20	2 (90)	-2.43 (2.86)	4.50 (3.42)	1.50 (2.91)	< 0.001*	-0.88	< 0.001*	1.37
Knee flexion-extension moments PC2	22	3 (82)	2.33 (3.09)	-4.31 (4.30)	-3.70 (3.44)	0.402	0.14	< 0.001*	-1.95
Ankle inversion-eversion angles PC1	86	4 (81)	4.18 (7.61)	-7.74 (6.96)	-2.90 (9.72)	0.001*	0.84	< 0.001‡*	-0.80
Anterior-posterior GRF PC1	49	5 (78)	3.10 (5.55)	-5.74 (5.92)	-2.34 (5.82)	< 0.001*	0.57	< 0.001*	-0.98
Vertical GRF PC1	58	6 (78)	3.42 (5.79)	-6.33 (6.71)	-2.29 (5.90)	< 0.001*	0.60	< 0.001*	-0.99
Knee power PC1	36	7 (77)	2.76 (4.84)	-6.29 (5.79)	-3.30 (4.53)	0.068‡	0.34	< 0.001*	-1.31
Ankle dorsi-plantar flexion angles PC3	14	8 (77)	1.39 (2.83)	-2.57 (3.82)	-2.25 (3.24)	0.706	0.08	< 0.001*	-1.28
Knee adduction-abduction moments PC2	15	9 (75)	1.58 (2.69)	-2.93 (3.95)	-1.91 (2.38)	0.09	0.26	< 0.001*	-1.30
Hip flexion-extension angles PC2	11	10 (75)	-1.32 (3.17)	2.44 (2.30)	0.85 (3.30)	0.001*	-0.69	0.006*	0.68
Ankle power PC2	17	11 (75)	2.20 (3.64)	-3.08 (5.74)	-0.88 (5.47)	0.055‡	0.22	< 0.001‡*	-0.95
Ankle dorsi-plantar flexion moments PC2	28	12 (73)	-2.29 (3.95)	4.23 (4.89)	1.06 (4.35)	< 0.001*	-0.65	0.001*	0.85
Pelvis elevation angles PC2	25	13 (71)	1.98 (3.43)	-3.68 (5.40)	-2.44 (3.57)	0.006*	0.41	< 0.001‡*	-1.01
Hip adduction-abduction angles PC2	11	14 (71)	-1.55 (3.50)	1.84 (3.81)	-0.19 (3.11)	< 0.001*	-0.53	0.240‡	0.32
Hip adduction-abduction moments PC1	56	15 (70)	2.72 (5.78)	-5.04 (7.77)	-1.31 (8.02)	0.059	0.48	0.013*	-0.70
Foot progression angle PC2	86	16 (70)	-0.91 (3.05)	1.68 (1.70)	1.07 (2.24)	0.003*	-0.71	0.021‡*	0.45
Mediolateral GRF PC3	9	17 (69)	-0.99 (2.63)	1.83 (2.92)	0.42 (3.09)	0.001*	-0.48	0.039	0.53
Hip internal-external rotation angles PC3	5	18 (68)	0.55 (1.96)	-1.02 (2.19)	-0.75 (3.91)	0.865	-0.02	0.010‡*	-0.83

GRF: ground reaction force; PC: principal component; NPs: non-pathological volunteers; TKR: total knee replacement; n: number; ‡: non-parametric statistics, median and interquartile ranges reported.

*: statistically significant at $p < 0.025$; in grey the effect sizes of non-significant differences.

The *feature accuracy* of a principal component corresponded to the proportion of subjects who were correctly classified out of the total sample, based on that principal component only. Principal components were ranked from largest to smallest *feature accuracy*.

Table 15 Comparison of the highest-ranking gait features and their biomechanical meaning between the Cardiff and Karolinska cohorts pre-TKR. Shaded in grey: the low PC interpretations were similar or the same between groups.

			Ranking of the PC based on accuracy (%)		**Interpretation of low PC score (in brackets, the group representing the low PC score)**	
			Cardiff	Karolinska	Cardiff	Karolinska
JOINT ANGLES						
Pelvis	Frontal plane	PC2	5 (85)	13 (71)	Difference: elevation of the pelvis towards the limb in mid-stance and elevating the pelvis towards the contralateral stance limb during all swing phase (opposite to what high PC scores - the NPs - displayed) (pre-TKR)	Difference: reduced drop of the pelvis towards the stance limb (around 15% gait cycle) and reduced elevation towards the contralateral stance limb (around 65% gait cycle): reduced ROM (Pre-TKR)
	Transverse plane	PC2	18 (70)	N/A	Difference: opposed to high PC, presence of rotation away from the stance limb at 0 and 100% and rotation towards the stance limb at 50% gait cycle: larger (peaks) ROM throughout gait (NP)	N/A
Hip	Flexion-extension	PC2	16 (74)	10 (75)	Difference: larger hip extension around 50% of the gait cycle, overall larger ROM (NP)	Difference: larger hip extension around 45% gait cycle, overall larger ROM (NP)
	Adduction-abduction	PC2	13 (81)	14 (71)	Difference: larger hip adduction between 20-45% gait and larger abduction around 70% gait: larger hip ROM (NP)	Difference: larger adduction between 15-45% gait cycle and larger abduction around 70% gait cycle: larger hip ROM (NP)
	Internal-external rotation	PC2	17 (71)	N/A	Difference: presence of internal rotation just after heel strike, external rotation at about 45% gait cycle and internal rotation from 85% gait cycle (all of the above opposite to high PC) (pre-TKR)	N/A
		PC3	N/A	18 (68)	N/A	Difference: presence of external rotation around 25% gait cycle and internal rotation between 55-75% gait cycle (Pre-TKR) (opposite to high PC scores of NP)
Knee	Flexion-extension	PC2	10 (83)	1 (94)	Difference: smaller knee flexion after heel strike and late swing, larger knee peak flexion during swing: larger (peaks) knee ROM (NP)	Difference and phase shift: earlier peak knee extension (around 40% gait cycle) and knee flexion (around 70% gait cycle), larger ROM (NP)
Ankle	Dorsi-plantar flexion	PC1	15 (77)	N/A	Magnitude: the waveform is shifted downwards (larger plantarflexion after heel strike, smaller dorsiflexion during stance, larger plantarflexion at initial swing, smaller peak dorsiflexion from about 85% gait cycle) (NP)	N/A
		PC3	N/A	8 (77)	N/A	Phase shift: delayed peak plantarflexion at the beginning of the swing phase (Pre-TKR)
	Inversion-	PC1	N/A	4	N/A	Magnitude: reduced eversion

	eversion			(81)		throughout the gait cycle (Pre-TKR)
Foot progression		PC1	N/A	16 (70)	N/A	Magnitude: larger external rotation of the foot throughout the gait cycle (NP)
JOINT MOMENTS						
Hip	Flexion-extension	PC2	11 (81)	N/A	Difference: larger hip flexion moment at 5% stance and larger extension moment at 95% stance (NP)	N/A
	Adduction-abduction	PC1	N/A	15 (70)	N/A	Magnitude: reduced adduction moment magnitude especially around 30 and 85% stance (Pre-TKR)
		PC2	8 (83)	N/A	Difference and phase shift: lower adduction moment magnitude in mid-stance and higher loading rate (NP)	N/A
Knee	Flexion-extension	PC2	7 (83)	3 (82)	Difference: reduced flexion moment around 20% stance and reduced extension moment around 70% stance (pre-TKR)	Difference: reduced flexion moment at 15% stance and reduced extension moment at about 65% stance (Pre-TKR)
	Knee adduction-abduction	PC2	12 (81)	9 (75)	Difference: smaller adduction moments at 20% and 90% stance and larger adduction moments around 45% stance (pre-TKR)	Phase shift: presence of one single peak adduction moment (around 50% stance), delayed time to peak and earlier termination of peak adduction (Pre-TKR)
Ankle	Dorsi-plantar flexion	PC1	9 (83)	N/A	Difference: larger dorsiflexion moment around 20% stance and a smaller peak dorsiflexion moment around 85% stance (pre-TKR)	N/A
		PC2	N/A	12 (73)	N/A	Difference: larger plantarflexion moment around 5% stance and larger dorsiflexion moment around 80% stance (NP)
JOINT POWERS						
Hip		PC2	2 (91)	2 (90)	Difference and phase shifts: earlier power absorption after heel strike, smaller power generation around 60% gait, earlier peak power generation at the end of swing: reduced magnitudes of the waveform during gait (pre-TKR)	Difference and phase shift: larger power generation around 20% gait cycle, earlier peak power generation around 65% gait, delayed peak power absorption around 88% gait cycle (NP)
Knee		PC1	4 (86)	7 (77)	Difference: larger peak power absorption and generation throughout the gait cycle (NP)	Difference: reduced power generation and absorption throughout the gait cycle (Pre-TKR)
Ankle		PC2	N/A	11 (75)	N/A	Difference: mainly larger power absorption around 20% gait cycle, delayed and reduced peak power generation around 55% gait cycle, difficult to interpret the meaning at 65-80-95% of the gait cycle (no apparent differences) (Pre-TKR)
		PC3	6 (84)	N/A	Difference: mainly larger power absorption around 10% gait cycle, smaller power generation around 55% gait cycle, difficult to interpret the meaning at 80 and 90% of the gait cycle (no apparent differences) (pre-TKR)	N/A
FORCES						
Ground reacti	Vertical	PC1	1 (94)	6 (78)	Phase shift: delayed at about 40% and anticipated at 60% of stance (reduced loading rate),	Phase shift: delayed at about 40% and anticipated at 60% of stance (reduced loading rate), higher

on force					higher magnitude around 50% stance: loss of by-phasic nature of the waveform (pre-TKR)	magnitude around 50% stance: loss of by-phasic nature of the waveform (Pre-TKR)
	Anterior-posterior	PC1	3 (87)	5 (78)	Difference: reduced peak anterior and posterior force (pre-TKR)	Difference: reduced magnitude of anterior and posterior force peaks (Pre-TKR)
	Medio-lateral	PC2	14 (79)	N/A	Difference and phase shift: reduced loading rate and presence of larger medial force around 50% stance (pre-TKR)	N/A
		PC3	N/A	17 (69)	N/A	Difference and phase shift: faster loading rate and presence of smaller medial force around 50% stance (NP)

NP: non-pathological participants; PC: principal component; TKR: total knee replacement; N/A: not available; ROM: range of motion.

4.3.6 Patients' gait biomechanics changes post-TKR in the two groups

Hypothesis 4: several gait features improve pre- to post-TKR but patients do not show a gait pattern equivalent to that of NPs at one-year follow-up in either group

There was a higher proportion of patients classed as having an NP gait in the Cardiff cohort (25.0%, n = 11) one-year after the surgery than in the Karolinska cohort (22.2%, n = 6). However, three (Cardiff) and two (Karolinska) of the patients had been classed as having an NP gait before the TKR surgery.

One-year after the recovery, fifteen (83.3%) and eleven (61.1%) of the gait features improved significantly for the Cardiff (Table 13) and Karolinska patients (Table 14), respectively, becoming closer to that of NPs. This confirmed the hypothesis. For the Cardiff patients, the improved gait features included mainly gait kinetics (n = 10, 66.7% of the improved gait features) (all three components of the

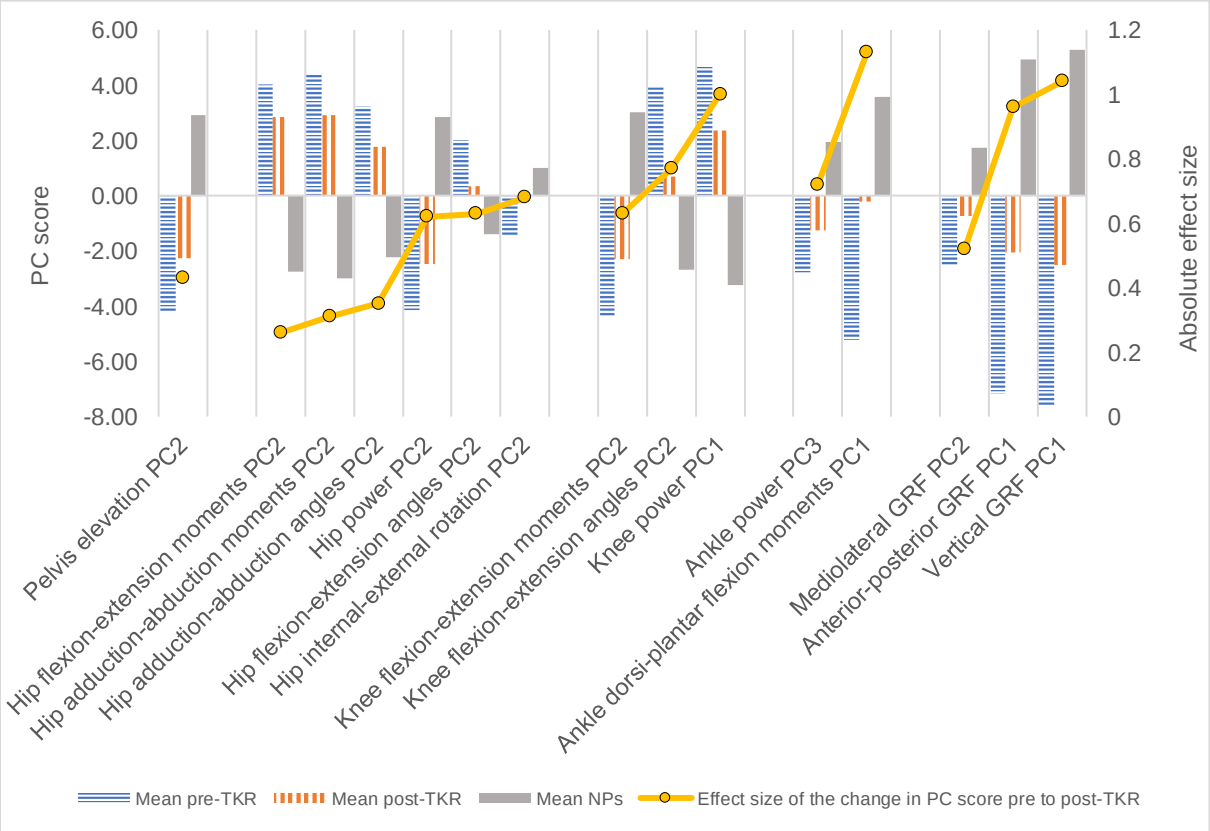


Figure 20 Cardiff cohort mean principal components (PC) utilised in the classifier for patients pre, post-total knee replacement (TKR) and non-pathological subjects (NPs). Only the PCs that changed significantly pre to post-TKR are displayed, by joint, in increasing order of improvement (i.e., from smallest to largest effect size).

GRF, all three PCs relative to joint powers and four PCs identifying joint moments); moreover, most of the features which improved included hip biomechanics, followed by knee, ankle and pelvis biomechanics (Figure 20).

In the Cardiff group, the gait features showing the largest change pre- to post-surgery were mainly related to kinetics and forces (ES > 0.8), including ankle dorsi-plantarflexion moments PC1, vertical GRF PC1, followed closely by knee power PC1 and anterior-posterior GRF PC1 (Figure 20).

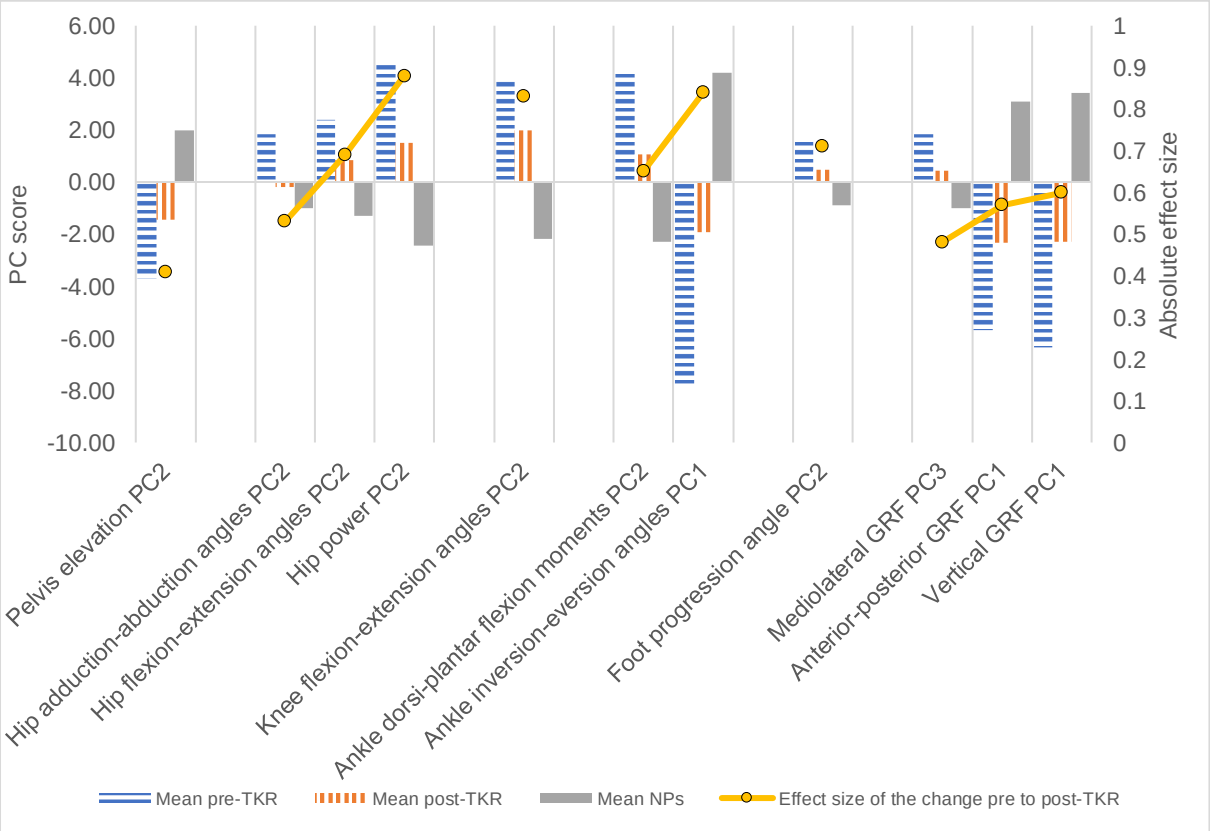


Figure 21 Karolinska cohort mean principal components (PC) utilised in the classifier for patients pre, post-total knee replacement (TKR) and non-pathological subjects (NPs). Only the PCs that changed significantly pre to post-TKR are displayed, by joint, in increasing order of improvement (i.e., from smallest to largest effect size).

Compared to the Cardiff patients, in the Karolinska ones, a larger number of gait features (mainly joint kinetics) did not improve significantly pre- to post-surgery. In the Karolinska patients, the improved gait features included a similar representation of gait kinematics and kinetics, with a slightly larger proportion of gait kinematics (six PCs identifying joint angles, 54.6% of the improved gait features). Similar to the Cardiff group, all three components of the GRF improved significantly

pre to post-TKR. Additionally, like in the Cardiff group, most of the improved gait features included hip biomechanics, ankle biomechanics followed, and then there was an improvement in pelvis, knee and foot progression angles (Figure 21). In the Karolinska group, the gait features with the greatest significant change (ES > 0.80) pre- to post-surgery were hip power PC2, followed closely by ankle inversion-eversion angles PC1 and knee flexion-extension angle PC2 (Figure 21).

Following the hypothesis, despite the improvements, no gait features were comparable between patients post-TKR and NPs in the Cardiff group. The gait feature showing the largest difference between NPs and Cardiff patients post-TKR was the pelvis elevation PC2 (describing elevation of the pelvis towards the limb in mid-stance and elevating the pelvis towards the contralateral stance limb during all swing phase - opposite to the high PC scores - the NPs - displayed), followed by hip adduction-abduction moments PC2 (describing NP's lower adduction moment magnitude in mid-stance and higher loading rate), hip adduction-abduction angles PC2 (depicting NP's larger hip adduction between 20-45% gait and larger abduction around 70% gait) and vertical GRF PC1 (describing patients' reduced loading rate, higher magnitude around 50% stance and loss of by-phasic nature of the waveform), which was also the gait feature most accurately differentiating NPs from patients with severe knee OA before surgery (94% accuracy).

Only hip adduction-abduction angle PC2 (describing hip adduction-abduction ROM) and mediolateral GRF PC 3 (describing loading rate and medial force around 50% stance) were comparable between Karolinska patients post-TKR and NPs, whereas the remaining patients' gait features were significantly different from controls.

Most of the differences between Karolinska patients post-TKR and NPs had a large effect size. The gait feature showing the largest difference was the knee flexion-extension moments PC2 (very large difference, describing patients reduced flexion moment at 15% stance and reduced extension moment at about 65% stance), followed closely by knee flexion-extension angles PC2 (very large difference, describing patients' reduced knee ROM throughout the gait cycle), and hip power PC2 (very large difference, describing NP's larger power generation around 20% gait cycle, earlier peak power generation around 65% gait, delayed peak power absorption around 88% gait cycle). Interestingly, the above-mentioned

features were the first three in terms of accuracy in differentiating NP from patients with severe knee OA before surgery.

4.4 DISCUSSION

4.4.1 Classifier internal responsiveness and comparison with GDI and GDI-kinetic internal responsiveness

Hypothesis 1.1: the Cardiff Classifier's BOA has a large responsiveness to change in two groups of patients pre- to post-TKR

One of the aims of this study was to determine the internal responsiveness of the Classifier in two cohorts of patients one-year after TKR surgery. The BOA showed a large responsiveness in both the Cardiff and Karolinska patients, indicating that the classifier is sensitive to gait biomechanics changes one-year post-operation. Pre-surgery, patients in both groups had a BOA > 0.70 and, meaning they had a pathological gait. The BOA decreased significantly in both patient groups post-TKR. Nevertheless, the BOA mean one year after the surgery was greater than 0.50 (Karolinska: 0.53; Cardiff: 0.58), indicating that patients did not go back to an NP gait and that the majority was still in the dominant OA area of the Classifier (BOA > BNP). This was in line with what was previously found by Biggs et al. (2019a) who observed a significant improvement in gait function in twenty-two patients, with a BOA mean of 0.51 (0.20 SD) post-TKR, showing that most of them had a pathological gait. The outcomes from both patient groups also confirm the additional findings by (Biggs et al. 2019a) demonstrating a large responsiveness of the BOA (r = - 0.848) pre to 9+ months post-TKR (n = 22). It must be noted that that 35.9% of the patients pre-TKR (n = 14), 35.9% of post-TKR participants (n = 14) and 82.4% of NPs (n = 28) included in the current study were the same as those included in (Biggs et al. 2019a), but the present investigation was carried out on a larger number and slightly different patients and NPs (Table 2).

Interestingly, the classifier responsiveness was similar between the Cardiff and Karolinska cohort, even if the Cardiff patients were compared to a non-age-matched control group and the Karolinska were compared to an age-matched reference group. Ageing produces changes in gait biomechanics that are similar to those observed in people with OA (more on this will be discussed below) (Boyer et al. 2017a; Herssens et al. 2018; Pol et al. 2021a), therefore, it may have been expected that Karolinska patients were going to change their classification within the Classifier to a larger extent than Cardiff's patients, who were compared to a much younger group of controls having a higher level of function. Nevertheless, Karolinska

patients had a lower BOA at baseline (=less pathological gait), compared to Cardiff patients, therefore, there was less room for improvement in the Karolinska patients and this may explain the similar responsiveness of the Classifier observed in the two cohorts.

In the current study, a moderate increase in the patients' classification uncertainty level was observed; this seems to be a recurring result that was found in previous research utilising the Classifier pre- and post-surgery for knee and hip OA. Biggs et al. (2019a) observed forty-one participants pre- and twenty-two post-TKR and found a moderate increase in the Uncertainty (significant mean change from 0.29 to 0.35) after the surgery. Likewise, Biggs et al. (2021) found that the classification uncertainty increased significantly pre (Uncertainty = 0.27) to one-year post-total hip replacement (Uncertainty = 0.39). Metcalfe et al. (2017) assessed twenty patients before and fifteen patients six months after TKR and UKR and found that the patients who were classed as NP post-surgery had a significantly higher Uncertainty than those who remained in the dominant area of the Classifier (BOA > 0.5). The results of the present study may indicate that although joint kinetics and kinematics do not go back to normal, the patients' biomechanics becomes more heterogeneous. In fact, a high level of uncertainty indicates the presence of contrasting information regarding the biomechanical data, with some variables not being able to successfully separate patients from NP due to mixed characteristics, some of which would class the participant as having an osteoarthritic and with some aspects that resemble NP gait features.

Hypothesis 1.2: the Cardiff Classifier's BOA is a more responsive measure of gait function change pre to post-TKR compared to the GDI and GDI-kinetic in both patient groups

In agreement with the hypothesis, the classifier was more responsive to gait biomechanics changes (i.e., BOA change) in two similar groups of patients post-TKR, than the GDI and GDI-kinetic. The second most responsive measure was the GDI, showing a moderate improvement in gait kinematics in both groups. However, the GDI-kinetic performed poorly in terms of responsiveness, the score increase was not significant, but it showed a trend towards a small recovery of gait kinetics in both cohorts. Although the same gait variables used in the GDI and GDI-kinetic were

utilised within the Classifier, the latter allowed adding additional biomechanics known to be relevant in knee OA patients' classification, such as GRF. In fact, it was found that vertical and anterior-posterior GRF were two of the six most accurate classifying features of patients with knee OA (Table 15). GRF is an important feature describing OA gait, as proven by previous reports (Yoshida et al. 2008; Biggs et al. 2019b), and the improvement in anterior-posterior and vertical GRF had a large ES in both patient groups post-TKR, indicating a considerable change in this variable. However, GRF is not included in the GDI-kinetic and this may reduce its sensitivity in detecting change pre to post-TKR. A previous study comparing individuals with knee OA to a reference group created a knee specific GDI, where only knee angles and moments were inputted to obtain a GDI score (Kobsar et al. 2019). It must be noted that this new GDI combined gait kinetics and kinematics and it is conceptually different from the original GDI, where only gait kinematics were included (Schwartz and Rozumalski 2008). This study (Kobsar et al. 2019) showed that it was possible to modify the GDI to add the variables relevant to the population studied. Therefore, future research may potentially include vertical and anterior-posterior GRF to the GDI-kinetic. This would allow to verify whether introducing GRFs, may make the new GDI-kinetic more sensitive in assessing change in gait pre to post-TKR.

Considering gait kinematics and kinetic improvement separately may give the advantage of determining which gait aspects improved after the surgery but reducing the ability to detect a change in gait function. The classifier had the advantage of combining gait kinematics and kinetics in one measure and based the classification of patients on the gait features that were most relevant in distinguishing patients from NPs. This made the classifier OA-specific, while GDI and GDI-kinetic were developed to evaluate the gait of children with cerebral palsy and therefore, utilised more generic gait features that did not include all the relevant biomechanical features for people with knee OA.

The results of the current investigation are in line with previous work (Naili et al. 2019b) where there was a moderate improvement in the GDI (ES = 0.7) and a small, not significant improvement in the GDI-kinetic (ES = 0.4) one-year after total hip replacement. Conversely, Naili et al. (2017b) found a small change in the GDI (ES = 0.4) and a moderate improvement in the GDI-kinetic (ES = 0.5) twelve months post-TKR. It must be noted that the patients included in Naili et al. (2017b) were the

same as those examined in the current study, however, for both patients and NPs, the GDI and GDI-kinetic were calculated against a different reference of fifty-nine subjects and this may justify the dissimilar findings from the present study.

GDI, GDI-kinetic and classifier determine gait function with different methods. GDI and GDI-kinetic compare each patient, separately, to a reference group of healthy volunteers, and establish how “far” the patient is relative to the reference group biomechanics (Schwartz and Rozumalski 2008; Rozumalski and Schwartz 2011). The classifier is trained to distinguish patients and NPs and then classifies them according to the main gait features characterising OA and NPs. This means that each subject is compared both to a group of NPs and a group of people with severe knee OA and this produces a system of Believes (BOA, BNP, Uncertainty) that the subject has NP/OA characteristics, also accounting for uncertainty in the classification. Having the option to check the level of uncertainty in the classification is important because it allows to make a more informed interpretation of the results and has an impact on the confidence the user may have in making a decision based on the output. For instance, if the classifier was to be implemented for clinical use, the occurrence of a high uncertainty level would prompt a clinician to examine the original gait data in more detail to verify the presence of potential errors, or to carry out further assessments on the patient to obtain additional details and clarify the observation made by the classifier.

Despite the different methods, the construct that GDI, GDI-kinetic and Classifier measure should be the same and corresponding to gait function in relation to healthy subjects and this justified comparing these measures.

The findings suggest that combining joint kinematics, kinetics and GRF into a single measure like the Classifier seems to capture changes in objective function more effectively than looking at kinematic and kinetic variables separately with the GDI and GDI-kinetic, respectively.

4.4.2 Relationship between the BOA vs. GDI and BOA vs. GDI-kinetic pre- and post-TKR

Hypotheses 2.1, 2.2: there is a significant, negative, moderate correlation between the BOA and GDI pre-TKR in both patient groups

In line with the initial hypothesis, pre-TKR, there was a significant association between having a high BOA (i.e., more compromised gait) and a low GDI (i.e., more compromised gait) for the Karolinska and Cardiff patients. When considering the GDI-kinetic, a moderate relationship with the BOA was found in both groups pre-TKR, in agreement with the hypothesis. In the Cardiff cohort, the classifier classification of patients and NPs was mainly based on gait kinetics features (joint moments, GRF and joint powers, about 60% of the total) (Table **13**). This could explain the stronger relationship between the Cardiff patients' GDI-kinetic and BOA (moderate correlation) and the weak relationship between the GDI and BOA. In the Karolinska patients, the classifier found that a similar proportion of joint kinematics and gait kinetics (about 55%) features were separating patients with severe knee OA from NPs. Therefore, the subjects' classification of gait was almost equally determined by kinetic and kinematic characteristics, justifying why a moderate correlation was found between the BOA and GDI and BOA and GDI-kinetic. The significant correlations between both the BOA and GDI, BOA and GDI-kinetic could also highlight that joint kinetics and kinematics were similarly affected before TKR surgery as noted in previous studies (Astephen et al. 2008a; Ro et al. 2019). Previous work on a larger OA cohort, described joint kinematics alterations such as reduced ROM at the hip, knee and ankle, coupled with joint kinetics differences from controls, like smaller knee extension and internal rotation moments in terminal stance, knee flexion moment in initial stance and hip adduction moments in mid-stance (Astephen et al. 2008a). The current results suggest that classifier, GDI and GDI-kinetic can inform on gait quality pre-TKR to a similar degree, and therefore the use of the Classifier or the combination of GDI and GDI-kinetic at this stage may be at the researcher's discretion, bearing in mind that when looking at longitudinal changes, the classifier is a more responsive measure of gait function change.

<u>Hypothesis 2.3: there is a significant, negative, weak correlation between the BOA and GDI post-TKR in both patient groups</u>

The hypothesis was disproven as no relationship was found between the BOA and GDI in either the Cardiff or Karolinska patients post-TKR. This may suggest that there was poor agreement between the measures. There were cases where patients showed a GDI approaching or exceeding 100 (= normal gait), indicating that gait

kinematics were comparable to controls, while also having a BOA > 0.5, demonstrating that a combination of gait kinetics and kinematics lead to a pathologic gait classification of most patients.

Hypothesis 2.4: there is a significant, negative, moderate relationship between the BOA and GDI-kinetic post-surgery in both patient groups

The hypothesis was accepted for the Cardiff patients only, as a moderate correlation was found between the classifier and GDI-kinetic; in the Karolinska cohort, a negative, linear relationship was observed between the BOA and GDI-kinetic, suggesting a similar trend to the Cardiff cohort, however, no significant correlation was found.

Surprisingly, the GDI showed that patients' gait quality was only moderately compromised post-TKR ($90 < GDI \leq 95$), given that their mean scores were less than one SD away from the NPs in the Karolinska (mean GDI: 95.0 ± 7.5) and the Cardiff cohort (mean GDI: 91.7 ± 9.2). Cardiff patients' GDI-kinetic mean score (85.4 ± 11.8) was about 1.5 SD away from the reference healthy group. Karolinska patients' joint kinetics, similar to what was found for the joint kinematics, were only mildly compromised as the GDI-kinetic mean score (92.2 ± 9.6) was less than 1SD away from the controls. Nevertheless, looking at the classifier, despite the improvement, gait function was far from normal one-year after the surgery, as shown by the relatively high BOA in both groups of patients (Karolinska: 0.53; Cardiff: 0.58), indicating that most were in the dominant area of the Classifier, meaning that their gait was severely affected and mainly still showing OA features. The advantage of the classifier is that it allows to examine which gait features were involved in determining the classification of gait function. When considering the patients biomechanics post-TKR, it was found that no kinetics or kinematics features were comparable to those of NPs in the two patient groups, except for Karolinska patients' hip adduction-abduction ROM, and medio-lateral GRF. GDI and GDI-kinetic do not allow to visualise which gait features were key in determining the index, and therefore, it is unknown what could have determined such a high GDI and GDI-kinetic index in both patient groups post-TKR.

One-year after the surgery, the uncertainty level of the Classifier increased significantly showing that classifying subjects becomes more complex at this stage of

recovery due to a higher degree of contrasting information in the gait features. The lack of correlation between the Classifier and GDI in both groups suggests that joint kinetics and kinematics were affected to a different degree one-year post-TKR. This agrees with previous research (Saari et al. 2005b; Yoshida et al. 2008; Levinger et al. 2013; McGinnis et al. 2013) reporting mixed recovery in terms of gait kinematics and kinetics and contrasting results following TKR surgery. Yoshida et al. (2008) studied twelve patients before and one-year after TKR and found that knee ROM, peak flexion and knee flexion at initial contact were comparable to that of twelve age-matched controls, while vertical GRF, knee and hip extension moments at peak knee flexion were significantly different in the patients' group. Saari et al. (2005) observed that up to two years post-TKR, thirty-nine patients had reduced hip extension compared to eighteen age-matched controls, however, peak hip flexion, adduction-abduction, peak knee flexion-extension moments and peak angles did not differ between groups. McGinnis et al. (2013) found significant differences in knee flexion-extension ROM at loading response but no differences in knee peak flexion-extension moments, hip and ankle flexion-extension excursion in patients six months after TKR compared to twenty-one, healthy volunteers. Levinger et al. (2013) observed smaller knee peak extension moments, reduced knee extension angles and knee power absorption in late stance, and increased ankle dorsiflexion in thirty-two patients one-year after TKR, nevertheless, no differences from the control group (twenty-eight subjects) were found in hip angles, moments and power, or ankle moments and power.

Considering that the Classifier combines gait kinematics and kinetics, the moderate correlation between the BOA and GDI-kinetic in the Cardiff patients only may suggests that gait kinetics may play a more substantial role in defining gait quality post-TKR compared to gait kinematics in certain patient groups. The current findings indicate that Classifier and GDI produce different outputs post-TKR and therefore it is not advised to utilise them interchangeably. Caution should be applied when considering utilising the GDI-kinetic in place of the Classifier as in the current cohorts of patients, the two measures were not associated in the Karolinska cohort but there was a moderate correlation in the Cardiff cohort, highlighting the presence of diverse biomechanical characteristics of the two groups of patients. Moreover, caution should be used when utilising the GDI and GDI-kinetic, since both indexes

showed a return to seemingly normal biomechanics one-year post-TKR, which was not corroborated by the findings of the classifier.

4.4.3 Relationship between BOA change vs. GDI change, and BOA change vs. GDI-kinetic change

Hypothesis 3.1: there is a significant, moderate correlation between the BOA change and GDI change pre- to one-year post-surgery in both patient groups

While a linear relationship could be observed between the BOA change and GDI variation post-TKR, the results negated the hypothesis as no significant correlation was found between the variables in either group. This means that the extent of gait kinematic change (GDI change), did not correspond to a similar change in gait function when gait kinetics and kinematics were considered together (BOA change).

Hypothesis 3.2: there is a significant, moderate relationship between the BOA change and GDI-kinetic change pre to one-year post-TKR in both patient groups

Regarding the association between the BOA and GDI-kinetic scores changes, the hypothesis could be confirmed for the Cardiff patients only, as a significant, moderate relationship was observed. This was likely because the gait features improving post-TKR and the ones with the largest change were mainly related to joint kinetics in this group. Contrarily, no correlation between the variables was found in the Karolinska group, although a linear, negative association could be observed between the BOA and GDI-kinetic change. In some cases, patients that were deemed to have had a gait function improvement with one outcome measure, were classed as having had a gait deterioration by the other outcome measure.

The lack of correlation between the BOA and GDI change in both groups and the mixed findings in the relationship between BOA and GDI-kinetic change suggest that patients had a mixed recovery, where a change in gait kinematics was not complemented with a similar variation in gait kinetics or vice versa, as previously discussed and as demonstrated by other reports (Saari et al. 2005b; Yoshida et al. 2008; Levinger et al. 2013; McGinnis et al. 2013). Therefore, it is not advised to utilise the GDI and classifier interchangeably to evaluate the gait quality difference pre to post-TKR. Additionally, it is advised to utilise the GDI-kinetic cautiously to

evaluate change pre to post-TKR as, in some groups of patients, it may not be able to measure a change in gait kinetics and additionally, this study demonstrated that the GDI-kinetic had poor internal responsiveness in two groups of patients one-year post-TKR.

GDI and GDI-kinetic are useful in determining a subject's gait quality, but they do not provide information on which gait features are deviating from the reference group, as opposed to the Classifier. This is valuable information to understand what could be potentially targeted by rehabilitation interventions. A recent NICE guideline recommended introducing a preoperative rehabilitation program before TKR; while rehabilitation may not improve pain and patient-perceived function pre-TKR, it revealed a reduction in hospital length of stay, which would be of clinical benefit to the patient (NICE 2020a). Measures such as the classifier may aid in informing pre-rehabilitation programs, by highlighting the gait features that most differ from healthy subjects.

4.4.4 Are there similarities in the eighteen highest-ranking features discriminating patients from NPs in two separate patient groups?

The Classifier out-of-sample classification accuracy was similar between the Cardiff and Karolinska cohorts, corresponding to 93.5% and 92.2%, respectively. The out-of-sample classification accuracy levels in this study were lower than what was observed previously by Biggs et al. (2019b) in thirty patients and thirty NPs (100% accuracy), however, the authors had not included pelvis kinematics and joint powers, and this may explain the differences from the current investigation. When including joint powers, Metcalfe et al. (2017) found that the Classifier accuracy was 95%, similar to the current report. It may be inferred that introducing joint powers lowers the Classifier classification accuracy, however, this hypothesis cannot be confirmed by the results in the current study and more research would be needed on this aspect, comparing the Classifier accuracy variation when introducing certain gait features.

The most prominent finding was that in two groups of patients from different countries over half (55.6%) of the eighteen most discriminative gait features separating patients and controls were similar. This suggests that the presence of severe knee OA produced characteristic gait alterations that are consistently

identified by the classifier. Some differences in the eighteen most discriminating gait features were found between cohorts. This could indicate that different groups of patients pre-TKR may have unique, and slightly different gait characteristics from one another. It must be noted that the Karolinska patients were compared to NPs of a similar age, while the Cardiff patients were compared to a significantly younger group of NPs. In the Cardiff cohort, this choice was made because older age is a risk factor for the development of knee OA (Blagojevic et al. 2010), and to ensure that the gait classification was going to be based on a “true” healthy gait, following previous work (Watling 2013; Biggs 2016), and a mixed-age group of NPs was included in the Cardiff cohort. Hence, it was unclear if the abovementioned differences were due to the Cardiff cohort’s reference group being significantly younger than the patients, as opposed to the age-matched controls in the Karolinska group, or if the differences were dependent on diverse gait biomechanics characterising the Cardiff and Karolinska participants. The results from the current study do not allow verification of these suggestions, but chapter 5 addresses the issue of utilising a much younger reference group versus using a NP group of similar age to the patients undergoing TKR.

As correctly emphasised previously by Metcalfe (2014), the order in which the highest ranking eighteen gait features appeared within the classifier does not play a role in the way participants are classified, as there is no weighing of the gait features depending on their rank. Therefore, the fact that similar gait features were in a different order between the two Classifiers was not fundamental in determining the subjects’ Belief values but was an indication of how accurately that feature classified the participants as NP or OA.

In the Cardiff cohort, most gait features had a classification accuracy > 80%, while in the Karolinska cohort, this was true only for four gait features, including knee flexion-extension angles PC2, hip power PC2, knee flexion-extension moments PC2 and ankle inversion-eversion angles PC1. This finding may be a result of Karolinska controls’ older age. It is widely agreed that biomechanical changes occur in healthy older peoples’ gait, such as a slower walking speed (Herssens et al. 2018), reduced ankle plantar flexion and power generation (Pol et al. 2021a), reduced knee extension and hip flexion at heel-strike, and smaller ground reaction forces (Boyer et al. 2017b). Therefore, it is likely that several gait features were “more comparable”

between the Karolinska patients and controls than between the Cardiff patients and NPs, resulting in a lower accuracy of the Classifier in being able to differentiate the severe OA from NP gait. In fact, some of the age-related gait changes are similar to those observed in people with knee OA, but Duffell et al. (2017) highlighted that certain kinematic and kinetic features at the hip-knee-ankle are not age-related but characteristic of suffering from OA. The following discussion will highlight the similarities observed in the two groups of patients pre-surgery.

Joints' kinematics features similarities between cohorts

When looking at joint kinematics, PCA revealed a reduced hip and knee flexion-extension ROM, and a decreased hip adduction-abduction ROM in both groups of patients. Hip kinematics alterations in the presence of severe knee OA have been reported in previous studies that compared patients to age-matched or younger NPs and were similar to what was found in the current work, including smaller hip flexion-extension ROM (Astephen et al. 2008a; Ro et al. 2018), smaller hip adduction angles at stance (Duffell et al. 2017), larger peak hip abduction during swing (Baert et al. 2013). The reduced hip flexion-extension ROM may be associated with the limited knee ROM in stance that was observed in the patients and that may be likely due to pain or knee flexion deformity, usually observed in people with knee OA; knee flexion-extension PC2, describing reduced knee flexion-extension ROM was highly discriminatory in the Karolinska group and this gait alteration was found in previous reports (Ouellet and Moffet 2002; Levinger et al. 2013; Ro et al. 2018) and the Cardiff patients alike. A lack of knee extension in terminal stance would limit consequently, the hip extension. This deduction is supported by the findings of Ro et al. (2019) who observed that hip ROM reduction was significantly correlated with knee ROM limitations in a group of eighty-nine patients with severe knee OA and this finding did not change after controlling for gait speed. Conversely, a previous study looking at older adults (Chehab et al. 2017), found correlations between walking speed, BMI and hip flexion angles during gait, suggesting that the pattern observed in this study may also be related to the patients' reduced walking speed and higher BMI.

Joints' moments and powers features similarities between cohorts

In terms of joint kinetics, in both patient groups, PCA depicted the presence of smaller knee flexion moments in early stance, as described in previous research (Astephen et al. 2008a; Milner 2009), and smaller extension moments in terminal stance, also observed in earlier reports (Levinger et al. 2013; Ro et al. 2018). This could be linked to the reduced knee ROM throughout stance, which would reduce the knee moment arm relative to the GRF vector, resulting in smaller knee flexion and extension moments.

In line with previous findings, both groups of patients displayed larger dorsiflexion moments at initial stance and smaller peak dorsiflexion moments at terminal stance (Ouellet and Moffet 2002; Biggs et al. 2019b; Ro et al. 2019) as depicted by ankle dorsi-plantar flexion moments PC1 (Cardiff) and PC2 (Karolinska). This finding may justify the reduced patients' gait speed, as an earlier study found that walking velocity in patients post-TKR was influenced by the amplitude of the ankle dorsiflexion peak moment and the knee ROM (Ro et al. 2018).

In accordance with Levinger et al. (2013), both Cardiff and Karolinska patients showed a reduced peak knee power generation and absorption throughout the gait cycle (knee power PC1) compared to reference groups and this characteristic was highly discriminatory between people with severe OA and NPs (within the first seven highest-ranking features in both groups). Joint power indicates when negative work (eccentric, negative sign of the waveform) or positive work (concentric, positive sign of the waveform) is performed around a joint; to aid interpretation of this variable, looking at the correspondent joint moment can explain if the concentric work was created by an external flexor or extensor moment (Robertson et al. 2013). Resende et al. (2012a) found reduced power absorption in late stance and power generation during swing but not in other aspects of the knee power waveform. However, the study (Resende et al. 2012a) only included women with mild knee OA and this may justify the modest alterations in their biomechanics. The reduced knee power profiles may have been due to the patients' reduced walking speed. In support of this proposition, McGibbon and Krebs (2002) found that the knee mechanical energy expenditure was similar between people with knee OA and controls when walking at the same speed.

The Cardiff and Karolinska patients in the current study manifested an increased ankle power absorption at the start of mid-stance, compared to NPs, as depicted by ankle power PC3 and PC2, respectively. These PCs described the eccentric work of the calf muscles guiding the movement of the tibia over the talus while the centre of mass moves forward (Houglum et al. 2012). When plotting the patients pre-TKR and NPs ankle power mean waveforms during a gait cycle, there were no discernible differences between these groups upon visual inspection, either in the Cardiff or Karolinska cohorts. This may be due to having averaged the data and eliminated the variability within it, while PCA was able to identify the pattern distinguishing patients form NPs. To the best of the author's knowledge, the patient's increased ankle power eccentric work at the start of mid-stance currently observed is a novel finding that has never been described in previous research. Patients with knee OA tend to walk with a stiff knee, meaning that they have reduced flexion-extension ROM during stance, with the tendency to maintain a flexed knee (Bytyqi et al. 2014; Toda et al. 2021). An increased knee flexion before mid-stance would cause the tibia to be in a more disadvantaged position (i.e., more forwardly tilted in the sagittal plane, relative to the foot) and would result in an increased ankle dorsiflexion [which was indeed depicted by Ankle dorsi-plantar flexion angle PC1 at this timepoint in the Cardiff cohort (Table **15**)]. The increased ankle eccentric work observed may be a strategy of the calf muscles to support the excessive tilt forward of the tibia, in an attempt to stabilise the ankle-knee before full weight acceptance. Yet, this hypothesis cannot be confirmed with the current results as the muscles' activity was not examined in this study. In the Karolinska patients, the ankle power PC2 also depicted a delayed and reduced peak power generation around 55% of the gait cycle (terminal stance), while in the Cardiff's the main aspect described by ankle power PC3 was a reduced concentric work at this stage. The peak ankle power generation at terminal stance is consistent with the calf muscles contracting to plantar-flex the foot (heel off) and allow the lower limb to be pushed forward in preparation for the swing phase (Houglum et al. 2012). Previous several reports identified a reduced ankle power generation at push-off in elderly people and the difference was present also when young and elderly adults walked at the same speeds (DeVita and Hortobagyi 2000; Silder et al. 2008). Similar findings were described in people with severe knee OA (Resende et al. 2012a), even at walking speeds matched with controls (McGibbon

and Krebs 2002), showing that the reduced ankle concentric work was not due to the reduced gait speed usually observed in people with knee OA.

GRF features similarities between cohorts

In both patients' groups, the GRF PCs identified similar patterns, for the vertical GRF there was a reduced loading rate, loss of the by-phasic nature of the waveform and a higher force in midstance. Alterations in the vertical GRF were described in previous studies, highlighting reduced vertical force at the first peak (Yoshida et al. 2012) and during push-off (Wiik et al. 2017), reduced vertical GRF peaks and a monophasic curve with higher GRF force in mid-stance (Costello et al. 2021). Cardiff and Karolinska patients equally displayed a reduced mediolateral GRF loading rate with a larger medial force around midstance, a finding in line with recent work including a large cohort of people with radiographic and symptomatic OA enrolled in the Multicenter Osteoarthritis Study (Costello et al. 2021). Nevertheless, mediolateral GRF was not as discriminatory between patients and NPs in either group, compared to other features (accuracy < 80%). In the above-mentioned studies (Wiik et al. 2017; Costello et al. 2021), the GRF differences were irrespective of patients and controls walking at the same velocity, showing that these changes are not simply gait speed-related, but specific to the presence of joint pain. Vertical and mediolateral GRF play an important role in determining the knee adduction moment, together with the knee lever arm (Hunt et al. 2006). The reduced vertical and mediolateral GRF loading rate may be an attempt to slow down the loading on the knee to limit the pain, however, contributing to an increased knee adduction moment in mid-stance (as depicted by knee adduction-abduction moment PC2 in both cohorts).

Both Cardiff and Karolinska patients displayed a reduced anterior and posterior GRF, but previous work did not support this finding (Costello et al. 2021) and observed that the minimal differences between people with knee OA and controls disappeared when adjusting for gait speed. However, the study (Costello et al. 2021) did not specify the knee OA severity and patients walked with their footwear, a factor known to modify the loading of the lower limb (Shakoor and Block 2006).

Aligning with a previous report (Astephen et al. 2008b), in the current study, most of the gait features differentiating NPs from patients pre-surgery were related to gait kinetics in the Cardiff group and a similar trend was noted in the Karolinska cohort, although the proportion of gait kinetic variables was lower than Cardiff's. This suggests that at this stage of knee OA, kinetic changes may be more relevant in differentiating patients from NPs.

Hip kinetics and kinematics similarities between cohorts

Assessment of the eighteen highest-ranking features revealed that in both cohorts, about 30% of gait features distinguishing between patients and NPs were hip-related, indicating that several compensatory strategies occur at this level. Excluding the similarities discussed above, most of the hip features depicted slightly different patterns between the two groups of patients. Cardiff patients had smaller hip flexion and extension moments (hip flexion-extension moment PC2) and a larger adduction moment in mid-stance (hip adduction-abduction moment PC2), whereas Karolinska patients showed a different pattern, with reduced hip peak adduction moments compared to NPs (hip adduction-abduction moment PC1). Hip kinetics alterations in the presence of severe knee OA have been reported in previous studies and were similar to what found in the current work, including smaller first and second peak hip adduction moments and larger hip abduction moment at heel strike (Mündermann et al. 2005) and reduced hip extension moments (Astephen et al. 2008a).

Interestingly, in both groups, gait features regarding hip joint powers in the sagittal plane were highly discriminatory between patients and NPs (accuracy ≥ 90%). Consistent with the current investigation, a high discriminatory value of hip and knee power using the Classifier was found by Metcalfe et al. (2017) in fifteen patients pre-TKR, where these gait features ranked first (92.5% accuracy) and second (87.5% accuracy), respectively, in terms of accuracy in classifying participants. Nevertheless, the authors (Metcalfe et al. 2017) did not specify which PC score was retained by the Classifier or what was the interpretation of the PCs.

Hip power PC2 was the second most discriminatory feature between patients and NPs in both the Karolinska and Cardiff groups, despite representing only 20% of the original data variance. The PC described slightly different aspects in the

Karolinska and Cardiff cohorts and reflected patients' hip power phase shifts and reduced power generation around toe-off (only in the Cardiff cohort), and midstance (Karolinska cohort only). In the Cardiff cohort, reduced power generation around toe-off was linked to the fact that patients had a reduced external hip extension moment, as depicted by hip flexion-extension moment PC2. This would have resulted in reduced concentric work of hip flexors, whose function at this stage is to flex the hip to clear the lower limb from the ground and accelerate the lower limb forward for the swing phase (Houglum et al. 2012). The reduced power generation at the start of mid-stance observed in the Karolinska patients may be linked to the smaller hip flexion moments at this stage, but not depicted as one of the main gait features by the Classifier. Hip extensors work concentrically to extend the hip and stabilise the trunk in this phase (Houglum et al. 2012). Reduced joint concentric and eccentric work could be linked to hip muscles' weakness which has been described previously in people with severe knee OA. Hinman et al. (2010) examined eighty-nine people with knee OA and found that compared to twenty-three controls, there was diminished strength in all hip muscle groups. Rightly, the authors (Hinman et al. 2010) highlighted that it was not possible to imply whether the muscle weakness was a consequence or if it was present before the occurrence of knee OA. It must be acknowledged that is not known if the participants of the present study had muscle weakness as muscle strength testing was not performed. Muscles strength testing and passive joints ROM assessment could be of value to interpret the results from 3D gait analysis, in order to understand if a reduced mobility and altered joint kinetics are due to muscles weakness or joint stiffness.

The current findings were also similar to Resende et al. (2012b), who explored significant differences in joint powers PC scores between patients with knee OA and controls; the study found that a statistical difference existed only in hip power PC2 scores, but not in PC1 and PC3 between groups, despite representing only 18.8% of the variance; the PC depicted patients' reduced power absorption (eccentric work) in late stance and power generation (concentric work) at toe-off (Resende et al. 2012b). The slightly different meaning of hip power PC2 from the current study could be because the report only included women that suffered from mild knee OA (Resende et al. 2012b). Similarly, previous research observed reduced maximum hip power generation at mid-stance and swing in forty-three patients with

severe OA compared to twenty-nine age-matched controls, nevertheless, the differences were not significant (Levinger et al. 2013). However, in the study (Levinger et al. 2013), gait velocity was adjusted for in the statistical analysis and this may justify the difference from the current investigation where gait speed was not corrected between patients and controls. It is well established that certain gait variables are gait-velocity dependent (Astephen Wilson 2012). Yet, the choice of not adjusting for gait speed in the current report was driven by the fact that while a reduced gait velocity is not specific to the presence of severe knee OA, it is a typical characteristic of the disease that tends to decrease with increased OA severity (Astephen et al. 2008a). Using walking speed as a covariate violates one of the assumptions of the statistical test, which is that the covariate (walking speed) must be independent of the other variables (joint biomechanics) (Portney and Watkins 2014).

The results of the current work demonstrated that most of the highest-ranking features discriminating severe knee OA from NPs found in previous research (Biggs et al. 2019b) and in the current study were not specific to the Cardiff cohort since they were present in a cohort of patients and NPs from a different country too. The results of the study will be able to inform future research utilising the classifier, as it has demonstrated which are the recurring, and therefore, key gait features that should be expected within the eighteen highest ranking ones distinguishing between severe OA and NPs.

4.4.5 Patients' gait biomechanics changes post-TKR in the two groups

Hypothesis 5: several gait features improve pre- to post-TKR but patients do not show a gait pattern equivalent to that of NPs at one-year follow-up in either group

In agreement with the initial hypotheses, the majority of the gait features improved pre to post-TKR in both groups of patients, showing that TKR surgery is successful at facilitating the recovery of gait biomechanics. The patterns of recovery were different between the two cohorts, as in the Cardiff patients, most of the improvements involved gait kinetics, while in the Karolinska patients, it entailed a slightly higher proportion of gait kinematics. This is not surprising, considering that the gait features that were utilised by the Classifier to evaluate gait changes pre to

post-TKR were mainly kinetic-related in the Cardiff group and kinematic-related in the Karolinska.

In line with a previous study (Biggs et al. 2019b), the gait improvement was mainly associated with joints other than the knee. The largest proportion of features which showed a trend towards recovery was seen in hip biomechanics in both patient groups, showing that there was a decrease in gait compensations at this joint. However, a return to normal could be seen only for the Karolinska patients, whose hip adduction-abduction ROM (hip adduction-abduction angles PC2) was comparable to NPs. This finding was in line with what was observed by Ro et al. (2020), who found that hip adduction angles throughout the gait were comparable to age-matched controls in eighty-four patients two years after TKR. It could be suggested that a return to normal of this gait feature could be expected only when patients are compared to a control group of a similar age. However, the current study does not allow to confirm this assumption as a group of patients should be compared to several groups of controls of varying ages to validate this point. In the Karolinska patients, the mediolateral GRF PC3 (loading rate and force magnitude at 50% stance) was comparable to reference subjects, nevertheless, the change pre to post-TKR was small, implying that this feature was not much different from NPs before the surgery. Moreover, mediolateral GRF was not as discriminative of patients and NPs pre-TKR in this group.

Overall, the patients' gait improvement was suboptimal one-year post-TKR, considering that no gait features were comparable to NPs in the Cardiff patients and only two gait features were equivalent to the controls in the Karolinska cohort. Some gait features returned to normal in the Karolinska patients, and it may be because they had been compared to an age-matched reference group. In accordance with the current study, it has been demonstrated that one year following TKR surgery, gait biomechanics do not go back to normal and several significant joint biomechanical alterations subsist, either when looking at peak values or gait features via PCA, including smaller peak vertical GRF (Yoshida et al. 2008), reduced knee ROM (McGinnis et al. 2013), reduced peak hip extension (Saari et al. 2005b), smaller knee peak extension angle in late stance, increased ankle dorsiflexion (Levinger et al. 2013), decreased knee extension angles and moments (Yoshida et al. 2008; Ro et

al. 2018) and hip flexion (Ro et al. 2018) and extension moments (Yoshida et al. 2008).

The Cardiff Classifier has been previously utilised in patients pre- to post-joint surgery and age-matched volunteers and a larger proportion of patients were classed as having a healthy gait after the operation when the NPs had a similar age to the patients. Worsley et al. (2015) looked at gait changes in a mixed group of thirty-one patients following TKR and UKR via the Cardiff Classifier; the reference group had a similar age (62.4 ± 5.9) to the patients (60.9 ± 67.2) and it was found that post-surgery, 20% of TKR patients were classed as NP. Metcalfe et al. (2017) observed twenty patients pre- and one-year post-TKR, compared to twenty age-matched controls and observed that 53.3% (8 out of 15) were classified as NP after the surgery (two of which had been classified as NP pre-surgery). When running the Cardiff Classifier on thirty patients pre (69.7 ± 8.6) and one-year post-TKR utilising a significantly younger cohort as reference (thirty NP, age 39.8 ± 17.6) Biggs et al. (2019b) found that one year after the operation only 10% of patients moved to the NP side of the Cardiff Classifier. Therefore, it was surprising to observe that a similar proportion of patients were classed as having an NP gait in the Cardiff group (25%) compared to the Karolinska (22.2%). Previous studies demonstrated that women tend to go back to normal gait patterns to a larger degree than men after TKR (McClelland et al. 2018; Paterson et al. 2020). Considering that the Karolinska group had a higher proportion of women compared to the Cardiff patients, it may have been expected that a larger number of patients would show a return to a normal gait in the Karolinska cohort. It is not clear whether the patients included in this study suffered from OA in other joints of the lower limbs which may have played a role in the gait recovery and future studies including this information may help interpret the results.

One of the strengths of this study is that it was conducted on two patient groups from different countries and therefore the results were not specific to a single cohort. However, there were some limitations. It is common practice to compare people with severe OA to age-matched controls, this ensures a more realistic evaluation, especially when considering comparing gait biomechanics post-TKR: a return to a completely normal gait is not expected after the surgery as the knee prostheses do

not perform the same as a native knee. Comparing patients to much younger people, like in the Cardiff cohort, may produce unrealistic expectations of gait recovery. Nevertheless, the choice of including younger controls was an attempt to remove one of the risk factors for knee OA, as previously discussed. In the Cardiff cohort, there was a prevalence of males, which is unusual, considering that knee OA is more prevalent in women (NICE 2022) and therefore the results of the Classifier regarding the Cardiff cohort should be interpreted accounting for this limitation. A further limitation is that NPs were not assessed radiographically for the absence of knee OA, however, any potential healthy volunteer was excluded in the presence of known risk factors, such as previous knee injuries or if they presented symptoms of a joint issue. The study was underpowered, and it included a relatively small number of patients, the latter is a common issue in studies utilising 3D motion analysis, considering the amount of time needed to collect data.

4.5 CONCLUSION

Adding to the limited existing evidence, this study showed that in two groups of patients from different countries, the classifier had a large responsiveness in measuring the gait biomechanics changes pre to one-year post-TKR. A novel finding was that the classifier was more responsive than the GDI and GDI-kinetic. It is suggested that the classifier may be a more appropriate measure to assess the objective gait function change in research evaluating TKR outcomes and that combining lower limb kinematics and kinetics in a single measure may produce larger responsiveness than assessing these separately as in the GDI and GDI-kinetics.

The classifier was compared to the GDI and GDI-kinetic absolute scores pre- and post-TKR. Pre-TKR, the significant correlation between both BOA and GDI, BOA and GDI-kinetic, suggested that both joint kinetics and kinematics were affected to a similar degree in both patient groups and an agreement between these outcome measures. This indicates that either the classifier or the GDI combined with the GDI-kinetic may be used interchangeably to assess gait function pre-TKR in future research. However, the classifier and GDI did not correlate post-TKR and the classifier and GDI-kinetic were not associated in one of the groups of patients, suggesting a disagreement between these measures. The lack of correlation

between the BOA and GDI change in both groups and the mixed findings in the relationship between BOA and GDI-kinetic change reinforced the idea that these measures did not agree. The findings may indicate that patients had a mixed recovery, where a change in gait kinematics was not complemented with a similar variation in gait kinetics or vice versa. Therefore, it is not recommended to utilise the classifier or the GDI – GDI-kinetic interchangeably to assess gait function and its change post-TKR. The fact that GDI and GDI-kinetic do not allow to explore what were the gait features determining the indexes, may result in favouring the use of the classifier to assess gait function, due to its potential to inform what could be targeted by rehabilitation to improve the outcomes.

A new contribution to the existing knowledge was that the majority of the gait features separating OA patients from their references were similar between the Cardiff and Karolinska groups, suggesting that severe knee OA produced distinctive gait alterations that are consistently identified by the classifier. Future studies employing the classifier should expect to observe these key gait features within the eighteen highest-ranking variables distinguishing between severe OA and NPs.

This study showed that the gait biomechanics, although improving one-year post-TKR, overall was not comparable to that of NPs in either group of patients, adding to the existing literature describing the recovery of gait function post-TKR.

Chapter 5: Relationship between preoperative, surgical and patients' factors and the gait function one-year post-TKR

5.1 INTRODUCTION

Preliminary work was conducted before developing the following chapter. This stemmed from the results in Chapter 4, where two separate groups of patients undergoing TKR were examined against NPs of similar age (i.e., in the Karolinska cohort) and significantly younger NPs (i.e., in the Cardiff cohort) and some differences in the eighteen most discriminating gait features were found between cohorts. Nevertheless, it was unclear if these differences were due to the Cardiff cohort's reference group being significantly younger than the patients, as opposed to the age-matched controls in the Karolinska group, or if the differences were dependent on diverse gait biomechanics characterising the Cardiff and Karolinska participants. Karolinska patients also had a lower BOA, both pre- and post-TKR, with respect to Cardiff patients, but a similar change in BOA pre to one-year post-surgery was observed between patient groups. However, no statistical analysis was carried out on these aspects as this was not one of the aims of the study. Nevertheless, the results from the study in Chapter 4 suggested that there might be two potential risks of errors if comparing patients' biomechanics to that of a much younger control group:

- Systematically underestimating the pre- and post-operative patients' objective gait function.
- The gait features that most accurately differentiate patients pre-TKR from NP may be the ones also indicating differences between older and younger people's gait rather than underpinning the main features differentiating strictly OA gait from a non-pathological gait.

These points should be investigated to determine potential sources of imprecision when comparing patients to a mixed-age group of NPs and on average, significantly younger NPs, and to highlight potential limitations when employing a mixed-age NPs group within the Cardiff classifier. This information would be valuable for the current study and future investigations employing and aiming to implement the Cardiff classifier.

In the course of the TKR study, there have been challenges in recruiting NPs who have a similar age to the patients undergoing TKR, as older people often suffer from joint pain and have had previous joint injuries and surgeries, both exclusion criteria of the study. Our research group at Cardiff University has often included NP participants with a wide array of ages when utilizing the Cardiff classifier (Watling 2013; Biggs 2016; Biggs et al. 2019a; Biggs et al. 2019b), the justification for this was that the main aim of utilising the Cardiff classifier was to evaluate the patients' change in biomechanics with respect to a "true" healthy gait, given that individuals with a high BMI and older age are at higher risk of having OA while potentially being symptoms-free, and that gait modifications are present in people with the abovementioned characteristics. Only two studies utilising the Cardiff classifier included NPs older than fifty (Worsley et al. 2015) or sixty years of age (Metcalfe et al. 2017). In these cases, data were collected at different Academic Institutions and looked at knee biomechanics of TKR and UKR pre- and six-months post-surgery (Worsley et al. 2015), and at hip, knee and ankle biomechanics pre- to one-year post-TKR in a smaller cohort of twenty patients (Metcalfe et al. 2017). Watling (2013) previously investigated differences in discrete biomechanics among young NPs (18-34 years old), mid-aged NPs (35-54 years old) and older NPs (55 years and older) to determine what was the best combination of NPs to utilise within the Cardiff classifier. The analysis was based on finding differences in discrete parameters of hip, knee and ankle joints and not on principal components or the Cardiff classifier (Watling 2013). The results highlighted that older NPs had mainly kinetics modifications compared to the younger NP groups, with significantly smaller ankle dorsi-plantarflexion ROM, peak hip flexion and adduction moments, knee flexion, extension and adduction moments, ankle plantarflexion and adduction moments (Watling 2013). Nevertheless, the author (Watling 2013) decided that the young and middle-aged group of NPs were appropriate to use as a control within the Cardiff classifier, as the main aim of his study was to be able to classify the patients against a healthy gait. Following this work, Biggs (2016) argued that only including young and mid-age NP subjects to classify patients, could result in some issues. In fact, the classifier would have most likely picked gait differences between NP and OA patients related to ageing too and not strictly related to the presence of knee OA. In his work, Biggs (2016) explored the potential correlation between the belief values of NP and

their age and found that the BOA and BNP of the NPs were not significantly associated with age, justifying the inclusion of elderly NPs within the classifier. However, it was not explored how the classification of patients may have varied when including only young and mid-aged subjects, mixed-age or only elderly individuals. No study, to date, has clarified whether utilising a group of NPs of mixed age or a group of NPs with a similar age to the patients undergoing TKR would produce different classification outcomes with the Cardiff classifier. Therefore, the preliminary work in this chapter included developing two Cardiff classifiers, and comparing the outputs when the patients from the Cardiff cohort were assessed against those NPs who were fifty years of age or older (from now onwards, defined as the "NP50 classifier"), and when patients were compared to a mixed-age group of NPs (from now onwards, defined as the "mixed-age classifier").

Aim 1: To explore whether there were differences in patients' classification between the NP50 classifier and the mixed-age classifier.
Hypothesis 1.1 – The patients' Belief of OA is lower using the NP50 classifier, as opposed to utilising the mixed-age classifier, both pre- and post-TKR.
Hypothesis 1.2 – Based on observations in the previous chapter, the change in Belief of OA for patients is similar when using the NP50 classifier or the mixed-age classifier.
Hypothesis 1.3 – The highest-ranking gait features discriminating patients from older NPs in the NP50 classifier, are similar to those separating patients from NPs in the mixed-age classifier (since the mixed-age classifier already includes the older NPs utilised in the NP50 classifier).

It was decided that if significant differences were found in at least one of the abovementioned hypotheses testing, the classifier trained on patients and the older NP group would be utilised in the subsequent work within the current chapter. This is because comparing patients pre- and post-TKR to an age-matched NP group may lead to a more realistic gait classification. In other words, patients' biomechanics would be compared to the biomechanics they would be expected to have at their present age, had they not had knee OA, rather than being compared to biomechanics expected when they were younger.

The Chapter 4 examined how gait function can be evaluated with diverse summary measures, showing that the Cardiff classifier is more responsive than the GDI and GDI-kinetic. The Cardiff classifier therefore provides an accurate method to classify patients pre- and post-TKR, and thus can be utilised to aid understanding of TKR outcome from an objective function perspective. As previously discussed within the Literature Review, there are several ways to measure TKR outcome, surgeons may be interested in revision surgery rates and implant survival, and clinicians usually assess the pain relief and function achieved following the surgery, for instance, via PROMs, performance-based tests (PBTs), joint ROM and muscles' strength (Vajapey et al. 2020).

There has been an ongoing interest in defining factors that may correlate with TKR outcomes. In the Literature Review, it was shown that age (Kramers-de Quervain et al. 2012; Bade et al. 2014), sex (Kennedy et al. 2006; Worsley 2011; Mandzuk et al. 2015; Pua et al. 2019), BMI (Stickles et al. 2001; Xu et al. 2018), having pain in several joints (Kramers-de Quervain et al. 2012), number of comorbidities (Kramers-de Quervain et al. 2012; Riddle et al. 2022), the function pre-TKR (Fortin et al. 1999; Fortin et al. 2002; Kennedy et al. 2006; Gandhi et al. 2009b; Watling 2013; Bade et al. 2014; Worsley et al. 2015; Pua et al. 2019), implant design (Watling 2013; Verra et al. 2015), receiving a patella resurfacing (Longo et al. 2018), may play a role in determining the outcome, measured mainly with subjective measures (Fortin et al. 1999; Fortin et al. 2002; Kennedy et al. 2006; Bade et al. 2014; Mandzuk et al. 2015; Xu et al. 2018; Riddle et al. 2022), post-TKR.

All the above preoperative factors may play a role in TKR outcome, however, the research on predicting objective function one-year post-TKR is very limited. Three earlier investigations within our research group explored the predictors of gait function post-TKR evaluated via the Cardiff classifier (Worsley 2011; Watling 2013; Biggs 2016). However, the statistical significance of the findings was not reported (Watling 2013), and participant numbers were small (n = 12 in Watling (2013); n = 22 or less, depending on the variable investigated in Biggs (2016) pre and 9+ months post-TKR), or when similar to the current investigation (n = 33), the study (Worsley 2011) included a cohort of mixed patients receiving either TKR or UKR, pre- to six-months post-knee replacement. The abovementioned reports looked at potential factors associated with absolute gait function post-TKR (Watling 2013) or explored

factors related to the change in gait function after knee arthroplasty (Worsley 2011; Biggs 2016). Nevertheless, only BMI, function pre-TKR, PROMs and implant design were investigated as potential predictors of gait function one-year post-TKR (Watling 2013; Biggs 2016).

Since PROMs and PBTs are poorly correlated (Stratford et al. 2010; Mizner et al. 2011; Onodera et al. 2020), and measures of gait quality do not always associate with PBTs (Naili et al. 2017a; Naili et al. 2017b), the predictors of objective gait function based on lower limb biomechanics may be different to those looking at predicting subjective or time-based objective outcomes. The aims of identifying such factors are to explore and target modifiable characteristics to optimise the outcome or to be aware of the non-modifiable factors affecting the outcome to facilitate making the best-informed decision regarding surgery. This information would also allow for discussion of the factors affecting the TKR outcome with the patient for a shared decision-making process, an approach recommended by the NICE guidelines (NICE 2022).

The current study aimed to explore predictors for both absolute gait function (BOA post-TKR) and its change (BOA change) one-year post-TKR. The purpose was to expand on the previous research by utilising a larger cohort of patients and exploring preoperative factors that had not been taken into consideration in previous investigations utilising the Cardiff classifier (Worsley 2011; Watling 2013; Biggs 2016) but potentially being relevant in influencing outcomes. The change in BOA can be interpreted as “getting better” or “improving” in the gait function domain, as opposed to “having a good or poor function”, depicted by the BOA post-TKR. The BOA change and BOA post-TKR describe slightly different aspects of the recovery of function post-TKR, and it was deemed of relevance to explore what factors could be predictive of both aspects of function after total knee arthroplasty.

Considering that the number of patients in this study was relatively small for prediction analyses, the investigation aimed to be exploratory. Nonetheless, the findings would be valuable in suggesting the preoperative factors that could be modified, or that influence the gait, or its improvement, and could offer a basis for further research. A previous study found that the Cardiff classifier classification of function correlated with the OKS (Biggs et al. 2019a), suggesting that patients' objective function has a relationship with the self-perceived level of function both at

baseline and one-year post-TKR. Given the relationship with the Cardiff classifier, the question arose of whether the baseline OKS may also be a predictor of the objective outcome post-surgery measured via the Cardiff classifier and this was investigated in the current study. This information would be relevant to clinicians since English clinical commissioning groups have previously suggested OKS thresholds pre-surgery to determine the suitability for TKR (Dakin et al. 2012).

Due to the limited literature looking at predictive factors of objective function post-TKR, it was difficult to formulate hypotheses on which preoperative factors were expected to be associated with gait function post-TKR or to choose the factors to be included in a regression model uniquely based on theory or existing knowledge. Due to the lack of strong evidence on predictors of gait functions, it was decided to determine the best predictors statistically. The following research questions were developed:

- *Which patient and clinical preoperative and surgical factors are associated with the change in BOA or BOA post-TKR?*
- *If associations exist between patient and clinical preoperative and surgical factors and BOA change or BOA post-TKR, could they be utilised to develop prediction models of BOA change and BOA post-TKR?*

Aim 2: Prediction of objective function post-TKR: to explore which factors could be predictive of the TKR outcome, where the outcome is the BOA change and the absolute BOA value one-year post-TKR.

Hypothesis 2.1 – A negative, weak relationship exists between BOA pre-TKR and BOA change pre- to post-TKR (i.e., the BOA pre-TKR can predict the BOA change).

This means that patients with worse gait function pre-TKR would have the largest improvement in gait (i.e., the greatest change in BOA) due to having more scope for functional gain. The strength of the relationship was hypothesised to be weak based on previous investigations using the Cardiff classifier, which have shown high heterogeneity in the change scores pre to post-TKR (Biggs 2016).

Hypothesis 2.2 – There is a strong, positive association between the BOA pre-TKR and the BOA one-year post-TKR (the BOA pre-TKR can predict the BOA post-TKR).

This means that the patients with worse preoperative function (i.e., a high BOA value) would have poor function post-TKR (i.e., a high BOA value). The

hypothesis was developed as such since the most recurrent finding of the abovementioned literature (Fortin et al. 1999; Fortin et al. 2002; Kennedy et al. 2006; Gandhi et al. 2009b; Watling 2013; Bade et al. 2014; Worsley et al. 2015; Pua et al. 2019), was that function pre- and post-TKR were associated.

The Cardiff classifier is a valuable tool to assess function in people with knee OA and the outcome post-TKR in a research setting, however, in its current form, its application in a clinical setting would be challenging, due to the need to carry out 3D gait motion analysis, with the associated expertise to process and analyse the data and outputs. During the TKR study, knee ROM was assessed as part of the data collection, and this is a variable routinely measured in the clinic by both surgeons and physiotherapists, pre- and post-TKR. Recently, knee ROM was identified as a key outcome measure to be reported in future TKR studies, together with overall function and pain levels (Singh et al. 2017; Capin et al. 2022). It was deemed to be informative for clinicians to know whether an objective measure of knee mobility that can be easily collected in a clinical setting, would be associated with the overall gait function measured with the Cardiff classifier, in the absence of other measures of function.

Aim 3: To explore the association between patient knee ROM and BOA.
Hypothesis 3.1 – Pre-surgery, knee ROM is associated with BOA.
Hypothesis 3.2 – Post-surgery, knee ROM is associated with BOA.
Hypothesis 3.3 – From pre- to post-surgery, the change in knee ROM is associated with the change in BOA.
Hypothesis 3.4 – Knee ROM pre-TKR is associated with the BOA post-TKR.

A regression analysis to predict the BOA post-TKR using knee ROM as a predictor was not conducted however, due to the limited number of patients with knee ROM data available.

5.2 METHODS

This was a retrospective cohort study on prospectively collected data. Additional data were collected from the available medical records for this study, for the patients who had explicitly consented to this.

5.2.1 Participants

Patients for this study were the same analysed in Chapter 4 in the Cardiff cohort, and therefore, the reader is directed to section 3.3.1 for the method of recruitment, inclusion and exclusion criteria. Thirty-nine patients who had gait data available for both the pre- and one-year post-TKR follow-up were included in the present study (same sample as Chapter 4). Five of these patients received a bilateral TKR (two patients had a simultaneous joint replacement), for a total of forty-four lower limbs.

A convenience sample of thirty-four mixed-age NPs were included, this was the same sample adopted in the study in Chapter 4 from the Cardiff cohort, collected between February 2007 and December 2021. NP data were extracted for both lower limbs, except for four participants who did not have data for the contralateral limb due to technical issues, for a total of sixty-four lower limbs.

For the purpose of this study, an additional group of NPs of similar age to the patient group was developed. More precisely, a sub-sample of fourteen NPs who were fifty years of age or older (NP50) was extracted from the mixed-age NP group mentioned above. Furthermore, three NPs were added to the NP50 group from an NP set of data collected between January and October 2022, based on their age (age ≥ 50 years) (the whole NP50 group included data collected between February 2007 and October 2022). A total of seventeen NP50 were identified and retained for further analysis. NP50 data were extracted for both lower limbs, except for two participants who did not have data for one lower limb due to technical issues, for a total of thirty-two lower limbs. The choice of including only NP aged fifty or older was to create a control group with an age that was more comparable to that of the patients included in this study. Table 16 shows the sample breakdown and the overlap between the participants utilised in Chapter 4 and the current work.

Table 16 Participants utilised in the current study and in Chapter 4, and details on the classifiers developed in the abovementioned chapters

Classifiers developed	Chapter where the classifier was developed	Chapter where the classifier output was utilised for analyses	Participants entered into the classifier and used for further analyses	
			Non-pathological subjects (n)	Patients pre-TKR (n)
Mixed-age classifier	Chapter 4	Chapter 4, current chapter	34 (mixed-age)	39
NP50 classifier	Current chapter	Current chapter	14* + 3 newly introduced (≥ 50 years of age)	39*

n: number; NP50 classifier: classifier including non-pathological volunteers ≥ 50 years of age; *: overlap with the sample utilised in Chapter 4.

5.2.2 Clinical data collection, patient-reported outcomes

The presence of contralateral knee pain, pain in other joints of the lower limb, and the presence of previous surgeries in lower limb joints were recorded preoperatively and at the post-TKR follow-ups during the 3D gait analysis sessions. For patients whose data were collected before January 2011 (six cases), the presence of pain in other lower limb joints was not collected in a standardised way, but subjects were asked to report on it, where nothing was found in the session notes, it was assumed that no pain was present in other joints.

For the patients included in this study who consented to access their medical records for research purposes, clinical data were collected by the author or by the research officers based at CAVOC between approximately May 2020 and June 2021 from the data existing within the patients' medical notes at CAVOC. Data were recorded systematically on a form developed for the current project (Appendix C) and included immediate complications post-surgery (i.e., deep vein thrombosis, pulmonary embolism, infection), the consultant's TKR surgery notes (surgical procedure, intra-operative events), knee implant components' labels (implant components manufacturer and posterior cruciate-retaining/sacrificing design, patella resurfacing), and comorbidities. Comorbidities were recorded via the American Society of Anaesthesiologists (ASA) score, which aids in evaluating the risk related to surgeries (Mayhew et al. 2019), and it was calculated by the anaesthesiologist at the patient's pre-surgical appointment. ASA 1 identifies a healthy person, ASA 2 is a

person with mild disease, ASA 3 is an individual with a debilitating disease, while ASA 4 is a subject with major diseases that are a risk to life (ASA 2020).

Patients completed the KOOS, described earlier in section 3.5.1, which is a subjective outcome measure evaluating five knee OA outcomes, including pain, symptoms, activities of daily living (ADL), sports and recreational activities, and knee-related quality of life (QoL) (Roos et al. 1998). Each KOOS subscale normalised score ranges from 0, meaning extreme knee issues, to 100, no knee issues. Patients also completed the OKS (section 3.5.2), evaluating pain and function with twelve questions, each offering four answers (0 = extreme issues; 4 = no issues), with a summative maximum score of 48 (i.e. no knee-related issues) (Fitzpatrick et al. 1998).

5.2.3 Three-dimensional motion analysis protocol

According to the protocol, patients' gait was assessed pre-TKR, and at the 3, 6 and 12 months post-surgery. Additionally, during the assessment, patients' knee ROM was evaluated via motion capture technology. Only pre and one-year post-TKR data were considered for the current study. 3D gait data collection and processing adopted for this study were previously described in Chapter 3, while the assessment of knee ROM is described followingly. The 23 gait variables extracted from the gait analysis were the same analysed in Chapter 4, and included pelvis and hip angles in all three planes, knee, and ankle kinematics in the sagittal and frontal planes; hip, knee, ankle sagittal and frontal planes distal moments (expressed as external moments); hip, knee, ankle sagittal distal joint powers; vertical, anterior-posterior, mediolateral ground reaction force (GRF); foot progression angles.

Knee active-assisted ROM was measured with the 3D motion analysis system (Qualisys, Sweden) with the patient sitting on a plinth in the centre of the laboratory, starting with their feet on the floor and their knees at approximately 90° of flexion; the examiner sat on the floor in front of the patient, with one hand under the foot and the other on top of the knee to avoid interfering with markers visibility, and guided the knee full extension and flexion to the maximum achievable ROM for three consecutive times, with the participant being instructed to actively engage in the movement. Two trials were recorded, for a total of six flexion-extension cycles.

The data collection was video recorded, and the knee ROM videos were visually inspected by the author to determine the quality of the assessment. For instance, an assessment of insufficient quality was deemed to be one where the examiner guided the knee continuously from extension to flexion at high speed, without stopping at the end of flexion or extension, potentially limiting the full knee ROM achievable by the patient. Other criteria to define the knee ROM assessment as “poor quality” included when the knee flexion stopped because the participant’s foot touched the floor limiting the knee flexion, or the hand of the examiner was kept in the knee popliteal fossa, potentially restraining full knee flexion. When a knee ROM assessment was deemed to be of poor quality for the above-mentioned reasons, videos from gait, stair negotiation and sit-to-stand were visually inspected for comparison to the knee ROM assessment full knee extension and flexion; when the videos of the other tasks clearly showed that a larger knee flexion or extension could be achieved by the patient, the poor quality of the knee ROM assessment was confirmed and data were not analysed further.

Knee ROM trials were labelled in Qualisys Track Manager (Qualisys, Sweden), exported in a “.c3d” file and then processed in Visual 3D (version 6, C-Motion Inc., Maryland, USA) with a custom-written pipeline. Marker's trajectories were filtered with a Butterworth low-pass filter (cut-off frequency 6 Hz). A biomechanical model of the participant was created based on the standing static calibration (similar to what was done for the walking trials discussed in section 3.6). Knee joint angles in the sagittal plane were calculated with an XYZ sequence (where X was the flexion-extension axis) (Cardan-Euler angles), with the proximal segment being fixed and the distal segment moving. Maximum, minimum knee flexion and knee ROM (the sum of maximum knee extension and maximum knee flexion) were calculated for each repetition and the average value among the three repetitions was exported in an ASCII file for the two trials, then transformed in a Microsoft Excel sheet (Microsoft, Redmond, WA, US) for extracting the relevant information. Data from the two trials were averaged for the operative side.

5.2.4 Participants’ classification of gait function

The outputs of the classifier developed for the Cardiff cohort in Chapter 4 (i.e., the BOA, BNP and Uncertainty level for 39 patients pre-TKR, 39 patients one year

post-TKR, and 34 mixed-age NPs), identified as the 'mixed-age classifier' from now onwards, were utilised in the current study, as shown in Table 16, to achieve Aim 1.

The work in this Chapter consisted in producing another classifier based on the patients' data pre-TKR deriving from Chapter 4 (i.e., 39 patients, 44 lower limbs) and the NP50 group created for the current study (n = 17 subjects, 32 lower limbs) (NP50 classifier). The patient's classification of gait function one-year post-surgery was also assessed in the NP50 classifier. The reader is redirected to section 3.7 for details on the Cardiff classifier method. Briefly, PCA was performed on patients and NP50 on 23 gait variables extracted from gait analysis (see section 5.2.3) for each participant. The first three PCs of each variable, and for each participant, were entered into the NP50 classifier. It is worth noting that the abovementioned 23 variables were the same ones chosen to perform PCA in Chapter 4, to then produce the gait features entered into the mixed-age classifier. The eighteen highest-ranking features, in terms of accuracy in differentiating between patients pre-TKR and NP50 during training of the classifier, were utilised to classify the participants in the NP50 classifier and to perform the patients' classification of gait function one-year post-surgery. The out-of-sample classification simplex plot was produced, and the out-of-sample classification accuracy was calculated as the proportion of out-of-sample subjects, out of the total, whose classification matched their class. The sensitivity (Eq. 27) and specificity (Eq. 28) of the classification based on the full dataset (i.e., 32 lower limbs from the NP50 group and 44 lower limbs from patients pre-TKR) was reported.

5.2.5 Statistical analyses

All analyses were performed in IBM SPSS Statistics (version 27, IBM, Armonk, NY, US). Normality distribution was determined via the Shapiro Wilk's test (p-value > 0.05: parametric distribution) and inspection of Q-Q plots. For all statistical tests, the significance level was $\alpha = 0.05$. To compare baseline characteristics between patients and NP50, continuous variables were examined with an independent t-test or its non-parametric equivalent (Mann-Whitney U test) and categorical variables (i.e., sex) with Pearson's Chi-Square test.

Aim 1 – Preliminary work

The patients' Belief values (BOA, BNP and Uncertainty) pre- and post-TKR, were calculated for the NP50 classifier, the Belief values for the mixed-age classifier were extracted from Chapter 4. For each classifier:

- The change scores of each Belief value pre to post-TKR were calculated by subtracting the score obtained at baseline from the one calculated at one-year follow-up for each patient
- Then, the mean score of the change was calculated for the BOA, BNP, and Uncertainty. A negative BOA change and a positive BNP change denoted gait improvement. A positive Uncertainty change denoted an increase in the Cardiff classifier classification uncertainty level

To verify if differences in the patients' BOA, BNP and Uncertainty values pre and post-TKR and their change were significant between the mixed-age classifier and the NP50 classifier, a paired samples t-test or its non-parametric equivalent, the Wilcoxon signed-rank test, was run. The choice of considering the patients in the two conditions as paired was made because these were the same patients, with the same baseline biomechanics, but assessed against two different conditions (mixed-age NPs and NPs ≥ 50); this could be compared to a scenario where the same group of patients underwent two treatment conditions at different timepoints, testing differences between the two conditions in this case would require a paired statistical test. Additionally, Bland-Altman plots were produced for the BOA pre-TKR, post-TKR and the BOA change (main outcomes) to highlight trends between the mixed-age and NP50 classifiers and potential biases in the classification.

Similarity and differences in the eighteen highest-ranking variables classifying patients in the mixed-age classifier and NP50 classifier were examined.

Aim 2 – Prediction of objective function post-TKR

Following the results from the preliminary analysis, the prediction analysis was conducted utilising the NP50 classifier only. Multiple linear regression was adopted as it allows to explore the linear relationship between a continuous variable *y* to be predicted (outcome or dependent variable), and *x* predictors (or independent variables), and is represented by Eq. 30:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \quad \text{Eq. 30}$$

Where β_0 corresponds to the intercept, and β_1 is the regression coefficient for x_1 (or the slope of that regression line). A regression coefficient represents the value of the predicted outcome (e.g., β_1) with respect to the reference category (e.g., male, x_1) when dealing with categorical data; for continuous predictors, it represents the value of the outcome when the continuous variable (e.g., BMI, x_2) increases by one unit. The significance of the relationship between outcome and predictors is indicated by the p-value. The adjusted R^2 (ranging from 0, a poor fit of the model, to 1, a perfect fit of the model) indicates to which degree the independent variables are associated with the dependent variable (and represents the effect size), and the amount of the outcome's variance described by the independent variables in the general population (as opposed to R^2, representing the degree to which the predictors predict the outcome in the actual group studied). An adjusted $R^2 < 0.1$ was interpreted as a poor fit, $0.11 <$ adjusted $R^2 < 0.3$ as an acceptable fit, $0.31 <$ adjusted $R^2 < 0.5$ as a moderate fit, $R^2 > 0.5$ as a strong fit (Muijs 2011).

The outcome variables (dependent variables) were the gait function change (= BOA change) pre to one-year post-TKR and the gait function post-TKR (= BOA post-TKR). Potential outcome predictors (independent variables) were chosen based on factors shown to affect function found in previous literature and discussed in the Introduction and included age at surgery, sex (male or female), BMI, previous major surgeries on other joints other than the knees pre-TKR (yes/no), the presence of contralateral knee issues at the baseline assessment (i.e. previous joint replacement, high tibial osteotomy, presence of knee OA or pain) (yes/no), pain at the hip and ankle joint/s at baseline (yes/no), gait function pre-TKR (BOA pre-TKR), perceived knee-related symptoms, self-perceived function, and pain at baseline (KOOS sub-scores), ASA grade at the time of surgery, implant design (posterior cruciate sacrificing or retaining), patella resurfacing (yes/no).

A preliminary correlation analysis was conducted to determine the relationship between the predictors and each of the outcome variables. It was decided to include in the correlation analysis all the cases, including those with incomplete data for certain variables because the statistical test may have revealed that the variables for which some of the patients did not have data, may have not been relevant predictors. However, the findings must be treated with caution as having had a larger

number of subjects for the variables with missing data, may have produced different results. The correlation analysis was set to exclude just the data for the variable in question that was not available (exclude pairwise, i.e., ASA score, implant design, patella resurfacing, post-surgery complications, and KOOS sub-scales). These cases were removed for the multivariable regression analyses if some of the variables for which they did not have data, were shown to be relevant for the prediction model.

Pearson's correlation was utilised for normally distributed, and for dichotomous data (i.e., sex), and its non-parametric equivalent (Spearman's correlation) was used for ordinal (i.e., ASA score, KOOS sub scores, OKS, etc.) or non-normally distributed data. Where a significant correlation between predictors and outcome was found, a multivariable regression model was developed utilising those predictors, therefore, adopting a data-driven approach, as anticipated in section 5.1. This approach was employed by previous studies (Brown et al. 2009a; Sims et al. 2009; Hamilton et al. 2012b; Twiggs et al. 2019) and ensures that only relevant predictors are entered into the model. Moreover, introducing more independent variables in a predictive model entails a larger sample size (Knofczynski and Mundfrom 2008; Algina and Olejnik 2010; Burmeister and Aitken 2012), which was not attainable due to the study design, and the choice of including only the relevant variables was made to increase the power of the analysis. The strength of the correlation, where significant, was interpreted as suggested by Dancey & Reidy (2011), where r is the correlation coefficient (indicating both r_s, Pearson's correlation coefficient, and ρ: Spearman's correlation coefficient):

- $|r| < 0.3$ meant a weak correlation
- $0.4 < |r| < 0.6$ was regarded as a moderate correlation
- $|r| > 0.7$ was interpreted as a strong correlation

Numerous predictors may have correlated significantly with the outcome, and it was decided to develop an initial regression model including all predictors that correlated significantly with the outcome in the correlation analysis. The nature of this initial model was exploratory, to reveal the predictors that significantly correlated with the outcome. However, the results should be treated with caution, considering the low number of patients versus the number of predictors.

There is no clear consensus on the most appropriate number of participants to be utilised in a linear multivariable regression model. Some researchers have recommended considering the correlation coefficient target as a reference to determine the number of participants, rather than the number of predictors alone. According to Algina & Olejnik (2010) for a 95% probability that the sample's regression coefficient R^2 will be 0.55 ± 0.20 (i.e., between a moderate to strong fit), the most appropriate sample size should be 43 subjects. However, Knofczynski & Mundfrom (2008) suggested that for a good prediction model, an adjusted R^2 (i.e., the population regression coefficient) of 0.50 (moderate fit), and 4 predictors, the sample size should be 55 subjects. Historically, and as a rule of thumb, the predictors-to-participants ratio should be at least 1:10 (Miller and Kunce 1973), and this was the approach utilised in the current investigation, given its exploratory nature. Only the main predictors (those which significantly correlated with the outcome in the initial regression model or whose significance level was the closest to p = 0.05) were entered into the final model, to align with the 1:10 predictors to participants ratio. All predictors were inputted simultaneously in each regression model.

Multivariable regression requires the data to satisfy the following assumptions, which were tested adhering to the guidance of (Muijs 2011) and (Nayebi 2020):

- The presence of a linear relationship between each continuous, statistically significant predictor and each of the outcomes was visually investigated with scatterplots
- Independence of residuals (= independence of the observed values between different patients and lower limbs, for those patients who received a TKR bilaterally) was tested with the Durbin-Watson statistics and the acceptable range was 0 to 4
- The absence of correlation between predictors (= multicollinearity) was assessed via Pearson's correlation ($r_s > 0.7$ was an indication of collinearity), and additionally, tolerance for each of the independent variables was examined for the presence of any values < 0.1, indicating collinearity. Tolerance measures the degree of variance in a predictor that is not explained by the other independent variables; values approaching zero indicate that the

variance of the predictor in question is almost fully explained by the remaining predictors, therefore, indicating multicollinearity

- The presence of homoscedasticity, in other words, the variance of the residuals (the difference between the real values and the predicted ones, i.e., the prediction error) being evenly scattered around the zero line, was tested by plotting the unstandardized predicted values and the studentised residuals and observing the presence of patterns in the data
- The absence of outliers (studentised deleted residual with a value > ±3)
- Approximately normally distributed residuals, determined with the visual inspection of the histogram displaying the distribution of the residuals with a superimposed bell curve and inspection of the P-P plot

For the multivariable regression model, age, BMI, KOOS sub-scores, OKS, and BOA pre-TKR were coded as continuous variables. ASA score is a categorical variable that cannot be entered into the regression model in its original form due to presenting more than two categories (Nayebi 2020). Dummy variables were created for the ASA scores. ASA scores 3 and 4 were combined into ASA ≥ 3 to allow an adequate number of subjects in this group as only one patient in the cohort had an ASA of 4; this approach was adopted in previous studies similar to the current investigation and looking at TKR functional outcome prediction (Jiang et al. 2017; Sanchez-Santos et al. 2018). ASA ≥ 3 was utilised as a reference and therefore, was not included in the model. Each ASA score (1 and 2) was, therefore, dichotomised (yes/no).

Aim 3 – Association between knee ROM and objective function

A correlation analysis was conducted for the patients who had knee ROM data available. The following correlations were explored:

- Knee ROM and BOA at baseline
- Knee ROM at baseline and BOA post-TKR
- Knee ROM and BOA one-year post-TKR
- Knee ROM change and BOA change one-year post-surgery

Pearson's correlation was utilised for normally distributed data, and its non-parametric equivalent (Spearman's correlation) was used for non-normally

distributed data. The strength of the correlation was interpreted as suggested by Dancey & Reidy (2011) and as previously mentioned.

5.3 RESULTS

5.3.1 Preliminary work

The results for the mixed-age classifier were already reported in Chapter 4 and therefore, only the relevant aspects for comparing the findings between the two classifiers will be reported.

Characteristics of NPs utilised in the NP50 and mixed-age classifiers

Table 17 shows that a larger proportion of women was found in the NP group utilised in the mixed-age classifier, where most participants (58.8%) were in the 19 to 49 years old range, and 41.2% of subjects in this group were 50 years old or older. Subjects in the mixed-age classifier had a normal weight, while NP50 were, on a mean in the pre-obesity range (WHO 2010). NP50 were only 3.2% slower than the mixed-age NPs, but on a median, were heavier than the mixed-age NPs.

Table 17 Comparison of the NP participants' characteristics utilised in the two Cardiff classifiers

	Mixed-age classifier non-pathological volunteers (n = 34)	NP50 classifier non-pathological volunteers (n = 17)
Age (years), mean (SD)	42.9 (19.2)	64 (9)
range	*19 - 79*	*51 - 79*
19-29, n (%)	14 (41.2)	0
30-39, n (%)	3 (8.8)	0
40-49, n (%)	3 (8.8)	0
50-59, n (%)	6 (17.8)	6 (35.3)
60-69, n (%)	4 (11.7)	5 (29.4)
70-79, n (%)	4 (11.7)	6 (35.3)
Height (m), mean (SD)	1.70 (0.10)	1.68 (0.09)
Weight (kg), mean (SD)	69.3 (14.3)	73.0 (14.7)
BMI (kg/m^2), mean (SD)	24.4 (3.7)	25.9 (3.9)
range	*17.7 – 31.5*	*20.3 - 32.5*
Females, n (%)	22 (64.7)	9 (52.9)
Walking speed (m/s), mean (SD)	1.23 (0.12)	1.19 (0.08)

NP: non-pathological volunteer; m: meters; kg: kilograms; BMI: Body Mass Index; n: number; *: statistically significant at $p < 0.05$

Characteristics of participants utilised in the NP50 classifier

The proportion of females in the NP50 group was lower with respect to the patients, who were slightly higher and older than NP50 but all these differences were not significant (Table 18). Nevertheless, the age distribution was different between groups since NP50 had not been age-matched (due to data availability). While most patients (45.5%) were between 70-79 years old at the time of surgery, there was an even distribution of NP50 in the age group 50-59 and 70-79 age ranges (both 35.3%). Moreover, patients were significantly heavier and with a higher BMI than the NP50.

Table 18 Characteristics of the participants within the NP50 classifier

Participants characteristics	Non-pathological volunteers ≥ 50 years of age (n = 17)	Patients pre-TKR (n = 44)	Significance level
Age (years), median (IQR), *range*	67 (16) *51 - 79*	69 (8) *41 - 77*	0.221‡
40-49, n (%)	0	1 (2.3)	
50-59, n (%)	6 (35.3)	5 (11.4)	
60-69, n (%)	5 (29.4)	18 (40.9)	
70-79, n (%)	6 (35.3)	20 (45.5)	
Height (m), mean (SD)	1.68 (0.09)	1.69 (0.09)	0.565
Weight (kg), median (IQR)	73.0 (24.7)	94.0 (30.2)	< 0.001‡*
BMI (kg/m^2), median (IQR), *range*	24.9 (6.8) *20.3 - 32.5*	31.4 (10.0) *21.3 - 52.4*	< 0.001‡*
Females, n (% within group)	9 (52.9)	14 (31.8)	0.127
Walking speed (m/s), mean (SD)	1.19 (0.08)	0.82 (0.20)	< 0.001*

TKR: Total Knee Replacement; IQR: interquartile range; SD: standard deviation; m: meters; kg: kilograms; BMI: Body Mass Index; n: number; ‡: non-parametric test;
*: statistically significant at $p < 0.05$.

Compared to NP50, patients walked significantly slower, both pre-TKR (mean walking speed 31.1% slower than NP50) and post-TKR (mean walking speed 17.6% slower than NP50: 0.98 ± 0.16 m/s; $p < 0.001$), however, their walking speed increased significantly pre to one year after the surgery ($p < 0.001$, non-parametric test).

Comparison between the NP50 and mixed-age classifiers

Figure 22 shows that the mixed-age classifier had a 93.5% out-of-sample classification accuracy in classifying Cardiff participants pre-TKR as OA and NP (as shown in Chapter 4), and the NP50 classifier out-of-sample classification accuracy was similar, although slightly lower (90.8%).

In the mixed-age classifier, the sensitivity was 93%, and the specificity 98%, as seen in Chapter 4. In the final NP50 classifier classification based on the whole dataset, the sensitivity was 91%, and the specificity was 94%. Two NPs were classed as having an OA gait. It must be noted that no knee radiographs were performed on the NP50 subjects, and therefore, it was unknown whether the false positives may potentially have had radiographic signs of knee OA. One NP50 (ID: “P4” – right lower limb) was closer to the boundary between BOA and BNP. Additionally, the results of the NP50 classifier highlighted the presence of altered biomechanics on the left side for NP50 ID 02812 (not previously included within the mixed-age classifier), and this is consistent with this participant reporting recently having some potential knee issues, which were deemed to be negligible by the author, due to their sporadic nature and duration (the participant reported “an occasional clicky knee on the left side when walking downhill, causing a twinge of pain that only lasted for that moment”, which was not present at the moment of recruitment, just a few months before the data collection).

When considering the patients within the NP50 classifier, four were classified as having an NP gait (ID 02575-left side, 02624, 02644, 02766), and two of these (ID 02575-left side, 02766) had been previously misclassified within the mixed-age classifier too. When observing the raw gait biomechanics outputs, these patients had overall “normal” joint angles, moments, powers, and GRF (within 1SD from the NP group), compared to NP, justifying their classification as non-dominant NP gait. The only gait abnormalities in these patients included an increased hip flexion throughout gait (ID 02575-left side), the absence of knee extension moments in the second half of stance (ID 02575-left side, 02624, 02644), increased hip abduction (ID 02766), delayed knee flexion during swing (ID 02624, 02644, 02766), increased ankle dorsiflexion (ID 02766). When visually comparing the simplex plots from the two classifiers, Figure 22b shows that a larger number of patients were shifted towards

or were within the OA non-dominant area, however, patients were slightly closer to the uncertainty vertex of the plot.

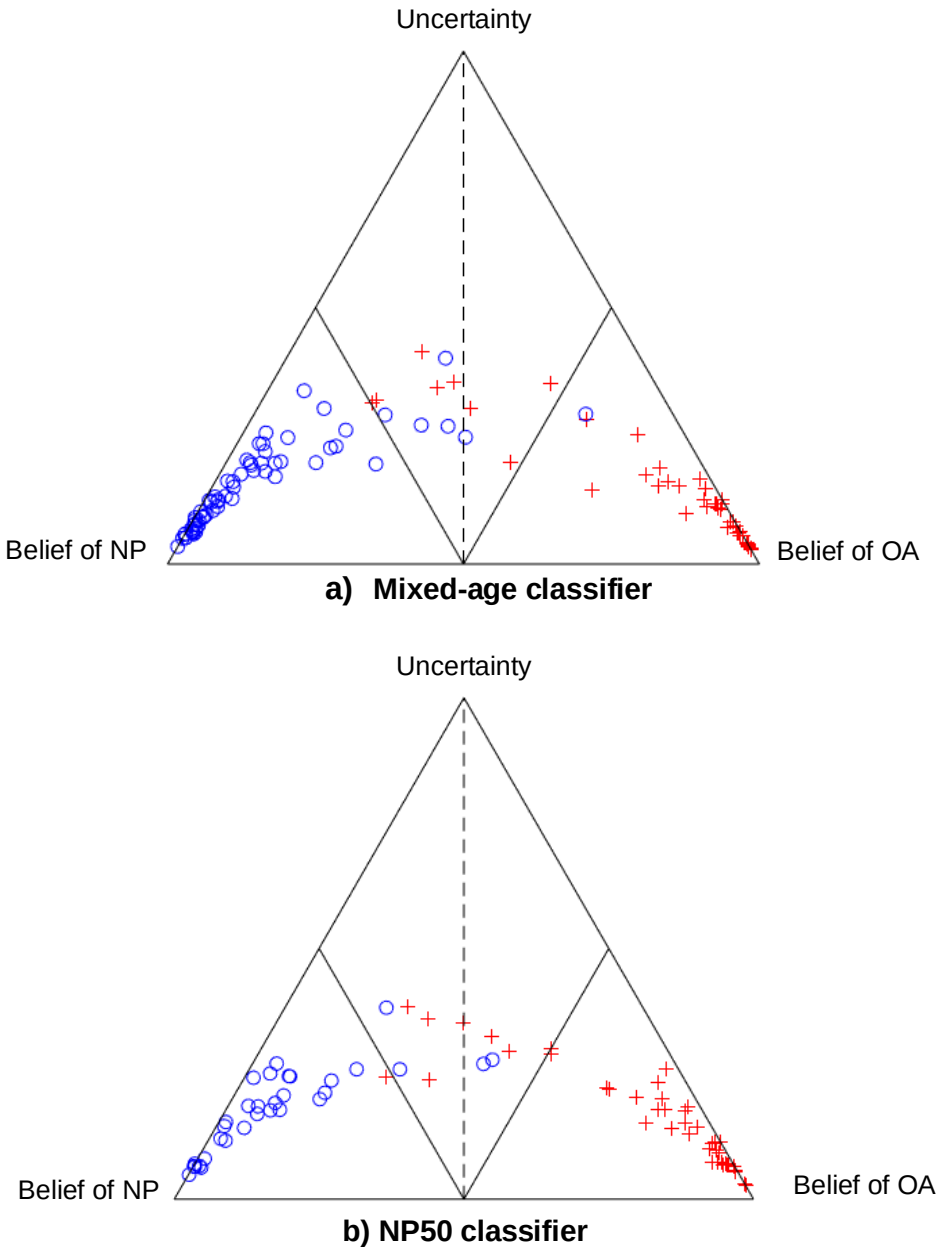


Figure 22 Simplex plots representing the out-of-sample classification accuracy. **a)** Mixed Age classifier simplex plot; **b)** NP50 classifier simplex plot. The blue circles represent the non-pathological volunteers, the red crosses the patients pre-surgery. The black, dashed line indicates the boundary between the non-pathological and osteoarthritic area of the plot. The solid lines within the plot, indicate a Belief of osteoarthritis (BOA) and Belief of non-pathological (BNP) of 0.50 (the boundary between dominant BOA and dominant BNP).

Hypothesis 1.1 – The patients' Belief of OA is lower using the NP50 classifier, as opposed to utilising the mixed-age classifier, both pre- and post-TKR

Table 19 reports the patients' Belief values pre- and post-TKR, and their change over one-year post-TKR. The results showed that patients compared to NP50 had a significantly better gait function (= lower BOA value) both pre (mean BOA 0.76 ± 0.20) and post-TKR (mean BOA 0.55 ± 0.25) compared to the same patients assessed against a mixed-age NP group, confirming the hypothesis. Patients compared to NP50 had a significantly higher BNP (= better gait function) than those compared with a mixed-age NP group, but post-TKR only. The uncertainty level in the patients' classification was significantly higher in the NP50 classifier at both time points, as was seen in Figure 22b pre-TKR.

Table 19 Comparison of the gait function classification between the mixed-age classifier and the NP50 classifier pre, post-TKR and the change in classification pre to 12 months post-TKR

	Patients compared to NP ≥ 50 years (n = 44) Mean (SD)	Patients compared to mixed-age NP (n = 44) Mean (SD)	Significance level
Pre-TKR			
BOA	0.83 (0.22)	0.87 (0.21)	0.007‡*
BNP	0.02 (0.09)	0.02 (0.10)	0.205‡
Uncertainty	0.14 (0.15)	0.12 (0.12)	0.001‡*
Post-TKR			
BOA	0.55 (0.25)	0.58 (0.27)	0.010*
BNP	0.20 (0.33)	0.12 (0.39)	0.041‡*
Uncertainty	0.22 (0.08)	0.21 (0.09)	0.041*
Change scores			
BOA	-0.20 (0.23)	-0.17 (0.25)	0.162‡
BNP	0.14 (0.18)	0.10 (0.17)	0.071‡
Uncertainty	0.07 (0.12)	0.06 (0.14)	0.515‡

NP: non-pathological volunteer; n: number; SD: standard deviation; TKR: Total Knee Replacement; BOA: Belief of osteoarthritis; BNP: Belief of non-pathological; ‡: non-parametric test, results reported as median (interquartile range);
*: statistically significant at $p < 0.05$.

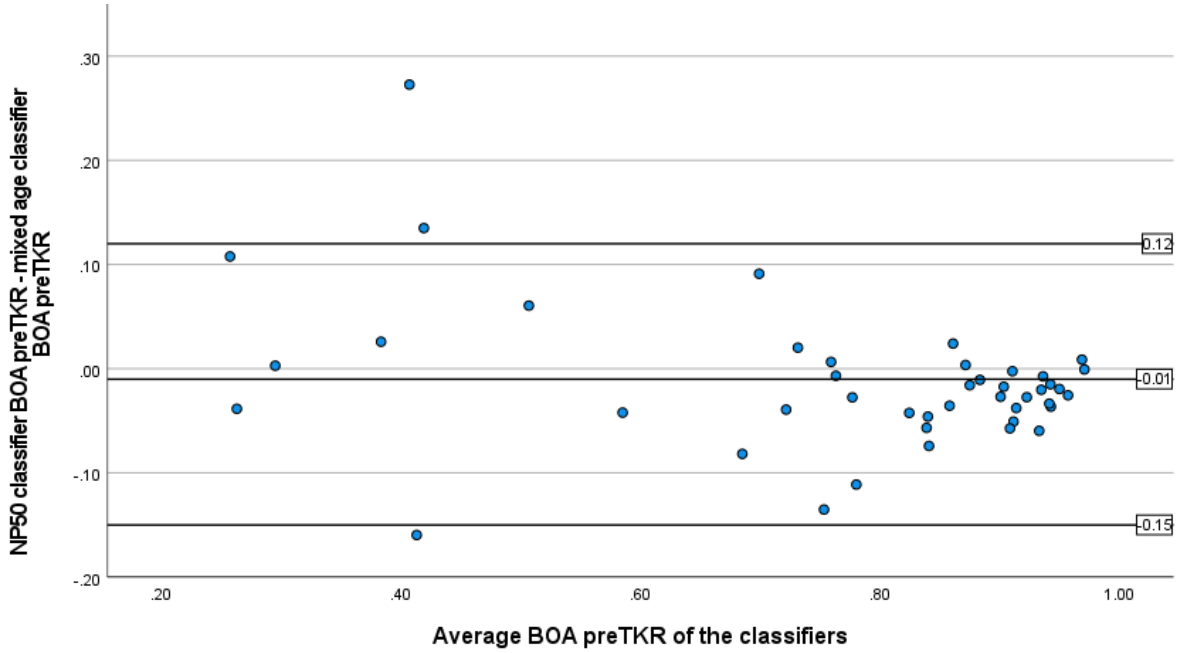


Figure 24 Bland-Altman plot of agreement between the mixed-age and NP50 classifier (BOA pre-TKR). NP50 classifier: classifier where the non-pathological (NP) group was aged 50 or older; BOA: Belief of osteoarthritis; TKR: total knee replacement.

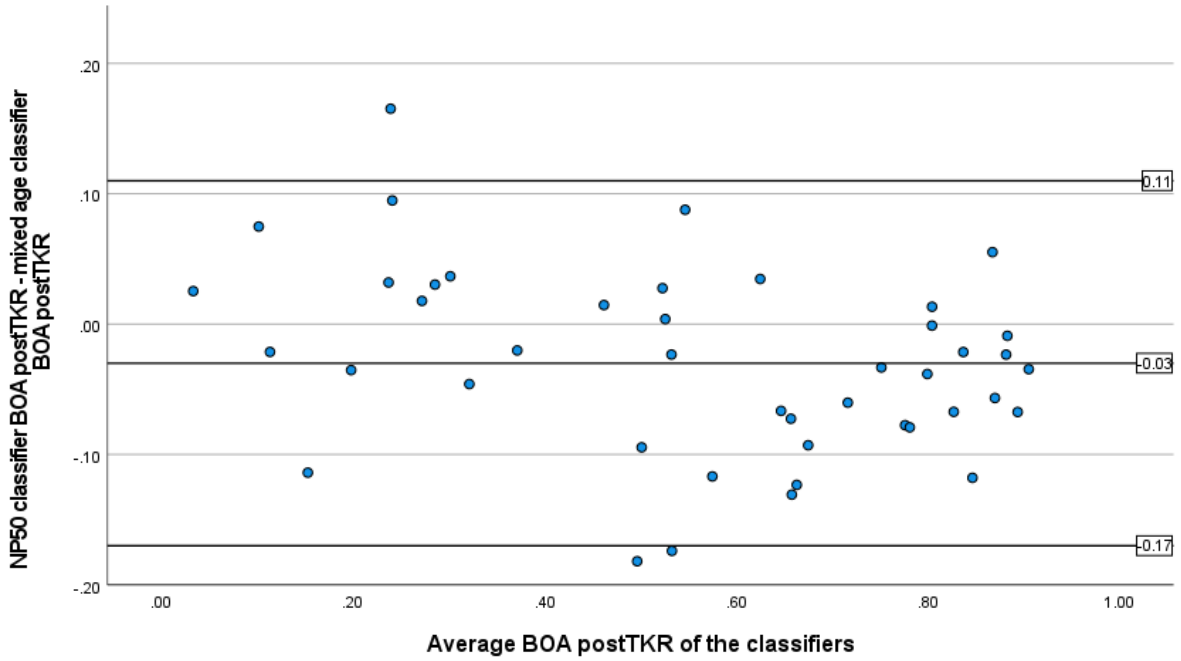


Figure 23 Bland-Altman plot of agreement between the mixed-age and NP50 classifier (BOA post-TKR). NP50 classifier: classifier where the non-pathological (NP) group was aged 50 or older; BOA: Belief of osteoarthritis; TKR: total knee replacement.

The plot in Figure 24 shows that on average, the NP50 classifier measured -0.01 BOA points less than the mixed-age classifier. There was a trend between data, in fact, the variability around the mean was not constant and became smaller at average BOA values above 0.80, indicating better agreement between the classifiers for participants with a more compromised gait function. Most of the points were within the limits of agreement (0.12; -0.15).

The plot in Figure 23 shows that on average, the NP50 classifier measured -0.03 BOA post-TKR points less than the mixed-age classifier. There was a slight trend between data, in fact, the variability around the mean was not constant and became slightly smaller at average BOA values above 0.70, indicating better agreement between the classifiers for participants with a more compromised gait function post-TKR. Additionally, there seemed to be a trend where the NP50 overestimated the BOA value as the estimated BOA approached 0.40 and lower values. Most of the points were within the limits of agreement (0.11; -0.17).

Hypothesis 1.2 – The change in Belief of OA for patients is similar when using the NP50 classifier or the mixed-age classifier

Even though the patients compared to the NP50 had a larger improvement in gait function (mean BOA change: -0.21 ± 0.18) (Table 19), the change in all three belief values was not statistically different from the patients assessed against a mixed-age NP group. This agreed with the hypothesis.

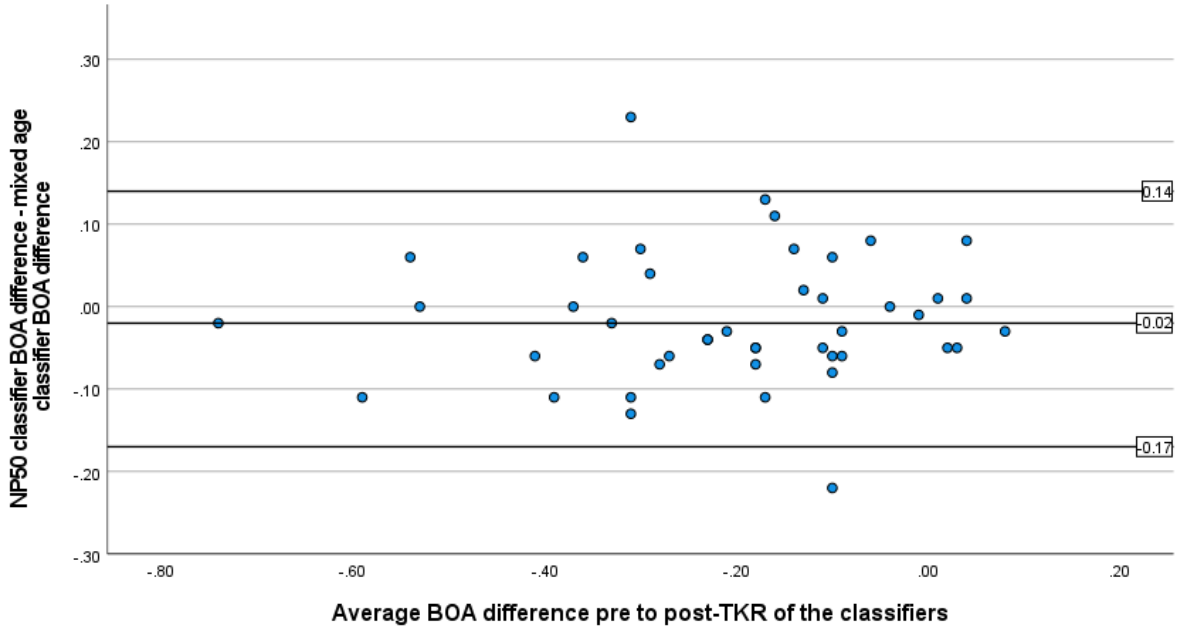


Figure 25 Bland-Altman plot of agreement between the mixed-age and NP50 classifier (BOA change pre to post-TKR). NP50 classifier: classifier where the non-pathological (NP) group was aged 50 or older; BOA: Belief of osteoarthritis; TKR: total knee replacement.

The plot in Figure 25 shows that on average, the NP50 classifier measured -0.02 BOA change points less than the mixed-age classifier. There was a slight trend between data, where the NP50 overestimated the BOA change value as the estimated BOA change approached 0.00 and higher values. Most of the points were within the limits of agreement (0.14; -0.17).

Hypothesis 1.3 – The highest-ranking gait features discriminating patients from older NPs in the NP50 classifier, are similar to those separating patients from NPs in the mixed-age classifier

Table 20 shows the eighteen highest-ranking PCs discriminating gait between patients pre-TKR and NP50. 77.8% (= 14 PCs) of the most discriminative gait features coincided or were similar between the two classifiers, confirming the hypothesis. The ankle dorsi-plantar flexion moments feature between the two classifiers described the same gait characteristic but were depicted by PC1 in the mixed-age classifier, and PC2 in the NP50 classifier, describing a smaller plantarflexion moment around 15-20% stance and a smaller peak dorsiflexion moment around 80-85% stance. Two gait features were depicted by the same PC but had slightly different meanings for the two classifiers:

- Hip power PC2
- Ankle power PC3

61.1% of the retained gait features (11 of the total 18) identified the same PCs and the same gait features in both cohorts, including:

1. Pelvis kinematics in the frontal plane PC2 (patients elevation of the pelvis towards the stance limb in mid-stance)
2. Hip flexion-extension angles PC2 (patients smaller hip flexion-extension ROM)
3. Hip flexion-extension moments PC2 (patients reduced flexion moment during stance and reduced extension moment in final stance)
4. Hip adduction-abduction angles PC2 (patients reduced hip adduction-abduction ROM)
5. Hip adduction-abduction moment PC2 (patients larger adduction moment in mid-stance and slower loading rate)
6. Knee flexion-extension angles PC2 (patients smaller knee flexion-extension ROM)
7. Knee flexion-extension moments PC2 (patients reduced knee flexion moment around 15-20% stance and reduced extension moment around 65-70% stance)
8. Knee power PC1 (patients reduced knee power generation and absorption throughout the gait cycle)
9. Vertical GRF PC1 (patients delayed peak to about 40% and anticipated at 60% of stance - reduced loading rate, higher magnitude around 50% stance: loss of by-phasic nature of the waveform)
10. Anterior-posterior GRF PC1 (patients reduced peak anterior and posterior force)
11. Medio-lateral GRF PC2 (patients reduced loading rate and larger medial force around 50% stance)

The NP50 classifier identified four gait features differentiating patients from NP50, which were absent from the mixed-age classifier, but were also the least

accurate in classifying participants (accuracy < 75%), comprising (in descending accuracy order):

- Ankle power PC2
- Hip adduction-abduction moments PC1
- Ankle inversion-eversion angles PC3
- Hip flexion-extension angles PC1

Of note, for each of the above variables, ankle power, hip adduction-abduction moments and hip flexion-extension angles, two PCs appeared as the highest-ranking features and just one of them was not in common between the classifiers.
Features that appeared in the mixed-age classifier but were missing in the NP50 classifier were (in descending accuracy order):

- Knee adduction-abduction moments PC2 (81% accuracy)
- Ankle dorsi-plantarflexion angles PC1 (77% accuracy)
- Hip in/external rotation angles PC2 (71% accuracy)
- Pelvis angles in the transverse plane PC2 (70% accuracy)

In both classifiers vertical GRF PC1 (the interpretation was equal in the two classifiers and described the patients' loss of the biphasic nature of the waveform with a peak GRF at 50% stance), hip power PC2 (the interpretation was slightly different between the two classifiers, with the common feature of a reduced magnitude of the waveform during gait) and anterior-posterior GRF PC1 were the first three most discriminating features between NP and patients, and the accuracy levels were nearly the same.

Most of the gait features differentiating NP from patients pre-surgery were related to gait kinetics and GRFs both in the NP50 classifier (66.7%, n = 12), and the mixed-age classifier (61.1%, n = 11).

In both classifiers, excluding GRF, most gait features were related to hip biomechanics, especially in the NP50 classifier (mixed-age classifier: n = 6; NP50 classifier: n = 8), followed by the knee (n = 4) and ankle (n = 3) in the mixed-age classifier, while there were more ankle features (n = 4) discriminating NP50 from patients than knee features (n = 3) in the NP50 classifier. Pelvis features were the least represented in both classifiers (mixed-age classifier: n = 2; NP50 classifier: n =

1). Additionally, excluding joint powers and GRF features, most features were relative to sagittal plane joint angles and moments (n = 6) in both classifiers. A similar number of frontal plane features separated patients from NP in the NP50 classifier (n = 5) than in the mixed-age classifier (n = 4). Only pelvis and hip angles in the transverse plane were entered in the classifier, therefore, two features were related to the transverse plane in the mixed-age classifier, while none appeared in the NP50 classifier.

The gait features accuracy in classifying patients pre-TKR and NP was higher in the mixed-age classifier, with thirteen out of eighteen features (72.2%) having an accuracy ≥ 80%. In the NP50 classifier, ten gait features (55.5% of the total) had these accuracy levels. In both classifiers, most gait features differentiating patients from NP were PC2 and PC3, depicting variations in gait characteristics (such as delayed peaks and not only magnitude differences), rather than PC1, usually depicting a waveform magnitude difference between groups (Deluzio and Astephen 2007) such as Vertical GRF PC1.

Table 20 Comparison of PC interpretation and ranking between the mixed-age classifier and the NP50 classifier. shaded in grey: the low PC interpretations were similar or the same between groups.

			Ranking of the PC based on accuracy (%)		**Interpretation of low PC score (in brackets, the group the low PC score is referred to)**	
			Patients compared to mixed-age NP	Patients compared to NP50	Patients compared to mixed age NP	Patients compared to NP50
JOINT ANGLES						
Pelvis	Frontal plane	PC2	5 (85)	7 (83)	Difference: elevation of the pelvis towards the limb in mid-stance and elevating the pelvis towards the contralateral stance limb during all swing phase (opposite to what high PC scores - the NPs - displayed) (pre-TKR)	Difference: elevation towards the stance limb between 10% and 60% of the gait cycle and lowering of the pelvis towards the stance limb during swing (the above is opposite to what found in high PC) (pre-TKR)
	Transverse plane	PC2	18 (70)	N/A	Difference: opposed to high PC, presence of rotation away from the stance limb at 0 and 100% and rotation towards the stance limb at 50% gait cycle: larger (peaks) ROM throughout gait (NP)	N/A
Hip	Flexion-extension	PC1	N/A	17 (70)	N/A	Magnitude: waveform shifted downwards (smaller hip flexion at heel strike and toe-off and larger hip extension at about 50% of the gait cycle) (NP)
		PC2	16 (74)	13 (74)	Difference: larger flexion at 10 and 85% gait; larger hip extension around 50% of the gait cycle: larger ROM (NP)	Difference: larger hip flexion at 10% and 85% gait; larger hip extension at about 50% of the gait cycle: larger ROM (NP)
	Adduction-abduction	PC2	13 (81)	11 (78)	Difference: larger hip adduction between 20-45% gait and larger abduction around 70% gait: larger hip ROM (NP)	Difference: adduction between 20 and 45% gait cycle (opposed to hip abduction in high PCs), abduction from about 70% of the gait cycle (where high PC have adduction) (NP)
	Internal-external rotation	PC2	17 (71)	N/A	Difference: presence of internal rotation just after heel strike, external rotation at about 45% gait cycle and internal rotation from 85% gait cycle (all of the above opposite to high PC) (pre-TKR)	N/A
Knee	Flexion-extension	PC2	10 (83)	10 (80)	Difference: smaller knee flexion after heel strike and late swing, larger knee peak flexion during swing: larger (peaks) knee ROM (NP)	Difference: smaller knee flexion after heel strike and late swing, larger knee peak flexion during swing: larger (peaks) knee ROM (NP)

Ankle	Dorsi-plantar flexion	PC1	15 (77)	N/A	Magnitude: the waveform is shifted downwards (larger plantarflexion after heel strike, smaller dorsiflexion during stance, larger plantarflexion at initial swing, smaller peak dorsiflexion from about 85% gait cycle) (NP)	N/A
	Inversion-eversion	PC3	N/A	16 (70)	N/A	Difference: reduced inversion around 60% of the gait cycle (pre-TKR)
JOINT MOMENTS						
Hip	Flexion-extension	PC2	11 (81)	8 (82)	Difference: larger hip flexion moment at 5% stance and larger extension moment at 95% stance (NP)	Difference: larger peak flexion moment around 10% of stance and larger peak extension moment at 95% stance (NP)
	Adduction-abduction	PC1	N/A	15 (72)	N/A	Magnitude: waveform shifted upwards (larger adduction moment), especially around 30% and 70% of the stance (NP)
		PC2	8 (83)	6 (83)	Difference and phase shift: lower adduction moment magnitude in mid-stance and higher loading rate (NP)	Difference and phase shift: lower adduction moment magnitude in mid-stance and higher loading rate (NP)
Knee	Flexion-extension	PC2	7 (83)	5 (84)	Difference: reduced flexion moment around 20% stance and reduced extension moment around 70% stance (pre-TKR)	Difference: smaller flexion moment at heel strike and 95% stance, but especially at 20% of stance; smaller/absent extension moment around 70% stance (pre-TKR)
	Knee adduction-abduction	PC2	12 (81)	N/A	Difference: smaller adduction moments at 20% and 90% stance and larger adduction moments around 45% stance (pre-TKR)	N/A
Ankle	Dorsi-plantar flexion	PC1	9 (83)	N/A	Difference: larger dorsiflexion moment around 20% stance and a smaller peak dorsiflexion moment around 85% stance (pre-TKR)	N/A
		PC2	N/A	4 (87)	N/A	Difference: smaller plantarflexion moment around 15% stance and smaller peak dorsiflexion moment around 80% stance (pre-TKR)
JOINT POWERS						
Hip		PC2	2 (91)	2 (93)	Difference and phase shifts: earlier power absorption after heel strike, smaller power generation around 60% gait, earlier peak power generation at the end of swing: reduced magnitudes of the waveform during gait (pre-TKR)	Difference: larger power generation after heel strike and at about 15% gait cycle, smaller power generation around 60% gait, large power generation at 80% gait and reduced power generation at 90% gait: reduced magnitudes of the waveform during gait (pre-

						TKR)
Knee		PC1	4 (86)	12 (75)	Difference: larger peak power absorption and generation throughout the gait cycle (NP)	Difference: larger peak power absorption and generation throughout the gait cycle (NP)
Ankle		PC2	N/A	14 (74)	N/A	Phase shift: delayed peak power generation (NP)
		PC3	6 (84)	18 (68)	Difference: mainly larger power absorption around 10% gait cycle, smaller power generation around 55% gait cycle, difficult to interpret the meaning at 80 and 90% of the gait cycle (no apparent differences) (pre-TKR)	Difference: smaller power absorption between about 10-25% gait. Difference around 85% gait of difficult interpretation (NP)
FORCES						
Ground reaction force	Vertical	PC1	1 (94)	1 (95)	Phase shift: delayed at about 40% and anticipated at 60% of stance (reduced loading rate), higher magnitude around 50% stance: loss of by-phasic nature of the waveform (pre-TKR)	Phase shift: delayed at about 40% and anticipated at 60% of stance (reduced loading rate), higher magnitude around 50% stance: loss of by-phasic nature of the waveform (pre-TKR)
	Anterior-posterior	PC1	3 (87)	3 (87)	Difference: reduced peak anterior and posterior force (pre-TKR)	Difference: reduced peak anterior and posterior force (pre-TKR)
	Medio-lateral	PC2	14 (79)	9 (82)	Difference and phase shift: reduced loading rate and presence of larger medial force around 50% stance (pre-TKR)	Difference and phase shift: reduced loading rate and presence of larger medial force around 50% stance (pre-TKR)

NP: non-pathological participants; PC: principal component; NP50: non-pathological participants ≥ 50 years old; TKR: total knee replacement; N/A: not available; ROM: range of motion.

5.3.2 Prediction of objective function post-TKR

Given the differences between the mixed-age classifier and the NP50 classifier, it was decided to utilise the NP50 classifier to evaluate the objective function and utilise this for building a prediction model of objective TKR outcome.

Table 21 shows the potential factors affecting TKR outcome utilised for the correlation analysis. On the mean, patients had obesity class I (WHO 2010), and 63.6% were obese. It was not possible to access the medical notes for six patients (seven knees) due to the type of consent they had signed (relative to a protocol which did not contemplate accessing medical information for the study). Additionally, the medical notes were not retrievable for two patients who had consented for the medical notes to be accessed. These nine cases did not have significant age differences ($p = 0.909$), BMI ($p = 0.731$), BOA pre-TKR ($p = 0.076$), BOA post-TKR ($p = 0.527$), BOA change ($p = 0.237$) compared to the remaining thirty-five cases. Thirty-one patients, with thirty-five TKRs (i.e., four patients had a bilateral TKR, two at different times, two simultaneously) (79.6% of the total 44 knees) had available medical records. Of these, four patients (five knees) had an incomplete set of notes in their medical information (two regarding the knee patella resurfacing; four regarding the ASA score).

Table 21 Potential factors affecting TKR outcome tested for their relationship to the BOA post-TKR and the change in BOA pre to post-TKR

Potential factors affecting gait function	Patients pre-TKR (n = 44)
Age (years), mean (SD)	67 (7)
Females, n (% of the total)	14 (31.8)
BMI (kg/m^2), mean ± SD	33.7 ± 7.0
Obesity class III, n (% of the total)	7 (15.9)
Obesity class II, n (% of the total)	10 (22.7)
Obesity class I, n (% of the total)	11 (25.0)
Pre-obesity, n (% of the total)	14 (31.8)
Normal weight, n (% of the total)	2 (4.6)
Previous major knee surgeries (affected knee), n (% of the total)	4 (9.1)
Contralateral knee issues/previous surgeries, n (% of the total)	34 (77.3)
Surgeries in other lower limb joints excluding knees, n (% of the total)	9 (20.5)
Hip/s issues/previous major surgeries, n (% of the total)	16 (36.4)
Painful ankle/s, n (% of the total)	7 (15.9)
ASA, n (% of the total)	
1	6 (13.6)
2	19 (43.2)
≥ 3	6 (13.6)
Missing data	13 (29.5)
Patella resurfacing, n (% of the total)	
Yes	16 (36.4)
No	17 (38.6)
Missing data	11 (25.0)
Implant design, n (% of the total)	
Posterior-sacrificing	4 (9.1)
Cruciate-retaining	31 (70.5)
Missing data	9 (20.5)
Post-surgery complications, n (% of the total)	
Yes	5 (11.4)
No	30 (68.2)
Missing data	9 (20.5)
BOA pre-TKR, mean (SD)	0.76 (0.20)
OKS pre-TKR, median (IQR)	n = 42 21 (16)
KOOS pre-TKR, mean (SD)	n = 37
Symptoms	48.4 (19.6)
Pain	44.8 (20.7)
ADL	50.2 (20.2)
Sport/Rec (n = 36)	22.9 (18.2)
QoL	26.4 (21.9)

TKR: Total Knee Replacement; n: number; SD: standard deviation; ASA: American Society of Anaesthesiologists; ROM: range of motion; BOA: belief of osteoarthritis; OKS: Oxford Knee Score; IQR: interquartile range; KOOS: Knee Injury and Osteoarthritis Outcome Survey; ADL: activities of daily living; QoL: quality of life.

Only 42 patients filled out the OKS pre-TKR, the median value was 21 (16 IQR), and ranged from a minimum of 6 to a maximum of 36. Of the patients with an available KOOS questionnaire pre-TKR (n = 37), one patient left too many blanks within the KOOS sports and recreational activities section to allow calculating the KOOS pre-TKR score. The data missing due to the questions being left blank was related to the characteristics of the questions, as this patient did not carry out the sports activities investigated (i.e. squatting, running, jumping, twisting/pivoting on the affected knee, and kneeling). Among all KOOS sub-scales, patients perceived their sport and recreational activities and knee-related quality of life as being affected the most by the presence of knee OA pre-TKR (lowest sub-score).

Research question: Which patient and clinical preoperative and surgical factors are associated with change in BOA and BOA post-TKR?

Hypothesis 2.1 – A negative, weak relationship exists between BOA pre-TKR and BOA change pre- to post-TKR

There were no statistically significant correlations between any of the preoperative or surgical factors and the BOA change one year after the surgery, as shown in Table 22. This is in contrast with the initial hypothesis stating that the gait function pre-

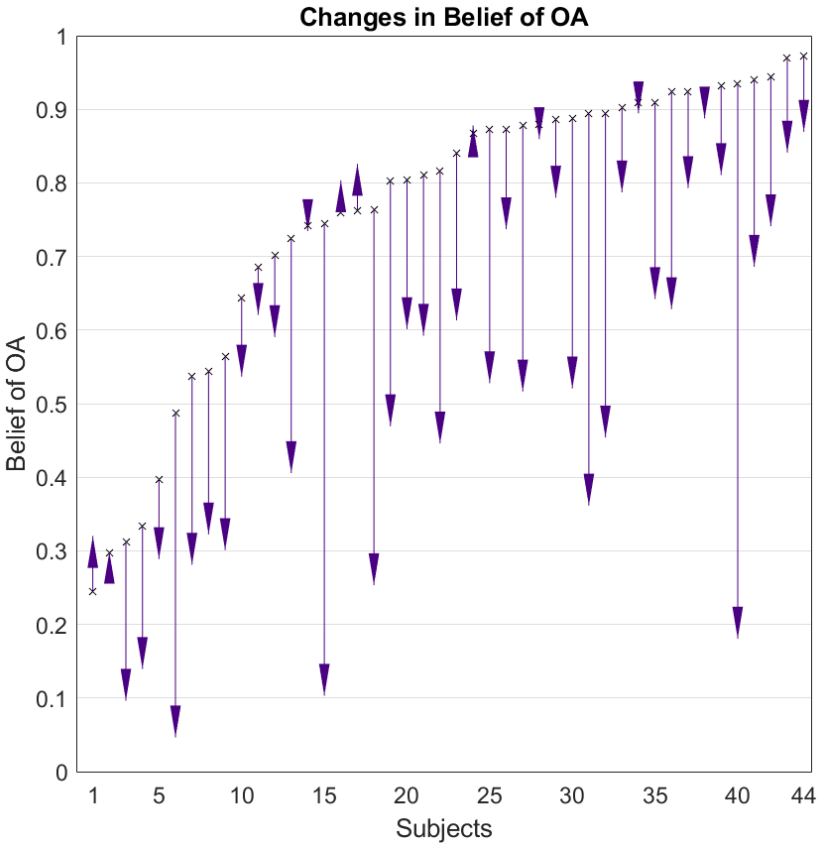


Figure 26 BOA change pre to one-year post-TKR for the patients within the NP50 classifier, in order of increasing BOA pre-TKR. The cross represents the BOA pre-TKR, the tip of the arrow is the BOA post-TKR, the arrow represents the change in BOA pre to post-TKR. On the abscissa, the subject ID, on the ordinate, the BOA, ranging from 0 to 1.

TKR was associated with the change in function post-TKR.

Figure 26 shows the change in BOA pre to post-TKR, where the “x” identifies the BOA pre-TKR, the purple arrow the change in BOA and the tip of the arrow represents the BOA post-TKR. By observing the figure, the change in gait function did not seem to follow an obvious pattern, showing that the patients whose BOA pre-TKR was higher, did not always achieve the largest change in gait quality pre to post-TKR.

Hypothesis 2.2 – There is a strong, positive association between the BOA pre-TKR and the BOA one year post-TKR (the BOA pre-TKR can predict the BOA post-TKR)

Table 22 shows that several statistically significant correlations were found between BOA post-TKR and various potential predictors, the relationship was:

- Weak to moderate with sex: being a female was associated with a higher BOA post-TKR or worse function
- Weak to moderate with previous major surgeries on the affected knee: having had major surgery on the affected knee before TKR was correlated with a lower BOA post-TKR or better function
- Weak to moderate with BMI: a higher BMI pre-TKR was correlated to a higher BOA post-TKR
- Moderate with ASA score (n = 31): having a higher ASA score, or worse health, was associated with a higher BOA post-TKR
- Moderate to strong with BOA pre-TKR (in accordance with the hypothesis): a higher BOA, or worse function, pre-surgery was associated with a higher BOA post-TKR

These potential predictors were, therefore, utilised to develop multivariable regression models for predicting BOA post-TKR.

Table 22 Correlation between predictors and BOA post-TKR and predictors and change in BOA one-year post-TKR

	Correlation coefficients				
Predictors	BOA post-TKR (n = 44 unless otherwise stated)	Significance level	Interpretation (where significant)	BOA change (n = 44 unless otherwise stated)	Significance level
Age	0.128‡	0.409		0.095‡	0.539
BMI	0.366‡	0.015*	Weak to moderate	0.156‡	0.313
Sex (0: male; 1: female)	0.309	0.041*	Weak to moderate	0.071	0.649
Previous major surgeries on the affected knee (0: no; 1: yes)	-0.367	0.014*	Weak to moderate	-0.117	0.451
Contralateral knee issues (0: no; 1: yes)	-0.010	0.949		0.144	0.352
Surgeries in other lower limb joints (other than knees) (0: no; 1: yes)	0.060	0.698		-0.039	0.800
Painful hip/s (0: no; 1: yes)	0.236	0.122		0.005	0.973
Painful ankle/s (0: no; 1: yes)	0.151	0.328		0.108	0.487
ASA score (n = 31)	0.445‡	0.012*	Moderate	0.163‡	0.381
Patella resurfacing (0: no; 1: yes) (n = 33)	0.151	0.328		0.108	0.487
Implant design (0: cruciate retaining; 1: cruciate sacrificing)	0.265	0.124		0.267	0.122
Post-surgery complications (0: no; 1: yes)	0.216	0.212		0.124	0.477
BOA pre-TKR	0.641‡	< 0.001*	Moderate to strong	-0.115‡	0.456
OKS pre-TKR (n = 42)	-0.144‡	0.362		0.080‡	0.614
KOOS	n = 37			n = 37	
Symptoms	-0.157‡	0.354		-0.021‡	0.904
Pain	-0.025‡	0.882		0.168‡	0.320
ADL	-0.121‡	0.477		0.150‡	0.377
Sport/Rec (n = 36)	-0.086‡	0.619		0.044‡	0.798
QoL	-0.241‡	0.150		0.114‡	0.501

BOA: Belief of Osteoarthritis; TKR: Total Knee Replacement; n: number; BMI: Body Mass Index; ASA: American Society of Anaesthesiologists; ROM: range of motion; KOOS: Knee Injury and Osteoarthritis Outcome Survey; ADL: activities of daily living; QoL: quality of life; ‡: Spearman's rho. *: statistically significant at $p < 0.05$.

Research question: If associations exist between patient and clinical preoperative and surgical factors and BOA change or BOA post-TKR, could they be utilised to develop prediction models of BOA change and BOA post-TKR?

Due to the lack of association between preoperative and surgical factors and the change in BOA, the regression model predicting the BOA change was not developed.

A regression model including the following predictors: sex, BMI, ASA score, previous major surgeries on the affected knee, and BOA pre-TKR was developed on the patients who had data available across all predictors (n = 31) (outcome: BOA post-TKR, which was normally distributed). The prediction model satisfied all assumptions (Appendix D). It is worth noting that the model included a large range of BMI scores, but most subjects (61.3%) were obese, 32.3% were pre-obese, and 6.5% had normal weight. Additionally, a minor proportion of females was represented (16.1%) and only 9.7% of the patients had had previous major surgeries. Additionally, only 19.4% of patients had an ASA of 1, while the majority (61.2%) had an ASA of 2 and the reference category, ASA ≥ *3*, included only 19.4% of subjects, as shown in Table 23.

Table 23 Predictors entered in the multiple regression model and patients' details for each predictor

Predictors	Patients pre-TKR included in the model (n = 31)
BMI, mean ± SD (range)	33.7 ± 7.5 (21.3 - 52.4)
Females, n (% of the total)	5 (16.1)
Previous major knee surgeries (affected knee), n (% of the total)	3 (9.7)
ASA, n (% of the total)	
1	6 (19.4)
2	19 (61.2)
≥ *3*	6 (19.4)
BOA pre-TKR, mean (SD)	0.72 (0.22)
BOA post-TKR, mean (SD)	0.51 (0.26)

TKR: Total Knee Replacement; n: number; BMI: Body Mass Index; SD: standard deviation; ASA: American Society of Anaesthesiologists; BOA: Belief of Osteoarthritis.

Table 24 shows that the multiple regression significantly predicted the BOA post-TKR ($p < 0.001$). The model explained 63.0% of the variance in the sample examined ($R^2 = 0.630$). The coefficient of determination in the sample, R^2, is usually larger than the one found in the larger population, and therefore positively biased. The adjusted R^2 is a more suitable parameter to consider when aiming to apply the findings to the wider population, and it showed that the combination of the independent variables explained 53.8% of the BOA post-TKR variance. Only the preoperative gait function (BOA pre-TKR) added significantly to the prediction ($p = 0.011$). Therefore, looking at the regression coefficient, when the other independent variables were kept constant (i.e., while controlling for the other variables), having a higher BOA pre-TKR was associated with a significant increase in the BOA post-TKR (i.e., an increase of the BOA pre-TKR by 0.10, was correlated with an increase in the BOA post-TKR by 0.06, or rather, having a poorer gait function pre-TKR, was associated with a worse gait function post-TKR).

The remaining pre-operative factors were not significant in the model as their regression coefficient did not differ significantly from zero. This indicated no strong linear relationship between the predictors and the outcome, or rather, when the other variables were controlled for, a change in that predictor would produce a non-significant change in the outcome.

Table 24 Multiple linear regression results for predicting the BOA post-TKR (n = 31)

BOA post-TKR	β	Significance level	95% CI for β		R^2	Adjusted R^2
			Lower Limit	Upper Limit		
Model		< 0.001*			0.630	0.538
Constant	-0.075	0.747	-0.551	0.401		
BMI	0.006	0.272	-0.005	0.018		
Sex (reference: 0, males)	0.126	0.195	-0.069	0.320		
Previous major knee surgeries (affected knee) (reference: 0, no surgeries)	-0.121	0.357	-0.386	0.145		
ASA 1	-0.104	0.393	-0.352	0.143		
ASA 2	-0.077	0.439	-0.279	0.125		
BOA pre-TKR	0.595	0.011*	0.149	1.042		

BOA: Belief of Osteoarthritis; TKR: total knee replacement; Model: all predictors were entered at once in the model; β: unstandardised regression coefficient; CI: confidence interval; R^2: coefficient of determination; BMI: Body Mass Index; ASA: American Society of Anaesthesiologists.
*: statistically significant at $p < 0.05$

The predictors (n = 7) to participants (n = 31) ratio was approximately 1:5. Based on their significance in the previous regression model (either significant, or the closest to significance with respect to the other predictors), and to keep the 1:10 ratio, three predictors were brought forward in the final regression model. BOA pre-TKR, BMI and sex were utilised as predictors of BOA post-TKR in the final regression model. The model was developed on the patients who had data available across all predictors (n = 44) and it showed that patient ID 90 was an outlier (±3 SD studentised deleted residuals), with a studentised residual of -3.10. Patient ID 90 was a female, with a BMI of 26.1 (overweight), who had not had previous major knee surgeries; moreover, the patient had an ASA grade of 3, and had a higher BOA pre-TKR (0.93) than other patients, but a very low BOA post-TKR (0.18). Patient ID 90 was removed, forty-three patients were included in the final model, and this allowed to satisfy the multiple linear regression assumptions (Appendix E). The model included n = 13 females (30.2% of the total 43 cases), almost double the number included in the previous model. Moreover, the mean BMI was 33.9 ± 7.1 (range 21.3 - 52.4), slightly lower than the previous model. The mean BOA pre-TKR was 0.76 ± 0.20 and the mean BOA post-TKR was 0.56 ± 0.25, both somewhat higher than the previous model.

Table 25 Multiple linear regression results for predicting the BOA post-TKR (n = 43)

BOA post-TKR	β	Significance level	95% CI for β		R^2	Adjusted R^2
			Lower Limit	Upper Limit		
Model		< 0.001*			0.591	0.560
Constant	-0.245	0.082	-0.523	0.032		
BMI	0.006	0.132	-0.002	0.013		
Sex (reference: 0, males)	0.095	0.100	-0.019	0.210		
BOA pre-TKR	0.769	< 0.001*	0.495	1.044		

BOA: Belief of Osteoarthritis; TKR: total knee replacement; Model: all predictors were entered at once in the model; β: unstandardised regression coefficient; CI: confidence interval; R^2: coefficient of determination; BMI: Body Mass Index; ASA: American Society of Anaesthesiologists.
*: statistically significant at $p < 0.05$

Table 25 shows that the multiple regression significantly predicted the BOA post-TKR ($p < 0.001$), confirming the hypothesis. The model explained 56.0% of the variance in the larger population, slightly higher than the previous model. As observed in the previous model, only the preoperative gait function (BOA pre-TKR)

added significantly to the prediction ($p < 0.001$) and the 95% confidence interval was shifted more to the positive side compared to the previous model, confirming the strength of the finding. Therefore, looking at the regression coefficient, when controlling for the other variables, having a higher BOA pre-TKR was associated with a significant increase in the BOA post-TKR (i.e., an increase of the BOA pre-TKR by 0.10, was correlated with an increase in the BOA post-TKR by 0.08).

5.3.3 Association between patient knee ROM and BOA

Knee ROM was measured with a clinical goniometer in three of the patients, due to a protocol change throughout the TKR study, and therefore, their data was not accounted for due to the different methods utilised. One patient had a TKR in both knees but the data was of sufficient quality only for one of the knees (ID 02429 only R side). Knee ROM good quality data was available only for twenty-one knees (47.7% of the total forty-four knees). Patients pre-TKR with available knee ROM data had knee ROM limitations as the mean ROM was 97˚ ± 17 (range 56˚ to 123˚) pre-TKR. The knee ROM post-TKR did not improve significantly ($p = 0.143$, paired t-test) and was 99˚ ± 12 (range 73˚ to 122˚), and the mean improvement from pre-surgery was 4˚ ± 11 (range -12˚ to 26˚). The knee ROM improvement was due to a slight recovery of knee extension (mean recovery of 6˚ ± 9), while maximum knee flexion decreased pre to post-TKR (mean maximum flexion decrease: -3˚ ± 10).

Table 26 Correlation between knee ROM and BOA values

	Correlation coefficients (interpretation, where significant), *p-value*		
(n = 21)	BOA pre-TKR	BOA post-TKR	BOA change
Knee ROM pre-TKR	-0.618‡ (Moderate) *0.005**	-0.504‡ (Moderate) *0.028**	-
Knee ROM post-TKR	-	-0.277 *0.251*	-
Knee ROM change pre to post-TKR	-	-	0.121‡ *0.622*

BOA: Belief of Osteoarthritis; TKR: Total Knee Replacement; n: number; ROM: Range of Motion; ‡: Spearman's rho.
*: statistically significant at $p < 0.05$.

Hypothesis 3.1 – Pre-surgery, knee ROM is associated with BOA

It was found that having a larger knee ROM pre-surgery was associated with a smaller BOA pre-TKR or rather a better gait function, confirming the hypothesis.

Hypothesis 3.2 – Post-surgery, knee ROM is associated with BOA

Hypothesis 3.3 – From pre- to post-surgery, the change in knee ROM is associated with the change in BOA

No relationship was found between the knee ROM post-surgery and BOA post-TKR, nor between the knee ROM change and the BOA improvement pre to one-year post-TKR. The results contrasted with the hypotheses.

Hypothesis 3.4 – Knee ROM pre-TKR is associated with the BOA post-TKR

It was found that having a larger knee ROM, pre-TKR was moderately associated with a better BOA one year post-TKR, this confirmed the hypothesis.

5.4 DISCUSSION

5.4.1 Preliminary work

One of the aims of the preliminary work in this chapter was to determine whether comparing the same group of patients analysed in chapter 4 to a control group older than 50 years of age would produce a different gait classification, than comparing them to a group of controls of mixed age ranging from 19 to 79.

Hypothesis 1.1 – The patients' Belief of OA is lower using the NP50 classifier, as opposed to utilising the mixed-age classifier, both pre- and post-TKR

By the hypothesis, it was found that comparing patients to a group of NPs which included mostly participants younger than 50, resulted in underestimating the pre- and post-operative gait function of patients. This meant that the BOA value was significantly lower when patients were compared to a group of NPs of similar age. The finding was not surprising as it is widely accepted that older age causes significant changes in gait, and these have common characteristics with OA gait, such as slower walking speed (Herssens et al. 2018), reduced pelvic rotations (Cury et al. 2021), increased hip flexion and reduced hip extension (Boyer et al. 2017c), alterations in knee flexion-extension angles (Chehab et al. 2017), reduced plantarflexion angles at toe-off (Pol et al. 2021b), and dorsiflexion moments (Boyer et al. 2017c). Therefore, comparing patients with older NP who would have manifested gait alterations due to ageing, naturally resulted in patients having a "better" gait than when compared to younger controls.

The BOA values in the NP50 classifier (mean BOA pre-TKR: 0.76 ± 0.20; mean BOA post-TKR 0.55 ± 0.25) were similar to those observed for the Karolinska cohort in Chapter 4, both pre (mean BOA Karolinska: 0.73 ± 0.23) and post-TKR (mean BOA Karolinska: 0.53 ± 0.25). However, it is difficult to compare the current findings to previous studies including older NP within the Cardiff classifier (Worsley 2011; Metcalfe 2014). Worsley (2011) failed to report the BOA pre-TKR, while the BOA post-TKR was 0.48 ± 0.17, which was lower than the present study. However, Worsley included patients undergoing UKR and entered a different set of variables in the classifier, including knee passive ROM, and rectus femoris atrophy, this may justify the difference. Metcalfe (2014) found a lower BOA both pre (mean 0.66 ± 0.26) and post-TKR (mean 0.33 ± 0.28). Nonetheless, the exclusion criteria for

Metcalfe's study (2014) were much more restrictive than the current one, as patients with pain in any joint other than the operative knee, those who had previous trauma or surgery in the lower limbs, or lower back pain, were excluded. The patients in the study of Metcalfe (2014) would have been functionally better compared to the patients involved in the present study, due to the lack of issues in other joints, therefore, their BOA values were much lower.

Hypothesis 1.2 – Based on observations in the previous chapter, the change in Belief of OA for patients is similar when using the NP50 classifier or the mixed-age classifier

In line with the hypothesis, patients had a comparable improvement in gait function between the two classifiers. Both the BOA pre-TKR and post-TKR in the NP50 classifier were smaller than BOA scores calculated against mixed-age NP. However, the magnitude of the difference between BOA pre-TKR of the NP50 classifier VS. the mixed-age classifier and the BOA post-TKR of the NP50 classifier VS. the mixed-age classifier was similar. Therefore, the BOA change score from the NP50 classifier was not statistically different to the one calculated in the mixed-age classifier.

Most of the previous studies adopting the classifier failed to report on the BOA change score across patients, describing only BOA values at baseline and follow-up, or reporting these values for single patients (Whatling 2009; Watling 2013; Biggs 2016; Biggs et al. 2019b). One previous report observed a -0.15 BOA change pre to 9-24 months post-TKR (Biggs et al. 2019a), which was slightly smaller than the current study. Some of the subjects in Biggs et al. (2019a) were the same as in the current study [35.9% of the patients pre-TKR (n = 14), 35.9% of post-TKR participants (n = 14) and 64.7% of NPs (n = 11)], but patients were compared to a group of NPs of mixed-age, and it was not clear what were the input variables used in the classifier. The use of a different set of input variables, and a slightly different cohort may justify the difference in BOA change from the current study. Nevertheless, the improvement in gait function found in the NP50 classifier (mean BOA change -0.21 ± 0.19), was similar to that observed in the Karolinska patients, who had been assessed against age-matched NP (mean BOA change -0.20 ± 0.14) in Chapter 4.

It must be noted that no previous research explored how the BOA at baseline/follow-up or its change may be affected by comparing OA patients gait function to different sets of NPs when using the classifier, and, therefore, the findings of this work were novel and it was difficult to compare them to similar earlier work.

Hypothesis 1.3 – The highest-ranking gait features discriminating patients from older NPs in the NP50 classifier are similar to those separating patients from NPs in the mixed-age classifier

Most of the gait features discriminating patients from NPs were the same or similar between the NP50 and the mixed-age classifier. The hypothesis was confirmed, and the analysis highlighted that the differences in the eighteen highest-ranking gait features between the classifiers concerned those with the lowest accuracy. This showed that knee OA produces some main modifications in gait patterns that are typical of the disease, and that appear even when older controls, who could have some age-related gait modifications, are utilised. It must be noted that the older NPs utilised in the mixed-age classifier were almost the same as the ones entered in the NP50 classifier. Therefore, certain gait characteristics relative to "young age" (due to the presence of NP < 50 years old) had already been mitigated by the presence of "older age" gait characteristics in the mixed-age classifier. Because of this, it was not expected to observe a substantial difference in the highest-ranking gait features discriminating between patients pre-TKR and NP within the NP50 classifier. In the NP50 classifier, gait features for the variables hip flexion-extension angles (PC1 and PC2), hip adduction moments (PC1 and PC2), and ankle joint power (PC2 and PC3) appeared twice, described by different PCs for the same variable, and highlighting slightly different features discriminating patients from NP50s for that variable, suggesting that aspects of gait at these levels are very relevant in differentiating patients from NP of similar age. Some of the 18 highest-ranking gait features were present only in the mixed-age classifier and not in the NP50 classifier but had relatively low accuracy, such as ankle dorsi-plantarflexion angles PC1 (describing controls' larger plantarflexion after heel strike, smaller dorsiflexion during stance, larger plantarflexion at toe-off, smaller peak dorsiflexion from about 85% gait cycle), hip in/external rotation angles PC2 (the presence of internal rotation just after heel strike, external rotation at about 45% gait cycle and

internal rotation from 85% gait cycle, a pattern present in controls and opposite in patients, who maintained external rotation throughout), pelvis angles in the transverse plane PC2 (controls' larger ROM).

Various systematic reviews showed that older adults have a significantly reduced ankle plantarflexion at toe-off (Boyer et al. 2017c; Pol et al. 2021b) and smaller ankle ROM throughout gait (Boyer et al. 2017c), a reduction in the pelvic rotations in the transverse plane, but not in the sagittal and frontal planes (Cury et al. 2021). These movement alterations related to older age are similar to those of people with knee OA, who have reduced ankle ROM (Ro et al. 2019), smaller ankle dorsiflexion during stance, and a reduced plantarflexion at initial swing (Ismailidis et al. 2020); this may justify why the abovementioned PCs were discriminative between patients and mixed-age NP but not between patients and NP50. Conflicting findings exist regarding hip movement differences between subjects with knee OA and controls, some studies found an increased hip internal rotation in OA (McKean et al. 2007), while others did not (Esrafilian et al. 2013). The transverse plane is subject to lower data reliability in 3D motion analysis (McGinley et al. 2009) and therefore, the inconsistencies of findings in the literature could be due to measurement inaccuracies. It could also be suggested that hip rotations may be subjected to a lot of variability between individuals, and differences between groups may depend on the subjects included in the analysis rather than being related to the disease investigated.

It must be highlighted that while differences in the patients' gait classification were found between the mixed-age and the NP50 classifiers, it is not clear whether these were clinically significant. Due to not having an availability of patients' views regarding their improvement post-TKR, it was not possible to determine the minimal clinically important change of the BOA. Further research should explore this aspect to expand the knowledge of the measuring capabilities of the Cardiff classifier.

Given that the findings highlighted that utilising NP50 in the classifier potentially yielded a better representation of the patient's gait function at least pre- and post-TKR, and given that function pre-TKR was one of the key predictors to be tested, it was decided to develop prediction models of function post-TKR and change in BOA post-TKR based on the NP50 classifier. Moreover, it is recommended that when comparing participants in a cohort study including several groups, these

should have as much similar characteristics as possible, to ensure the effect investigated is the main one differing between them, in this case, the biomechanics of the lower limb in the presence of knee OA and post-TKR, to reduce systematic bias (Greenhalgh 2019).

There are challenges in recruiting NP subjects aged 60 or older who are free from musculoskeletal pain or conditions. This is because musculoskeletal disorders are frequent in older individuals, tend to increase with age, and often involve multiple sites (Fejer and Ruhe 2012). While the author attempted to increase the cohort of NPs throughout the PhD project, this process proved challenging (1) for the abovementioned reasons, (2) due to limited interest shown from the public within the Cardiff area following recruiting messages, (3) because of the COVID-19 pandemic, which took place three months into the start of data collection. Further investigations utilising the classifier should aim to increase the NPs older than 50 groups, considering that recruiting older NPs for gait studies utilising motion capture technologies is achievable, as demonstrated by other research groups exploring biomechanical differences between age-matched controls and patients with knee OA (Deluzio and Astephen 2007; Crossley et al. 2018; Na et al. 2018), or patients post-TKR (Ro et al. 2020), although it may take longer time and more recruiting resources than including NP subjects of much younger age.

5.4.2 Prediction of objective function post-TKR

The main aim of this study was to explore which factors could be predictive of the gait function one-year post-TKR.

Hypothesis 2.1 – A negative, weak relationship exists between BOA pre-TKR and BOA change pre- to post-TKR (i.e., the BOA pre-TKR can predict the BOA change)

In contrast with the hypothesis, it was found that having a substantially compromised gait function as a result of severe knee OA was not associated with a larger improvement in gait function following TKR. Yet, the correlation coefficient, although not statistically significant, suggested that gait function pre-TKR and BOA change were inversely correlated, as hypothesised. The previous chapter and an earlier report (Biggs 2016) showed that the gait recovery (i.e., BOA change) can be very heterogeneous; this finding was confirmed in the current study, and Figure 26

shows that subjects with a similar BOA pre-TKR had different levels of gait improvement. When looking at other factors that may have been associated with the change in gait, such as pre-operative, surgical, anthropometric, or demographic factors, neither of them was found to have a relationship with the level of gait improvement/worsening. Moreover, neither baseline pain levels, nor self-perceived function measured via the KOOS and the OKS was associated with the objective level of functional improvement (BOA change). The OKS is routinely collected in the NHS as a key measure of outcome. Patients involved in the current study had a preoperative OKS (median 21, 16 IQR), in line with scores found by previous authors (mean OKS of 110 English patients undergoing TKR: 18) (Dakin et al. 2020). The OKS score found in the present study indicated the appropriateness of the surgery in all cases (OKS range 6 – 36), as recent research suggested that OKS < 41 have an 88% chance of reaching a meaningful improvement (Price et al. 2020).

The current findings confirm what was found by Biggs (2016), where no correlation was observed between the gait improvement (i.e., BOA change) of 22 patients 9+ months post-TKR and BOA pre-TKR, or OKS, age, or BMI at baseline. The present results, however, were in contrast with what was discovered by Worsley (2011), where the combination of sex, age, pre-surgery activity, BMI, baseline knee ROM, the degree of rectus femoris atrophy, and the function pre-TKR significantly contributed to predicting the BOA change ($R^2 = 0.579$, $p = 0.006$). Subjects with better function pre-surgery had a smaller improvement in gait post-arthroplasty (including TKR and UKR) (Worsley 2011). Nevertheless, it was not clear which independent variable was a significant predictor of BOA change, as only the regression coefficients, but not the statistical significance of each predictor in the model, were reported. Additionally, the number of predictors to patient ratio in the regression model was very small (less than 4 patients per predictor), and therefore, the results should be interpreted with a high degree of caution. With a small number of patients, the addition of just a few more subjects could lead to quite different results, as ascertained during the analysis for the current study (results not reported). However, Worsley (2011) entered variables in the Cardiff classifier that were much different from the present study, including mostly the biomechanics of sit-to-stand, alongside gait variables, gait double support time, knee ROM and rectus femoris muscle atrophy. This may justify the differences between Worsley's findings

and the current ones. Moreover, the study (Worsley 2011) looked at the BOA change 6 months post-surgery, and half of the group of patients (n = 15 out of 31) included subjects who had a UKR, and therefore, medial OA as opposed to generalised knee OA. At 6 months post-surgery, the recovery of function is not complete and there is still scope for improvement, while 12 months is usually the time point where, generally, a patient may be considered recovered, as evidence showed that no significant biomechanical changes happen past this point (McClelland et al. 2014). There is very limited literature exploring predictors of change in the objective function. Some of the previous research has been carried out on patients who received a hip replacement. One study looked at factors predicting the change in GDI two- and six-month post-hip replacement and found that only baseline GDI (but not age, sex or preoperative WOMAC) was a significant predictor (Jensen et al. 2015a). Another recent report included a cohort of thirty-two patients undergoing total hip replacement and found that the combination of preoperative BOA, age and BMI were predictive of the BOA change one-year post-surgery (Biggs et al. 2021). However, the model showed that BOA pre-surgery was not contributing significantly to the prediction, confirming the results from the present investigation, and the main predictor was age (Biggs et al. 2021). Unlike the present study, age was predictive of the outcome, however, the population investigated had hip OA and underwent a different procedure (Biggs et al. 2021), and this makes it difficult to compare the findings.

Most of the previous literature predicting function post-TKR focussed on PROMs and on predicting the absolute score post-surgery, or the minimal clinically important change, rather than the change in the outcome score itself. A previous study on a cohort of 22 691 patients undergoing TKR found that the change in OKS was most strongly predicted by the baseline OKS (a higher preoperative score, namely, a better function, was associated with a smaller change in OKS), the subject's perceived health, the presence of depression and anxiety (Baker et al. 2012). Moreover, it was found that patients with a better PROM score pre-TKR do not achieve a clinically relevant change in their perceived function (Berliner et al. 2017; Goh et al. 2022).

A patient would expect to be suggested what could be done to ensure the gait quality "gets better" post-TKR. When looking at predicting a change in objective gait

function, however, the current findings, combined with the existing literature looking at gait biomechanics, make it challenging to draw a definite conclusion to inform patients on what preoperative factors may aid them to be better in terms of the objective function. This study may suggest that, due to a lack of a clear pattern in the change of gait function, collecting data on a much larger sample of patients may be advisable to establish predictors of change in objective gait function, and this could be the direction for future research. These results may also indicate that no specific factors related to the surgery, a person's characteristics or baseline status seemed to play a role in the amount of recovery a patient could expect. It might be suggested that those with a higher baseline BOA score who should theoretically have a greater recovery potential with TKR surgery, do not necessarily achieve a better improvement than those with a higher baseline objective function.

Hypothesis 2.2 – There is a strong, positive association between the BOA pre-TKR and the BOA one-year post-TKR (the BOA pre-TKR can predict the BOA post-TKR)

The findings confirmed the hypothesis as BOA pre-TKR strongly correlated with the outcome both in the unadjusted (i.e., correlation) and the adjusted analysis (i.e., multiple linear regression). Other factors showed a relationship with the outcome, including BMI, having had previous major surgeries on the affected knee, and ASA score (i.e., comorbidities).

It must be noted that multiple comparisons were carried out between predictors (n = 19) and outcome in the correlation analysis, and no Bonferroni correction was applied. The p-value Bonferroni correction is considered very rigorous (Greenhalgh 2019), and the correlation analysis was conducted purely to identify factors that may have correlated with BOA post-TKR to then enter them in a more robust regression model analysis. Had a Bonferroni correction been applied, the correlation with the outcome would have been statistically significant only for BOA pre-TKR (i.e., 0.05 (significance level) /19 (number of predictors) = 0.003 (new Bonferroni-adjusted significance level)). However, previous research suggested that sex (Kennedy et al. 2006; Liebs et al. 2011; Astephen Wilson et al. 2015; Mandzuk et al. 2015; Pua et al. 2019; Paterson et al. 2020), BMI (Stickles et al. 2001; Xu et al. 2018), and ASA (Kramers-de Quervain et al. 2012; Riddle et al. 2022) correlated

with the TKR outcome and it was deemed appropriate to explore their contribution in predicting the BOA post-TKR.

Since earlier research found a significant correlation between the OKS and BOA values both pre- and post-TKR (Biggs et al. 2019a), it was of interest to investigate whether the preoperative OKS could have any predictive abilities of objective function one-year post-TKR. This information would have been of value to clinicians who do not have access to 3D motion analysis and could have utilised the OKS as a way to inform on the objective function post-surgery. However, no correlation was found between the self-perceived function pre-surgery and the objective function achieved post-TKR. This suggests that it may not be reasonable to utilise the baseline OKS to advise on the objective postoperative function. It may be that the sample size was too small to detect a significant correlation between these variables, but the current findings suggest a trend (although far from significance), where a higher OKS pre-TKR (i.e., better-perceived function) may be related to a lower BOA post-TKR (i.e., better gait function).

The initial linear regression model showed that having a worse gait function pre-TKR predicted having a worse gait quality one-year post-TKR, even when controlling for sex, BMI, comorbidities and having had previous major surgeries. Nevertheless, BMI, sex, previous major surgeries on the affected knee, or comorbidities did not predict the BOA post-TKR when adjusting for the effect of other covariables. The findings from the initial regression model must be interpreted with caution due to the low participants-to-predictors ratio. Moreover, only three subjects had had previous major surgery on the affected knee in the initial model. Although these patients showed a relatively low BOA post-TKR compared to the remaining patients, and therefore, a better gait, this finding may have been due to chance and future research including more participants with this characteristic should be carried out to corroborate the current result. It seems plausible to think that the subjects included in this study, who had had corrective knee surgery before TKR (i.e., HTO or patella realignment surgery) may have been advantaged from a biomechanical point of view, due to having had a lower limb realignment that allowed them to walk with a “better” pattern for many years. It has been demonstrated that HTO is successful at reducing the load on the medial compartment of the knee and improving gait (Whatling et al. 2020) even after five years after the surgery (Birmingham et al.

2017), and only about 21% of patients have a TKR ten years after an HTO (Primeau et al. 2021).

A final model including only three predictors (BOA pre-TKR, sex, and BMI) and forty-three patients was developed which confirmed the results of the initial model. The proportion of variance explained did not change drastically from the initial (53.8%) to the final model (56%), although it is preferable to have a simpler model, when possible, to explain an outcome. Only the gait function pre-TKR contributed significantly to predicting the gait function one-year post-TKR, even when controlling for BMI and sex. The results regarding sex not being relevant in predicting TKR outcome should be considered in the contest of the small proportion of women included in the model (31.8%). The small number of women in this study is unusual, considering that the number of females undergoing TKR in the UK (56.2%) is higher than males, as reported by the NJR (Brittain et al. 2022). Nevertheless, the number of males willing to enrol in the TKR study has routinely been higher than the number of females over the years, but the reasons for this are unknown. The 95% confidence interval for sex, was reduced from the initial to the final model, indicating that the differences between sexes were even smaller when adding more females. Therefore, it is suggested that sex may not be a significant predictor of gait function one-year post-TKR, but re-running the analysis with a larger number of females in the model would be advisable, to increase the external validity of the findings.

As highlighted by other authors (Devasenapathy et al. 2019), limited literature exists on predicting the objective outcome of TKR and the present investigation contributes to the existing knowledge, confirming some earlier findings of similar research. The current results are in line with previous reports when considering the ability of a pre-surgical function to predict the objective function post-TKR. However, to the best of the author's knowledge, this is the first study investigating a wide range of predictors of overall gait function one-year post-TKR, from a lower limb biomechanics perspective, utilising a regression model and a larger sample. Several studies found that having a lower performance-based test score pre-surgery (i.e., the time needed to carry out a test), was predictive of a longer time to complete the task post-TKR, indicating a worse outcome (Kennedy et al. 2006; Gandhi et al. 2009b; Stevens-Lapsley et al. 2010; Bade et al. 2014; Lee et al. 2020). Kluge et al. (2018) utilised portable sensors to evaluate an improvement in spatiotemporal gait

parameters and found that short strides length and time pre-surgery were predictive of patients who did not improve one-year post-TKR.

Patients suffering from knee OA, have usually had the condition for several years; maladaptive movement strategies, or compensatory movements, tend to develop in the presence of pain, and subjects suffering from knee OA may have had a certain maladaptive movement pattern for many years before undergoing TKR. A wealth of research showed that lower limb biomechanics do not go back to normal one-year post-TKR, despite improved pain levels, completing rehabilitation (Yoshida et al. 2008; Naili et al. 2017b; Zeni et al. 2019), and strengthening the quadriceps muscles whose weakness had been correlated to maintaining an altered gait (Yoshida et al. 2008). The fact that the objective function pre-TKR was correlated to the function post-TKR in the current study and previous research, may be the result of patients maintaining a certain acquired, and habitual motor behaviour that has not been modified with specific training. This idea seems to be supported by other researchers (Christensen et al. 2021b) and recently, a randomised controlled trial protocol was published, relative to a study that will explore the merit of “movement training” post-TKR in aiding a return to normal movement strategies (Bade et al. 2020).

A previous study utilising the Cardiff classifier on a small cohort of patients pre and one-year post-TKR (n = 15 patients) looked at differences between patients who moved towards the non-pathological side of the classifier (i.e., a good outcome, n = 8) and those who did not (i.e., a poor outcome) (Metcalfe et al. 2017). As in the current study, age and BMI did not differ significantly between those with a good (low BOA value) and poor outcome (high BOA value) (Metcalfe et al. 2017). Additionally, it was observed that those who had a non-pathological gait post-op had a significantly lower BOA pre-surgery than those who remained on the OA side of the classifier (Metcalfe et al. 2017). It must be noted that the outcome explored in the study of Metcalfe et al. (2017) was different from the current one (dichotomous: good/poor function VS. continuous outcome: BOA post-TKR), moreover, the authors looked at differences between groups rather than correlations. Nevertheless, the findings of Metcalfe et al. (2017) agree with the results of the current investigation, showing that those with a lower BOA pre-TKR also had better gait function post-surgery. Another study looked at factors correlating with the BOA post-TKR (Watling

2013), and there was a trend between pre-surgery BOA and poor outcome post-TKR (although no statistical test was run); the study additionally found that while age correlated with the gait function one year post-TKR, BMI did not; it was not clear, however, what was the statistical significance of these findings, as this information was not reported (Watling 2013); the sample size was small (n = 12), and, because of this, results should be treated cautiously, and the fact that age was correlated with the objective function post-TKR may have been due to a type II error (i.e., failing to reject the null hypothesis) (Portney and Watkins 2014).

Of the abovementioned studies looking at predicting PBTs following TKR, only two investigated the role of sex as one of the independent variables and found that being male was predictive of a longer distance covered during the six-minute walking test after surgery (Kennedy et al. 2006; Bade et al. 2014); however, females naturally tend to walk with shorter strides (Rowe et al. 2021) and this may justify the findings.

A high BMI is often a limiting factor in undergoing TKR, due to the higher risks of complications that may arise in individuals with severe obesity (D'Apuzzo et al. 2015; George et al. 2017). The current study suggests that when considering the objective gait function post-surgery, a higher BMI (within the safety range for risk of serious complications), should not be a deterrent in performing TKR as it does not seem to relate to the outcome when controlling for the baseline gait function and sex. Only a few studies looked at the effect of BMI on the objective outcome; two found that BMI was not a significant predictor of function (TUG, stair climbing test and six-minute walking test) one-month post-TKR (Lee et al. 2020) and one, three and six months post-TKR (Stevens-Lapsley et al. 2011a) in eighty-four, and one hundred forty patients, respectively; one study including one hundred fifty-five patients observed that a lower BMI was predictive of better TUG and stair climbing test one and two years post-TKR (Zeni and Snyder-Mackler 2010a). These studies had relatively large sample sizes, which would reduce the risk of a type II error (Portney and Watkins 2014), but also only looked at outcomes relatively early in the recovery phase and only one study evaluated the TKR outcome one-year post-TKR (Zeni and Snyder-Mackler 2010a). Nevertheless, the outcome evaluated was different from the current study, which considered the overall biomechanics of the lower limb, rather than tests measuring time to complete a task, which are not indicative of movement

quality. The relationship between the gait function examined via the Cardiff classifier and the PBTs has never been explored. Therefore, it is not possible to conclude how the discordant findings of the previous literature including such tests, relate to the current results. Therefore, chapter 6 will investigate if there is a correspondence between the PBTs and the Cardiff classifier, to understand whether PBTs may indirectly give information on movement quality.

Nowadays, TKR prostheses have a long life span of around twenty-five years (Evans et al. 2019), meaning that having surgery earlier in life may not entail needing a revision. Yet, TKR surgery is frequently postponed to prevent the need for a revision later in life, as studies on large cohorts showed that having surgery before the age of 55 had a 2% higher revision rate than those having surgery after the age of 65 (Julin et al. 2010). Nevertheless, previous literature found that delaying the TKR, decreased the quality-adjusted life years (i.e., the years spent in full health (NICE 2023)), and was not cost-effective (George et al. 2021). A previous qualitative study assessed the factors affecting the shared decision-making process of undergoing TKR and found that some patients believe they need to have severe mobility limitations and pain before they can have surgery (O'Neill et al. 2007). Following the findings of the current study, which are supported by previous literature discussed earlier, it is not advisable to wait for the knee function, and therefore the overall lower limb function, to be severely affected to undergo TKR, as this may result in a worse outcome, at least in terms of the objective gait function. Nevertheless, it remains challenging to define what would be the optimal timing to perform TKR to maximise the biomechanical outcomes, and pain, while minimising the revision rates.

5.4.3 Association between knee ROM and objective function

The last aim of the study was to determine whether there was a relationship between the knee ROM pre and post-TKR, and its change post-surgery with the BOA pre and post-TKR, respectively, and its change. The baseline knee ROM showed that there were movement limitations, and results were in line with previous reports, where patients had a knee ROM around 100 degrees pre-surgery (85.0 ± 16.9°) (Brown et al. 2009b), (111°) (Pua et al. 2019) (113.7°) (Mizner et al. 2011). The current patients' knee ROM was slightly smaller than what was found in other

studies one-year post-TKR (111.1 ± 13.9°) (Standifird et al. 2016) (117 ± 15°) (McClelland et al. 2017) but was similar to a study looking at knee ROM five years post-TKR (97 ± 16°) (Maempel et al. 2016). Comparing the results of previous literature, it may be inferred that knee ROM may vary across different patient groups, and it tends to be around 100° pre-TKR. The knee ROM limitations pre-TKR did not improve significantly post-TKR in the present investigation. This finding is consistent with previous literature where the knee ROM was similar between pre and one-year post-TKR, although there was a clinically meaningful improvement (less than 10 degrees improvement, on a mean) (Zeni and Snyder-Mackler 2010b; Mizner et al. 2011; Yoshida et al. 2012).

Hypothesis 3.1 – Pre-surgery, knee ROM is associated with BOA; *Hypothesis 3.2* – Post-surgery, knee ROM is associated with BOA; *Hypothesis 3.3* – From pre- to post-surgery, the change in knee ROM is associated with the change in BOA

In line with hypothesis 3.1, it was found that a larger knee ROM pre-TKR was moderately associated with a lower BOA before the surgery, hence a better gait function. Previous evidence also found that knee ROM pre-TKR had predictive abilities over function, namely the 30sCT, 6MWT and time to ascend or descend stairs in end-stage knee OA (Brown et al. 2009b). The patients in the current study experienced a non-significant improvement in the knee ROM post-TKR, while improvements were found in the gait function, as shown by the BOA values. This may justify why there was no relationship between the knee ROM and the BOA post-TKR or between the knee ROM change and the BOA change, as opposed to the initial hypotheses.

Hypothesis 3.4 – Knee ROM pre-TKR is associated with the BOA post-TKR

A greater preoperative knee ROM was found to be associated with better gait function one-year post-TKR. This is in contrast with previous research, where the baseline knee ROM had no predictive properties over the TUG or the SCT in patients one-year post-TKR (Mizner et al. 2005; Zeni and Snyder-Mackler 2010a). Nevertheless, the methods were different from the current investigation, where the function was measured via the Cardiff classifier, rather than the time needed to complete a task (Mizner et al. 2005; Zeni and Snyder-Mackler 2010a), and this may

justify the differences from existing literature. As previously mentioned, it is not clear how the findings from previous literature may relate to the current study, as PBTs measure the time to complete a task and not the movement quality (Naili et al. 2017b). It is suggested that preoperative knee ROM may correlate with the objective functional outcome post-TKR but the current evidence is scarce and difficult to contextualise.

The current study had several limitations, which should be considered when interpreting the results. The sample size was relatively small for a prediction analysis to be able to inform the clinical practice; this is a recurring problem in studies employing motion capture technology, which is the gold standard to measure biomechanics, but that requires a high level of expertise and is very time-consuming. Additionally, some factors that could plausibly aid in predicting the objective function post-TKR were not explored due to the lack of data, such as the rehabilitation process, the level of physical activity, the duration of symptoms before the surgery, the presence of depression and anxiety which may influence the engagement with rehabilitation and physical activity. The Cardiff classifier BOA value is not readily understandable on its own by a patient or a clinician. Perhaps, looking at the Cardiff classifier simplex plot may be more informative for patients and clinicians for interpretation purposes. Following this, it may be easier to discuss whether it is possible to predict what are the factors that play a role in determining a relevant clinical improvement, rather than discussing in terms of a continuous variable (BOA value).

5.5 CONCLUSION

This is the first study showing that utilising a mixed-age control group within the Cardiff classifier results in underestimating the patients' gait function both pre and one-year post-TKR. This information can be taken into account in future studies utilising the Cardiff classifier, and it is suggested that when evaluating the absolute gait function pre- and post-surgery, a control group that is age-matched or has a similar age to the patients should be employed. However, when evaluating the change in gait function, a mixed-age control group may be utilised (provided there is

a high proportion of controls over the age of fifty), since the reduction in the BOA is comparable to that obtained when the control group has a similar age to the patients.

Previous studies looking at predictors of the gait function post-TKR utilising the Cardiff classifier adopted correlation analyses, which do not inform on the predictive ability of the variables of interest, or failed to report the significance of the findings, were conducted on small samples, and included only a few potential predictors. The existing literature looking at predicting objective outcomes of function (i.e., PBTs), is very limited. Therefore, the current results make a novel contribution to the limited existing knowledge on predicting objective gait function. It was found that no relationship existed between the change in gait function (i.e., BOA change) and anthropometrics, demographics, comorbidities, surgical factors, issues in other joints, knee ROM change, baseline gait function and pain. While it could be suggested that anthropometrics, demographics, surgical factors, the presence of issues in other joints, having had previous surgeries on lower limb joints and baseline function may not play a role in the amount of functional improvement post-TKR, it is recommended that an investigation on larger patients numbers should be carried out due to the high variability in the gait change observed in this study and the difficulty in identifying a clear pattern.

The main finding was that the objective function pre-surgery is significantly predictive of the gait function one-year post-TKR, even when controlling for sex and BMI. Therefore, waiting for the objective function to deteriorate severely before performing a TKR is not advised, as in the current patient group, having a worse function pre-TKR resulted in a poorer function post-TKR. It was also found that sex and BMI were not predictive of the gait function post-TKR and that no correlation existed between function post-TKR (BOA post-TKR) and a wide range of factors (age, issues and previous surgeries in other joints, comorbidities, implant design and patella resurfacing, post-surgery complications, previous surgeries on the affected knee, baseline PROMs), most of which had never been explored before when employing the Cardiff classifier to assess objective function (i.e. issues and previous surgeries in other joints, comorbidities, patella resurfacing, post-surgery complications, previous surgeries on the affected knee).

It is not clear how the current findings may relate to the previous research looking at predicting PBTs, as the relationship between the objective function

measured via the Cardiff classifier and PBTs has never been investigated. Therefore, this will be explored in the following chapter.

Chapter 6: Relationship between objective function, performance-based tests and patient-reported outcomes pre- and post-TKR

6.1 INTRODUCTION

The study contained in this chapter follows an ethics amendment of the TKR study research protocol, which introduced the PBTs suggested by OARSI. The OARSI PBTs are recommended to assess objective function of people with OA and post-TKR (Dobson et al. 2013). This study explores the PBT measurement properties against the Cardiff classifier. Within the protocol amendment, the marker set utilised thus far was implemented to allow measurement of trunk kinematics, which was previously found to be different between people with severe knee OA and those with no pathology (NP) (Hunt et al. 2010; Bechard et al. 2012; Creaby et al. 2012; Turcot et al. 2013). Alterations may be a consequence of the knee disease; there is evidence that trunk lean towards the affected limb may increase with the advancement of OA severity, and it is moderately associated with knee pain (Bechard et al. 2012). A recent systematic review (Iijima et al. 2019) found that trunk lean towards the ipsilateral limb was significantly larger in people with varying degrees of knee OA, compared to NPs of similar age. It was suggested that trunk lean, and flexion may reduce external knee moments in people with knee OA (Hunt et al. 2010). An experimentally increased trunk lean was found to reduce the knee adduction moment in healthy individuals (Hunt et al. 2011), and people with mild and moderate knee OA (Simic et al. 2012). Only a few studies have looked at trunk kinematics alterations in the sagittal plane in people with knee OA and found an increase in trunk flexion (Preece and Alghamdi 2021), but differences from controls were present only in patients with a varus knee alignment (Turcot et al. 2013). Evidence showed that this strategy may result in increasing external hip and ankle sagittal moments, and thigh muscle co-contractions (Preece and Alghamdi 2021). The research looking at trunk kinematics in people with end-stage knee OA is limited (Hunt et al. 2010; Sagawa et al. 2013; Turcot et al. 2013), and mainly discrete values (Hunt et al. 2010; Sagawa et al. 2013) or trunk ROM (Turcot et al. 2013) were reported, leaving a gap in knowledge on whether alterations may exist in the overall trunk flexion-extension and lateral lean patterns. It was suggested that more

research is needed to aid in understanding altered trunk kinematics as a possible distinctive characteristic of people with knee OA (Iijima et al. 2019).

Trunk kinematic features during gait have not been employed within the Cardiff classifier thus far, and this novel exploration stemmed from the following research question: *are features related to trunk lean and trunk flexion-extension during gait (i.e., trunk kinematics in the frontal and sagittal planes, respectively) included in the eighteen highest-ranking features discriminating between NPs and patients with severe knee OA?*

Aim 1: to determine whether trunk kinematics features are relevant for discriminating individuals with knee OA from NPs within the Cardiff classifier.

This would be useful to future researchers aiming to further develop the Cardiff classifier method. Moreover, this information would contribute to the existing limited knowledge on whether trunk kinematics may be a relevant variable to consider in people with severe knee OA.

The concept of physical function is elaborate and would fall under the item “Mobility” within “Activities” in the World Health Organization International Classification of Functioning (WHO 2023); it encompasses several activities of daily living such as walking, changing position, moving around, making it complex to assess. There is a degree of difficulty in defining an adequate test to evaluate physical function. It is not advisable to rely solely on the patient’s perception of it, which can be influenced by various factors, as presented in the Literature Review (Chapter 2). There is no gold standard to assess physical function to date, and in 2013, OARSI recommended the use of PBTs in the evaluation of objective function in people with OA and undergoing a joint replacement procedure (Dobson et al. 2013). The methods proposed by OARSI include a minimum battery of three PBTs which measure maximum walking speed (40m walk test), the time for negotiating stairs (SCT), or the number of sit-to-stand repetitions within 30s (30sCT), and two additional tests, recording the time to accomplish a walking, standing, changing direction and sitting (TUG), and the maximum distance covered in six minutes (six-minute walk test) (Dobson et al. 2013). The measuring properties of the proposed PBTs have been explored in various studies (Dobson et al. 2012; Naylor et al. 2014; Alghadir et al. 2015; Huber et al. 2016; Dobson et al. 2017; Mehta et al. 2019;

Mostafaee et al. 2022; Sharma et al. 2023). A recent review (Coleman et al. 2020), updating the earlier review of the evidence (Bennell et al. 2011), concluded that all the tests proposed by OARSI are appropriate for use in individuals with knee OA, as they are valid in measuring function, responsive to change, and have good reliability.

It is not clear whether the gait biomechanics measured by means of the BOA calculated via the Cardiff classifier may be associated with how well a subject would perform during a PBT pre- and post-TKR. Bolink et al. (2015a) proposed that patients with knee OA may complete a physical function test quicker post-TKR, even though limping and instability may still be present. Nevertheless, a previous study found that a lower GDI-kinetic, another measure of gait quality, was moderately correlated with longer times to complete the TUG, and five-time sit-to-stand pre-TKR, suggesting that the more severely joint loading is altered, relative to a reference group, the slower an individual performs during the tests (Naili et al. 2017a). Yet, no correlation was found between the GDI and the PBTs, indicating that the degree of gait kinematics impairment was not related to how fast TUG and the five-time sit-to-stand were completed (Naili et al. 2017a). The study of Naili et al. (2017a) analysed the correlations between gait quality and only one of the core tests suggested by OARSI (Dobson et al. 2013), and only pre-TKR, leaving a gap in knowledge of whether gait quality may be related to other PBTs, and if the relationship is similar post-TKR. Hence, another aim of this chapter was to expand the limited existing knowledge regarding the relationship between two types of objective measures of function. The findings may guide the development of future research focussed on measuring objective outcomes (e.g., for power calculation purposes). Utilising appropriate subjective and objective measures of knee function can aid the choice of treatments in the presence of knee OA, can help better define the outcome of TKR surgery, and may provide information to refine the rehabilitation strategies following TKR.

Since both the BOA and PBTs are measures of objective physical function, to avoid confusion between PBTs and BOA, while previously “objective function” was used as synonym for the BOA, from now onwards “gait quality” will correspond to BOA within this chapter.

Aim 2: to explore whether gait quality (i.e., BOA) correlates with the three core PBTs (30sCT, 40mFPWT, SCT) and the TUG proposed by OARSI.
Hypothesis 1 - At each time point in the study, a moderate correlation exists, in the following order of decreasing strength, between:

1. BOA and the 40mFPWT (negative relationship)
2. BOA and TUG (positive association)
3. BOA and SCT (positive correlation)
4. BOA and 30sCT (negative relationship)

The hypothesised strength of the correlation was based on the concept that the BOA and the 40mFPWT measure a similar activity (i.e., gait), the TUG had elements of gait in it, while the SCT and 30sCT do not.

The sample size of the current study was small pre-TKR, and especially post-TKR due to various challenges during data collection and the limited timeframe of the project, therefore, this study was exploratory in nature.

Earlier work in our research group showed a moderate correlation between PROMs including the OKS (which is routinely utilised in the NHS to evaluate TKR outcomes), the Knee Outcome Survey, and the BOA at baseline and approximately one-year post-TKR (Biggs et al. 2019a). Moreover, a strong association was found between the pre to post-TKR change in the abovementioned PROMs, and the change in BOA (Biggs et al. 2019a). Considering the poor correlation between PROMs and PBTs found in previous literature (Stratford et al. 2010; Mizner et al. 2011; Hamilton et al. 2012b; Graff et al. 2016), the authors (Biggs et al. 2019a) proposed that gait quality may be more closely related to how patients perceive their function rather than the timing to complete a task as in the PBTs, but this aspect has not been explored. Nevertheless, the evidence showing an association between PROMs and BOA is limited to the study of (Biggs et al. 2019a), which developed from Dr Biggs' thesis (Biggs 2016), and more research is needed to corroborate this finding.

Aim 3: to explore the relationship between BOA and PROMs.
Hypothesis 2.1 – A moderate, negative correlation exists between BOA and OKS pre and at three, six and twelve months post-TKR.

The KOOS is another knee-specific measure widely used in research (Vajapey et al. 2020), and it was decided to explore the association between BOA and perceived function (KOOS ADL) and pain (KOOS Pain) separately, rather than assessing these aspects together, as in the OKS.

Hypothesis 2.2 - A moderate, negative relationship exists between KOOS ADL, KOOS Pain and BOA pre and at three, six and twelve months post-TKR.

PBTs were previously found to be poorly correlated with PROMs, with the suggestion that these measures assess different constructs of physical function (Stratford et al. 2010; Mizner et al. 2011; Hamilton et al. 2012b; Graff et al. 2016). Moreover, an issue in comparing objective function with PROMs is that the latter investigate the degree of difficulty in carrying out a task such as walking, standing up from a chair, or turning in bed, which is strictly dependent on the patient's perception, which seems to be mostly influenced by pain rather than what can be achieved objectively (Boonstra et al. 2008; Mizner et al. 2011). To contextualise the current findings within the existing literature, and to aid understanding of whether the PBTs scores may have been associated with pain levels (KOOS Pain and OKS), the following aim was developed:

Aim 4: to explore the relationship between the PBTs included in the present study and PROMs utilised in the NHS and research. Given the existing evidence (Stratford et al. 2010; Mizner et al. 2011; Hamilton et al. 2012b; Graff et al. 2016) discussed earlier, the following hypothesis was developed.

Hypothesis 3.1 – A significant, weak correlation exists between the OKS and the 40mFPWT, the 30sCT, the SCT, the TUG.

Hypothesis 3.2 – A significant, weak correlation exists between the KOOS ADL/Pain and the 40mFPWT, the 30sCT, the SCT, the TUG.

6.2 METHODS

6.2.1 Participants

For this longitudinal, prospective, observational study, the participants' recruitment process, and inclusion and exclusion criteria were described in section 3.3. Patients listed for a primary TKR were approached at CAVOC by a Research Officer between November 2019 and September 2022, although the recruitment was halted in March 2020 and effectively restarted in February 2022 due to COVID-19-related face-to-face research restrictions, elective surgeries being paused and challenges in restarting the recruitment. Research Officers at CAVOC had restricted access to the hospital green zone areas to recruit patients for several months, and their duties were repurposed to give priority to COVID-19-related research. The initial aim was to collect data on thirty patients pre-TKR, to account for about 15% loss at follow-up, which was based on the ratio of patients interrupting their participation to patients completing the full study before recruitment for the current study started. Forty-nine patients signed a permission to contact and were contacted over the phone or via email by the main researcher or the research officer based at the MSKBRF and of these, eighteen agreed to enrol. Six, five and six patients returned for their follow-up at three, six and approximately twelve+ months after the surgery, respectively. Figure 27 shows the study flow. For this study, the initial aim was to collect control data on twenty-five people over the age of sixty, to have a similar age to the patient's group, to allow an optimal classification, given the results of Chapter 5. However, due to the challenges during the recruitment process (COVID-19 lockdown, and issues in finding individuals matching the inclusion-exclusion criteria), the data collection was changed to include a mixed-age NP group. The recruitment took place between January 2020 and December 2022. Sixteen NP subjects expressed an interest in participating in the study and were contacted to assess their eligibility. Two were not contactable, three were not eligible and two were not interested. Therefore, nine NPs were recruited. All participants signed a consent form before the collection of the data.

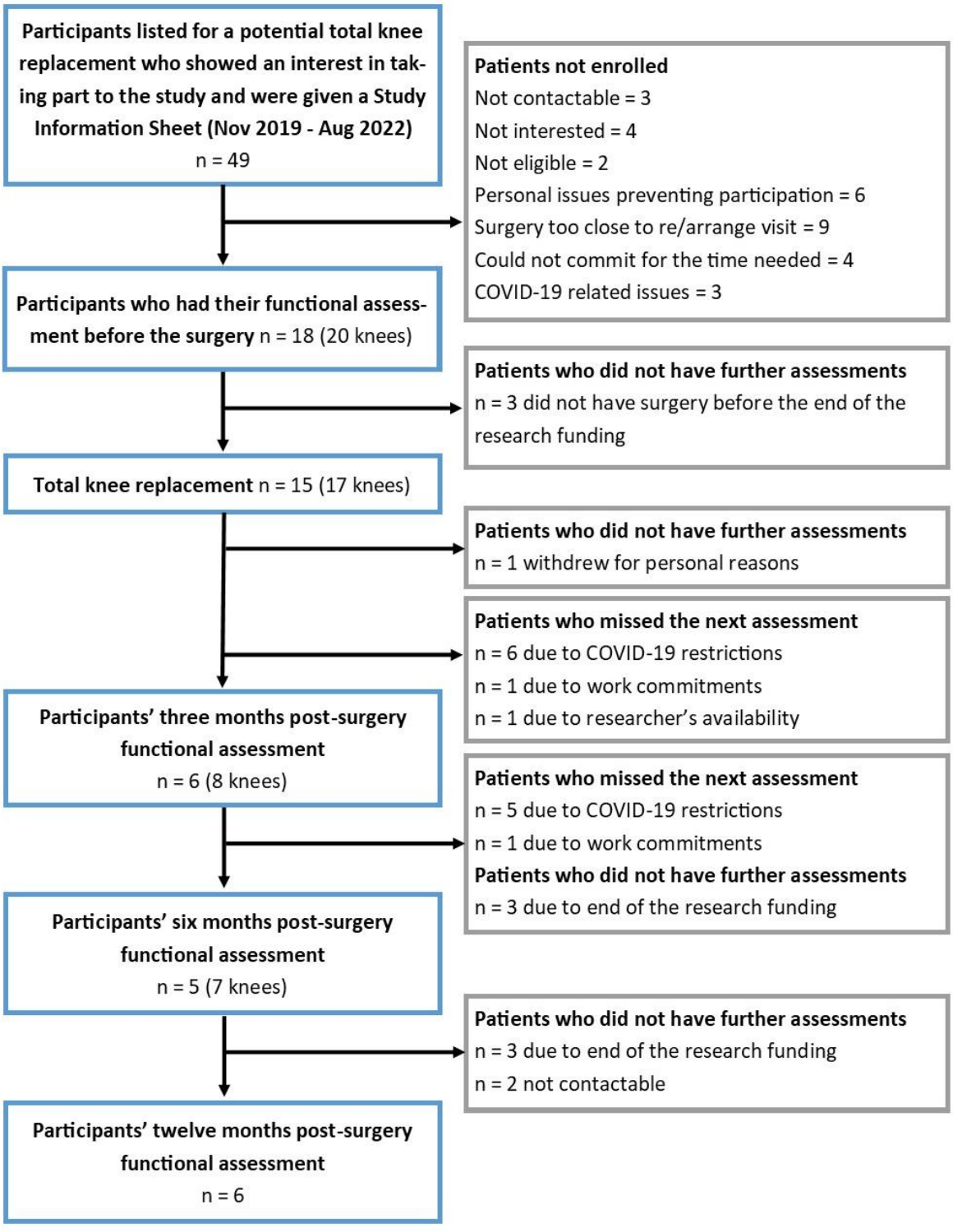


Figure 27 Patients study flow.

6.2.2 Data collection

PBTs were introduced following an amendment of the BBRCVA Centre Research Protocol, which was submitted in October 2018 and approved on 25th June 2019. All data were collected within the MSKBRF Clinical Lab by the current author with the technologies and methods described in section 3.3. Patients' data were collected pre-surgery and at three, six and twelve months post-TKR, while NP data was collected on one occasion. At each session, volunteers filled out the OKS (for both knees, but only the score for the knee receiving surgery was considered for further analysis) and KOOS (for the knee undergoing surgery only, in case of both knees receiving surgery, patients responded separately for each knee), the past, present medical history was recorded, as well as the rehabilitation at follow-up (duration, type: group/one-to-one/mixed).

All participants performed a 3D gait analysis. The marker set utilised was similar to the one described in section 3.3, consisting of a CAST marker set (Cappozzo et al. 1996a) with an additional marker on each iliac crest, and included six additional markers. For this study, the supplementary markers were placed on the thorax and shoulders to allow tracking of the trunk kinematics, for a total of twenty-eight markers and four marker clusters (Figure 28). The marker set suggested by the ISB for tracking the thorax kinematics encompasses markers on the seventh cervical vertebra's spinous process (C7), the spinal process of the eight thoracic vertebra (T8), the incisura jugularis and the xiphoid process (Wu et al. 2005). When implementing the new protocol, it was decided not to place a marker on the xiphoid process during data collection as this would have been difficult with overweight subjects and especially with women. The xiphoid process lies in the area where women's bra usually rests on the chest. It was deemed inappropriate to ask participants to expose this area for data collection. This is a common concern that was discussed by other researchers (Armand et al. 2014), who explored alternative marker sets for tracking trunk motion in ten subjects (50% females) during gait while performing large movements of the trunk and head in the three planes. The authors (Armand et al. 2014) tested fifty-two combinations of sixteen markers on the thorax and found that the mean error in marker positions was smaller for two marker sets, one including jugular notch, xiphoid process, and T8 (or nearing markers) (root mean square error 9.2 mm), one including jugular notch, C7 (or the spinous process of the

second thoracic vertebra (T2)) and T8 (or nearing markers) (root mean square error 9.7 mm). For the present study, the latter combination of markers was utilised (jugular notch, T2 and T8); T2 was chosen in place of C7 to track the movements of the thorax, following the recommendations from (Armand et al. 2014), since it was found that the orientation of C7 was affected by the head's movements and that T2 was a better landmark to track trunk movements. An additional marker was placed at about 50% distance between T2 and T8 (approximately at the level of the fifth thoracic vertebra's spinal process, T5) to offer an additional marker to track the segment, in case of potential marker occlusions. Participants were allowed to wear a vest top, which was lifted to allow visibility of T8. The placement of T5 may have been challenging in the presence of the vest top, and in this case, the top was secured with medical-grade double-sided tape to the back of the participant, and the T5 marker was placed on the vest's fabric. This allowed the marker to move with the thorax, avoiding artefacts due to the fabric movements relative to the thorax. Some male participants offered to remove the vest top as they felt comfortable doing so, in these cases the thorax markers were placed directly on the skin.

Following the gait assessment, participants were evaluated for OARSI PBTs (Dobson et al. 2013). The tests were performed in the following order and included the 30sCT, TUG, 40m (10 m x 4) fast-paced walk test, and SCT. 3D motion capture was utilised to record the activity during the tests, but the related biomechanical data were not utilised within the current study. The six-minute walk test is not included in the minimum set of core tests suggested by OARSI (i.e., 30sCT, 40m fast-paced walking test, SCT) and it was not incorporated in this study. During data collection, patients were required to walk several times for the biomechanical assessment at the beginning of the session, and adding more walking to the assessment day was believed to be inappropriate, given the knee joint pain and walking difficulties that patients with end-stage knee OA commonly present.

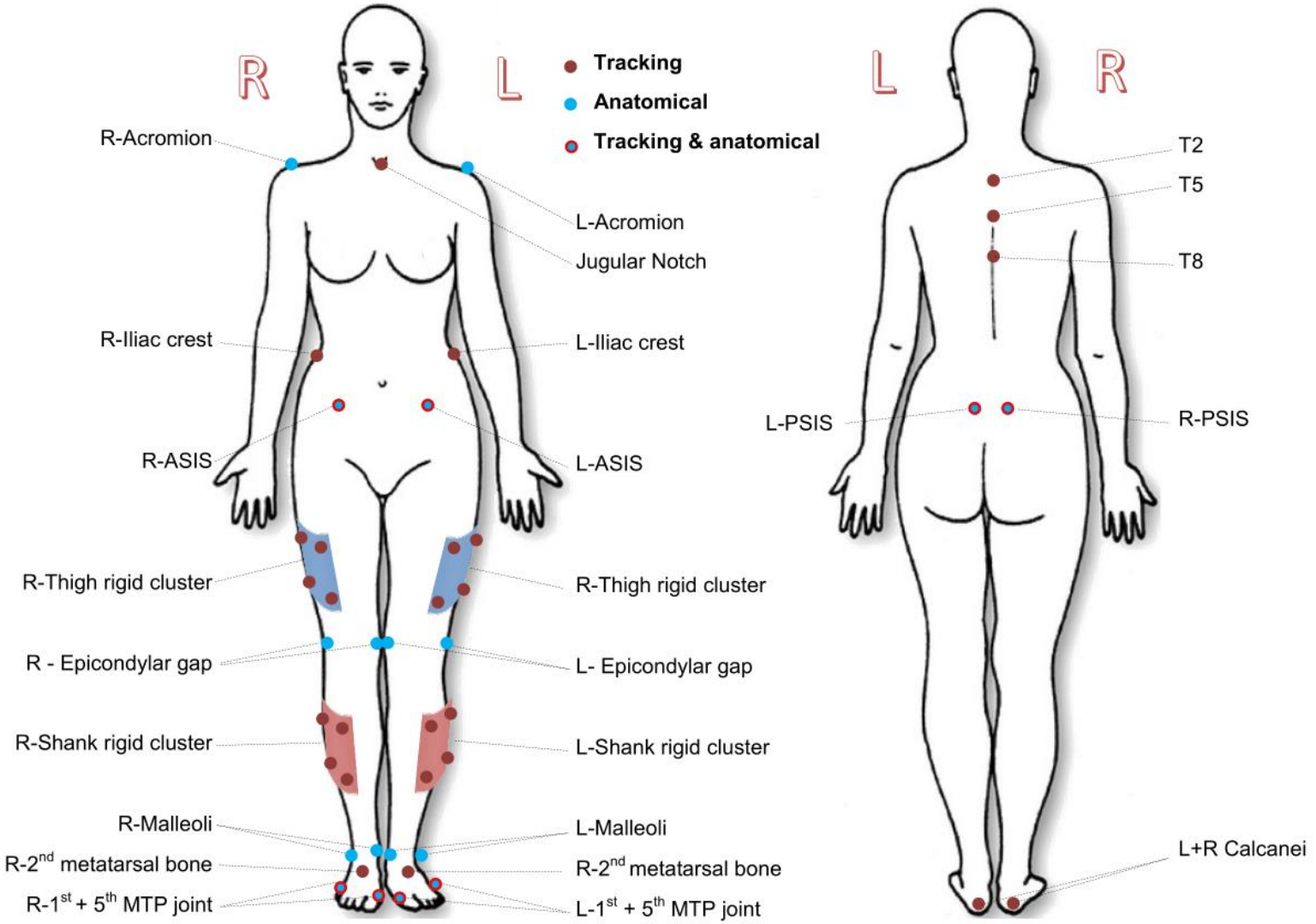


Figure 28 Marker set. R: right; L; left; MTP: metatarsophalangeal; T: thoracic vertebra; ASIS: anterior superior iliac spine; PSIS: posterior superior iliac spine.

Before each test, the researcher explained the procedure to the participant and ensured that they were comfortable carrying out the task. If this was the case, the researcher then read the standardised instructions on how to perform the test (OARSI 2023) and gave a demonstration. Participants were allowed to familiarise themselves with the test. All participants were assessed barefoot across the PBTs, to allow standardisation of the tests, as noted by OARSI (Coleman et al. 2020). Between tests, while the equipment was prepared for the following task, the participant was invited to rest on a chair for 1-2 minutes.

Performance-based tests

The equipment utilised, the test set up and the scoring for the PBTs followed the instructions in a handbook supplied by OARSI, which can be found online (OARSI 2023). A generic stopwatch application on the researcher's smartphone was utilised to track the time during the tests (to the nearest 10th of a second). The procedure and scoring for each test and the outcome are described in Table 27.

The 30 seconds chair test assesses sit-to-stand. The reliability of a test is deemed to be sufficient when the ICC > 0.8, and optimal when the ICC > 0.9 (Portney and Watkins 2014), and the 30 seconds chair test had excellent reliability (ICC 0.90) (Tolk et al. 2017) in people with severe knee OA. The intra-tester reliability was between optimal (ICC 0.97-0.98) (Gill and McBurney 2008) and sufficient (ICC: 0.85) (Dobson et al. 2017) in a mixed group of patients with hip and knee OA. The estimated minimal detectable change in patients with hip and knee OA (Dobson et al. 2017) and minimal important change in patients with severe knee OA undergoing rehabilitation (Mostafaee et al. 2022) was calculated to be approximately 2 repetitions. No information on the minimal detectable or important change could be found for individuals undergoing TKR. For this test, a plinth with a seat height of 44 cm from the floor was utilised. The plinth was used in place of a standard chair to allow full visibility of the markers during the task, and because the test required to be performed on a seat without armrests.

The TUG assesses "ambulatory transitions" (OARSI 2023); the test-retest reliability was found to be sufficient in people with hip and knee OA (ICC 0.81) (Dobson et al. 2017), and optimal in patients post-TKR (ICC 0.98) (Yuksel et al. 2017). The minimal detectable change was 2.27 s in forty-eight people who had a

TKR within six months from the assessment, but this was calculated by repeating the test on the same day, and not pre- to post-surgery (Yuksel et al. 2017). Another study found that the 95% confidence interval minimal detectable change was 36.7% in seventy-two people awaiting TKR (Naylor et al. 2014). In another study, the minimal detectable change was about one second shorter than what was found by Yuksel et al. (2017) post-TKR corresponded to 1.21 s (Dobson et al. 2017), this could be because the fifty-one patients suffering from knee or hip OA were examined one week apart and no intervention (either rehabilitation or surgery) was performed between the two examination time points. For the TUG, a chair with a seat height of 44 cm from the floor was utilised (Figure 29). The chair had the backrest removed leaving a metal bar backrest, to allow markers visibility during the test. This chair was used in place of the plinth because the test required to have availability of armrests and a backrest.

Figure 29 Modified chair utilised for the Timed up-and-go.

The 40m fast-paced walking test measures “walking short distances” (OARSI 2023); it had optimal test-retest reliability (10 m x 4, ICC 0.93) in patients undergoing TKR (Tolk et al. 2017) and with severe hip and knee OA (ICC 0.92) (Dobson et al. 2017). The minimal detectable change was 0.19 m/s in the presence of hip and knee OA (Dobson et al. 2017), while the minimal important change was 0.21 m/s in patients with knee OA post-rehabilitation (Mostafaee et al. 2022).

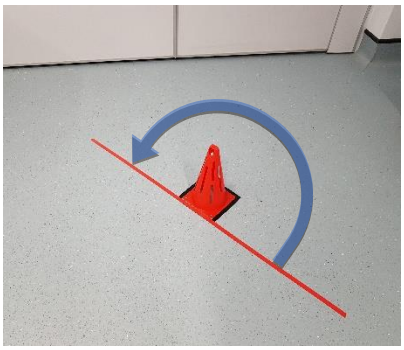


Figure 30 Black signage on the floor to guide cone placement and red lines to cross when turning around the cone. The curved arrow shows the path followed while turning around the cone, when the timing was paused. There was a minimum of 90 cm distance between the cone and the surrounding walls.

The SCT assesses stair negotiation (OARSI 2023) and it can be performed on stairs that have a varying number of steps, therefore, the research regarding this test is mixed, with studies reporting the test's measurement properties on four (Lin et al. 2001; Mostafaee et al. 2022), six (Sharma et al. 2023), nine (Kennedy et al. 2005), and twelve steps (Mizner et al. 2011). OARSI did not give specific recommendations on the optimal number of steps for the test (OARSI 2023). In the current study, a four-step staircase was utilised as it was readily available in the laboratory and because it can be easily found in most physiotherapy departments. The four-step Test had optimal intra-tester reliability, with an ICC between 0.94 (stair ascending) and 0.96 (stair descending) in people with hip and knee OA (Lin et al. 2001). There is very limited information on the minimal important or detectable change for this test (Coleman et al. 2020). The minimal important change (against a global rating of change scale indicating how much improvement/deterioration patients perceived) was tested on the four-step stair test in one study only (Mostafaee et al. 2022) and was found to be 3.21 s, one-month post-rehabilitation in people with knee OA. The minimal detectable change for the nine steps stair test was 5.49 s one month following TKR (Lin et al. 2001).

Table 27 Sequence of PBTs, the procedure for the test and its scoring, and outcome measure

Performance-based test	Equipment	Procedure (all tests performed once, unless stated otherwise)	Outcome
30 seconds chair test	A plinth, 44cm seat height from the floor	The participant started from a sitting position with arms across the chest and feet shoulder-width apart. On the signal to start (i.e., stopwatch started), the participant stood up to a fully standing position and sat down continuously as fast as possible, and then stopped at the 30s mark. If the volunteer stopped during the test, the timer was not paused. The participants could use their hands only if necessary.	Number of full chair stands within 30 seconds
Timed up and go	A chair, 44cm seat height from the floor and armrests 64cm from the floor; a line 3 m away from the chair's front legs	The participant started from a sitting position with the back resting on the backrest and arms on the armrests. On the signal to start (i.e., stopwatch started), the participant stood up, walked to a line 3m away at a self-selected speed, turned around 180°, walked back, and sat on the chair. The stopwatch was stopped when the participant's back touched the backrest. Two trials were recorded (with at least 30s break between trials), and only the fastest trial was considered.	Time to complete the task
40m fast-paced walking test	Two cones at least 90cm away from the walls, a 10m walkway across the laboratory (10m walked 4 times)	The participant started with their toes on the line, leaving the cone to their left. On the signal to start (i.e., stopwatch started), the volunteer walked as fast as possible, without running, to the second cone, turned around it, went back to the start line, turned around the first cone and repeated the procedure to cover in total 40m (2 laps). The stopwatch was paused when the patient crossed the line to turn around the cone and it was resumed when the participant's foot crossed the second line around the cone to walk back across the lab (Figure 30). The stopwatch stopped when the participant's foot crossed the 40m mark. Participants were allowed to utilise a walking aid if needed.	Walking speed in m/s: time to complete the task/distance
Stair climbing test	A staircase (4 steps: three 915mm wide x 305mm deep steps that are 160mm high, and a 915mm wide x 610mm deep landing) with a handrail on the left side	The participant stood in front of the stairs. On the signal to start, the participant ascended and descended the stairs once only, as fast as they could without running and in a safe manner. The use of the handrail was allowed, the participant could negotiate the stairs in their preferred manner (step by step, descending sideways, etc.) and if they needed to slow down or pause, the time kept going. The timing was stopped when the second heel touched the floor.	Time to complete the task

6.2.3 Data processing

Processing of the 3D gait data followed the steps described in section 3.6. The thorax model created was an adaptation of the model suggested by the ISB (Wu et al. 2005), based on (Armand et al. 2014):

- The origin of the segment was located at T2 rather than on the jugular notch. Moving the origin of the segment does not affect joint angles calculations for the trunk
- The z-axis (transverse plane rotations) was the line connecting T2 and T8 pointing upwards, rather than the line connecting the midpoint between the xiphoid process and T8 and the midpoint between the jugular notch and T2
- The x-axis (flexion-extension) was perpendicular to the plane created by the jugular notch, T2 and T8, rather than being perpendicular to the jugular notch, xiphoid process, T2 and T8.
- The y-axis (lateral lean) was the cross product between the z and y-axis, pointing forward

Figure 31 shows the thorax segment and the orientation of its coordinate system. The thorax was used for the calculation of joint kinematics only and was considered a rigid segment, i.e., not accounting for intervertebral movements. Calculations of thorax joint angles were relative to the virtual lab (as in the case of the pelvis) and had an X-Y-Z Cardan-Euler sequence (i.e., flexion-extension, lateral lean, rotation in the transverse plane).

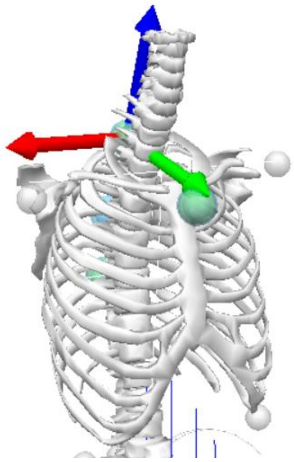


Figure 31 The orientation of the thorax coordinate system. Red: x-axis (flexion-extension); green: y-axis (lateral lean); blue: z-axis (rotations in the transverse plane).

6.2.4 Participants' classification of gait function

The reader is redirected to Chapter 3 (section 3.7) for details on the Cardiff classifier method. The gait features included in the classifier included the first three principal components related to the following variables: trunk and pelvis angles in all three planes, hip, knee, and ankle kinematics in the sagittal and frontal planes; hip, knee, ankle sagittal and frontal planes distal moments (expressed as external moments); hip, knee, ankle sagittal distal joint powers; vertical, anterior-posterior, mediolateral ground reaction force (GRF). Both lower limbs were considered for analysis in the NP group (n = 9), for a total of eighteen cases. Only the affected knee was considered in the patient group at each time point, resulting in twenty cases pre-TKR. The out-of-sample classification simplex plot was produced, and the out-of-sample classification accuracy was calculated as the proportion of out-of-sample subjects out of the total whose classification matched their class. The sensitivity (Eq. 27) and specificity (Eq. 28) of the classification based on the full dataset was reported.

The number of patients analysed at follow-up was smaller than that pre-surgery (Figure 27). The principal component scores for the patients post-TKR at each follow-up were calculated relative to the principal components initially defined for NPs and patients pre-TKR. The same eighteen highest-ranking gait features discriminating between patients pre-TKR and NPs were utilised in the Cardiff

classifier to obtain the Belief values only for the patients who attended their follow-up.

6.2.5 Statistical analyses

All analyses were performed in IBM SPSS Statistics (version 27, IBM, Armonk, NY, US). Normality distribution was determined via the Shapiro Wilk's test (p-value > 0.05: parametric distribution) and inspection of Q-Q plots. For all statistical tests, the significance level was $\alpha = 0.05$, unless otherwise specified. To compare baseline characteristics between patients and NP, continuous variables were examined with an independent t-test or its non-parametric equivalent (Mann-Whitney U test) and categorical variables (i.e., sex) with Pearson's Chi-Square test.

The differences in / correlations between the BOA values, the PROMs and the PBTs scores were tested separately pre-TKR, and at the three, six and twelve-month follow-up for the patients who attended the data collection. Specifically, the BOA, PBTs and PROMs scores were compared between the available patients post-TKR and the same patients pre-TKR only, therefore, not against the whole pre-TKR dataset, to allow for paired samples testing.

To verify if the patients' BOA value and PBTs scores changed significantly pre to post-TKR for the patients who attended their follow-up, a non-parametric test, the Wilcoxon signed-rank test was run, due to the small sample sizes at three, six and twelve months post-TKR. Results were reported as median values (IQR), but mean values were also described for comparative purposes with previous literature. Patients who attended the three-month follow-up were different from the ones participating in the six or twelve-month data collection, therefore, it was not possible to compare the patients' scores change between follow-ups.

Pearson's correlation was utilised for continuous variables which had a linear relationship and were normally distributed, or its non-parametric equivalent (Spearman's correlation) was employed for ordinal (i.e., KOOS sub-scores, and OKS) or non-normally distributed data. When multiple comparisons are carried out on a variable, there is the potential to increase the chance of a Type I error, or false positive (Portney and Watkins 2014). Hence, a Bonferroni correction was applied to the significance level to account for multiple correlations (Sedgwick 2014) (BOA correlated with four PBTs and three PROMs = 7 comparisons); considering $\alpha = 0.05$

for each statistical test before the adjustment, the Bonferroni-adjusted $\alpha = 0.05/7 = 0.007$ for the BOA versus PBTs and PROMs comparisons, while the Bonferroni-adjusted $\alpha = 0.05/4 = 0.013$ for the PBTs versus PROMs comparisons (accounting for the PBT-BOA comparison).

The strength of the correlation, where significant, was interpreted as suggested by Dancey & Reidy (2011), where r is the correlation coefficient (indicating both r_s, Pearson's correlation coefficient, and ρ: Spearman's correlation coefficient):

- $|r| < 0.3$ meant a weak correlation
- $0.4 < |r| < 0.6$ was regarded as a moderate correlation
- $|r| > 0.7$ was interpreted as a strong correlation

6.3 RESULTS

6.3.1 Participants characteristics

Table 28 shows that patients were significantly older, heavier, and with a higher BMI than NPs. While patients were between the age of fifty-eight and seventy-nine, with most of them being in the 60-69 years old range, 44.5% of NPs were between the age of twenty and forty-nine and 55.5% between the age of sixty and seventy-nine. Moreover, there were significantly more females in the NP group (77.8%), than in the patients' group (33.3%).

Table 28 Participants' characteristics

Participants characteristics	Non-pathological volunteers (n = 9)	Patients pre-TKR (n = 18)	Significance level
Age (years), mean (SD), *range*	53.0 (23.7) *20 - 79*	68.0 (6.0) *58 - 79*	0.017*
20-29, n (%)	2 (22.3)	0	
30-39, n (%)	1 (11.1)	0	
40-49, n (%)	1 (11.1)	0	
50-59, n (%)	0	1 (5.6)	
60-69, n (%)	1 (11.1)	10 (55.5)	
70-79, n (%)	4 (44.4)	7 (38.9)	
Height (m), mean (SD)	1.64 (0.07)	1.69 (0.07)	0.139
Weight (kg), mean (SD)	64.5 (12.4)	87.6 (17.4)	0.002*
BMI (kg/m^2), mean (SD), *range*	23.8 (3.7) *19.0 – 30.5*	30.9 (6.3) *21.3 – 42.4*	0.005*
Females, n (% within group)	7 (77.8)	6 (33.3)	0.029*

TKR: Total Knee Replacement; SD: standard deviation; m: meters; kg: kilograms; BMI: Body Mass Index; n: number;
*: statistically significant at $p < 0.05$.

Five surgeons performed the TKRs. Three patients did not have the TKR surgery before the end of the study, one patient withdrew before the clinical data could be retrieved; hence, surgical and medical notes were accessible for 14 patients (77.7% of the total). Of these, the median ASA was 2 (0 IQR), where three patients had an ASA 1, ten patients had an ASA 2, and three had an ASA 3. Two out of fourteen subjects did not have their patella resurfaced. For one patient, the implant design information was not retrievable due to the surgery being performed at a private hospital. Therefore, the implant design information was available for 13

patients. Of these, two received a posterior-cruciate-sacrificing prosthesis, and eleven a posterior-cruciate-retaining knee implant. Information for the surgical approach was available only for 11 patients, and in all cases, it was a medial parapatellar approach. Patients who were given a surgery date before the end of the study (n = 15) were assessed, on a mean, 4.6 months before their surgery (range: 0 to 20 months). Some patients who had their surgery scheduled at the time of the study assessment, had it delayed due to the COVID-19 pandemic or COVID-19-related issues.

6.3.2 Cardiff classifier including trunk kinematics

The Cardiff classifier showed a 92.1% out-of-sample classification accuracy in classifying participants (n = 3 misclassified subjects) (Figure 32).

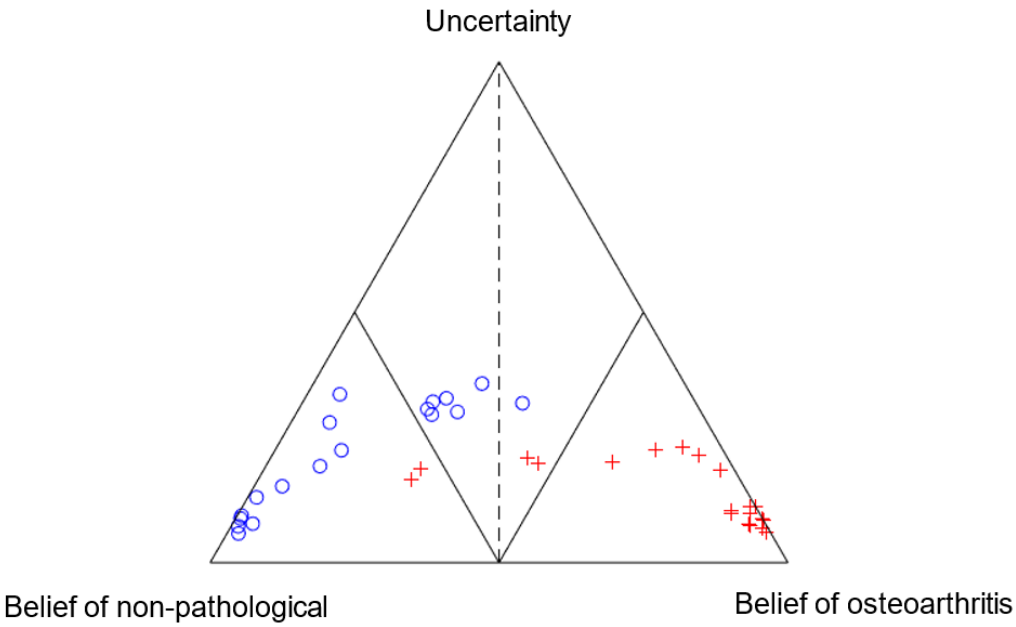


Figure 32 Classifier simplex plot representing the out-of-sample classification accuracy. The blue circles represent the non-pathological volunteers, the red crosses the patients pre-surgery. The black, dashed line indicates the boundary between the non-pathological and osteoarthritic area of the plot. The solid lines within the plot, indicate a Belief of osteoarthritis (BOA) and Belief on non-pathological (BNP) of 0.50 (the boundary between dominant and non-dominant BOA and BNP, respectively).

In the classifier classification based on the whole dataset, the sensitivity was 90%, and the specificity was 94%. One NP was misclassified as non-dominant OA (subject 02840, left side), while two cases pre-TKR were classified as non-dominant NP (patient 02867, both knees). It must be noted that no knee radiographs were performed on the NP subjects, and therefore, it was unknown whether the false positives may potentially have had radiographic signs of knee OA. When observing the raw data for the patient, for both sides' GRF, joint angles and moments in the sagittal and frontal plane were within one SD from the NP data, except the knee adduction angles on the left side, indicating an increased valgus, especially during stance, and the knee adduction moments on both sides, showing the presence of a first peak abduction moment, as opposed to the adduction moment observed in NPs. By observing the raw biomechanical data, it could be concluded that this subject had a seemingly non-pathological gait, and therefore, the classification was not the result of an error.

Research question: Are features related to trunk lean and trunk flexion-extension during gait (i.e., trunk kinematics in the frontal and sagittal planes, respectively) included in the eighteen highest-ranking features discriminating between NPs and patients with severe knee OA?

The findings from the classifier showed that no trunk kinematics were included in the eighteen highest-ranking features discriminating between NPs and patients with severe knee OA (Table 29). Trunk flexion-extension angle PC1, PC2 and PC3 did not discriminate well between patients and NPs; trunk flexion-extension PC1 described patients' larger flexion throughout the gait cycle, and PC2 and 3 represented 0% of the variance of the original data. Trunk lateral lean PC1 and PC3 were thirty-ninth (61% accuracy) and forty-eight (55% accuracy), respectively, in discriminating patients from NPs. Trunk lean PC1 (80% of the variance of original data) described a shift upwards of the patients' trunk lean waveform, indicating that patients' trunks leaned towards the ipsilateral limb throughout the gait cycle more than NPs. Trunk lean PC3 (7% of the original data variance) represented NPs' larger trunk lateral lean ROM during the gait cycle.

Figure 33 shows that the mean raw waveforms describing the trunk angles in the three planes had very subtle differences between patients (orange, solid line) and NPs (blue, dashed line).

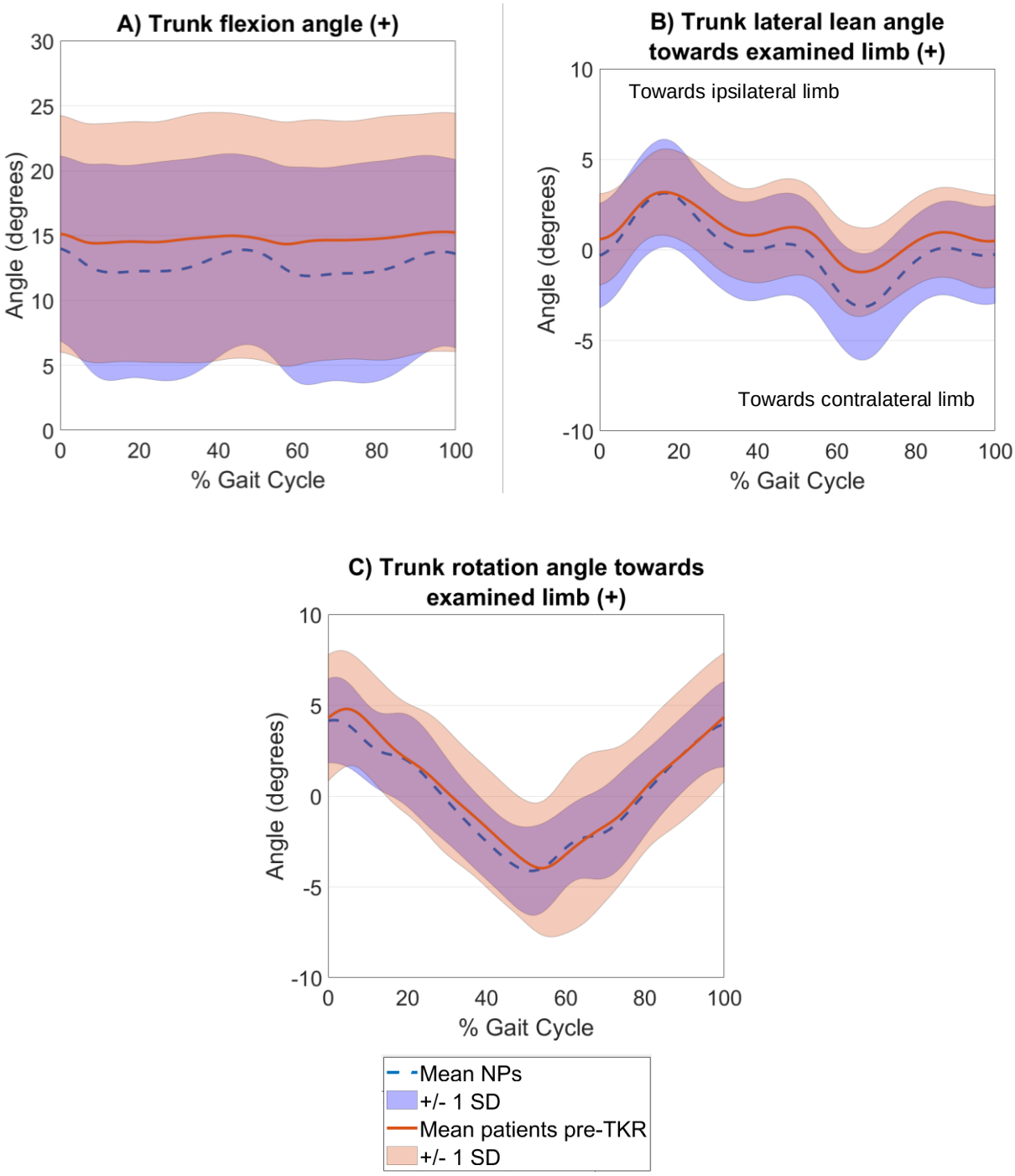


Figure 33 Ensemble averages of patients' and NPs' trunk kinematics in the (A) sagittal, (B) frontal, and (C) transverse planes. (B) e.g. when the right lower limb is examined, a positive value corresponds to the right shoulder being lower than the left; (C) e.g., when the right lower limb is examined, a positive value shows that the left shoulder is more forward than the right one.

Table 29 Eighteen highest-ranking gait biomechanics features most accurately discriminating between NPs and patients pre-TKR

Principal components	% of variance represented	Ranking based on accuracy	Accuracy (%)
Knee power PC1	29	1	92.1
Vertical GRF PC1	62	2	92.1
Ankle power PC1	24	3	86.8
Knee flexion-extension moment PC2	25	4	86.8
Pelvis obliquity (frontal plane) PC2	19	5	86.8
Ankle dorsi-plantarflexion moments PC2	32	6	84.2
Hip flexion-extension angle PC2	12	7	84.2
Anterior-posterior GRF PC1	61	8	81.6
Hip flexion-extension moment PC2	30	9	76.3
Knee abduction-adduction moment PC3	8	10	76.3
Hip abduction-adduction angle PC2	7	11	76.3
Knee flexion-extension angle PC2	16	12	73.7
Pelvis rotations (transverse plane) PC3	8	13	73.7
Ankle dorsi-plantarflexion angle PC2	22	14	73.7
Hip power PC2	21	15	71.1
Hip abduction-adduction moment PC1	39	16	71.1
Medio-lateral GRF PC2	14	17	71.1
Knee abduction-adduction angle PC3	3	18	68.4

PC: Principal Component; GRF: Ground Reaction Force

Considering that in most cases, at each follow-up, the patients included in the study were different, it was decided to present the results by follow-up time, for clarity, rather than by aims and hypotheses. A summary of the findings related to the hypothesis is presented at the end of the Results (section 6.3.7).

6.3.3 Relationship between BOA, PBTs and PROMs pre-TKR

Eighteen patients were assessed pre-surgery, and two of these were scheduled for a simultaneous, bilateral TKR, hence, twenty knees were evaluated. Pre-surgery, the patients' belief values were significantly lower than NP's, except for the BOA which was higher, the mean BOA was 0.77 ± 0.19 ($p < 0.001$, non-parametric), the BNP was 0.09 ± 0.15 ($p < 0.001$, non-parametric), and the Uncertainty was 0.13 ± 0.06 ($p = 0.048$, non-parametric). Patients had a median OKS of 20.5 (8.5 IQR), a median KOOS ADL of 47.1 (12.5 IQR), and a median KOOS Pain of 38.9 (13.9 IQR).

Patients performed significantly fewer sit-to-stand repetitions (median: 10, 4 IQR) than NPs (median: 14, 4 IQR) ($p < 0.001$, non-parametric) during the 30sCT. Two patients did not perform the SCT due to knee pain, and the remaining patients were 21.5% slower in the SCT (median 7.53 s, 2.92 IQR) than reference participants (median 5.91 s, 1.53 IQR) ($p = 0.006$, non-parametric); moreover, one patient needed to use the handrail while ascending only, three subjects used it only when descending and the majority ($n = 9$) used the handrail both ascending and descending. Additionally, patients took 36.1% longer (mean TUG: 12.66 ± 2.92 s) than NPs (mean TUG: 9.30 ± 1.37 s) ($p = 0.003$) to complete the TUG. Three patients did not complete the 40mFPWT due to lack of time during data collection, which was limited to three hours by the study protocol, and the remaining patients were 26.9% slower (mean: 1.28 ± 0.24 m/s) than reference volunteers (mean: 1.75 ± 0.20 m/s) ($p < 0.001$) to complete the task.

Table 30 shows that there was a significant, positive, strong relationship between BOA and the 40mFPWT pre-TKR, and a significant, negative, moderate relationship between the BOA pre-TKR and SCT. The correlation between the BOA and 30sCT was close to significance and revealed a negative, potentially moderate relationship. There was no relationship between the BOA and OKS or KOOS sub-scores pre-TKR. Similarly, no relationship was found between any of the PBTs and the PROMs pre-TKR. Surprisingly, most patients who had a high TUG timing, also had a high OKS, KOOS Pain and ADL (i.e., best perceived pain and function).

Table 30 Correlation coefficients and significance levels (in brackets) for the variables pre-TKR

	BOA (n = 20)	30sCT (n = 20)	TUG (n = 20)	40m fast-paced walk test (n = 17)	Stair climb test (n = 18)
30sCT (n = 20)	-0.525‡ (0.017)				
TUG (n = 20)	0.284‡ (0.224)				
40m fast-paced walk test (n = 17)	**-0.845‡ (< 0.001*)**				
Stair climb test (n = 18)	**0.510‡ (< 0.001*)**				
OKS (n = 20)	-0.203‡ (0.390)	0.138‡ (0.562)	0.162‡ (0.495)	0.333‡ (0.192)	-0.149‡ (0.555)
KOOS (n = 20)					
Pain	-0.150‡ (0.527)	0.154‡ (0.517)	0.273‡ (0.244)	0.104‡ (0.692)	-0.102‡ (0.688)
ADL	-0.432‡ (0.057)	0.335‡ (0.149)	0.079‡ (0.740)	0.436‡ (0.080)	-0.463‡ (0.053)

BOA: Belief of Osteoarthritis; 30sCT: 30-second Chair Test; TUG: Timed-up and go; n: number; OKS: Oxford Knee Score; KOOS: Knee injury and Osteoarthritis Outcome Score; ADL: Activities of Daily Living; ‡: Spearman's rho.
*: significant at $p < 0.007$

6.3.4 Relationship between BOA, PBTs and PROMs three months post-TKR

Six patients (at baseline, mean age: 68.7 ± 5.8, height: 1.70 ± 0.08, weight: 78.1 ± 11.4, BMI: 27.5 ± 6.5; 1 female) attended the three-month follow-up data collection (median follow-up time: 3 months, 0 IQR), of these, two had a bilateral TKR, therefore, eight knees were assessed (40% of the total knees analysed pre-TKR). All patients had already completed their rehabilitation program at the time of the assessment, with a median number of 7 sessions (2.8 IQR). An equal number of subjects joined group rehabilitation (n = 2), one-to-one sessions (n = 2) or a mix of individual and group rehabilitation (n = 2).

Table 31 shows that patients had a significant improvement in their gait quality (i.e., BOA value decrease). One patient used their hands halfway through the 30sCT, utilised a walking stick for the first 10 m of the 40mFPWT, and was unable to perform the SCT. None of the remaining five patients utilised the handrail during the SCT. There was a small improvement in time needed to complete the TUG (mean 0.70 ± 4.16 s faster), SCT (mean 0.27 ± 4.79 s faster), 40mFPWT (mean 0.06 ± 0.27 s faster), and 30sCT (median 2.5, 2.5 IQR, more repetitions). Although there was a

modest improvement in the PBTs' scores, none of them changed significantly pre to three months post-TKR. Conversely, the OKS, KOOS pain and ADL significantly improved three months post-TKR.

Table 31 Differences in gait function, PBTs, and PROMs change pre to three months post-TKR

	Pre-TKR Median (IQR) (n = 8)	3 months post-TKR Median (IQR) (n = 8)	Significance level
Gait function			
BOA	0.87 (0.25)	0.66 (0.67)	0.025*
Patient-reported outcomes			
OKS	22.5 (4.3)	36 (6.5)	0.049*
KOOS			
Pain	40.3 (13.2)	72.2 (16.7)	0.035*
ADL	50.0 (15.7)	87.5 (4.0)	0.035*
PBTs	(n = 6)	(n = 6)	
30sCT (n repetitions)	10.5 (3.8)	10.5 (4.8)	0.058
TUG (s)	12.18 (6.33)	10.97 (3.75)	0.917
40m fast-paced walk test (s)	1.38 (0.20)	1.36 (0.46)	0.715
	(n = 5)	(n = 5)	
Stair climb test (s)	6.56 (1.34)	5.58 (1.72)	0.068

TKR: Total Knee Replacement; IQR: Interquartile Range; n: number; BOA: Belief of Osteoarthritis; OKS: Oxford Knee Score (0 = extreme knee-related issues, 48 = no knee-related issues); KOOS: Knee injury and Osteoarthritis Outcome Score (0 = extreme knee-related issues, 100 = no knee-related issues); ADL: Activities of Daily Living; 30sCT: 30-second Chair Test; TUG: Timed-up and go.
*: significant at $p < 0.05$

Table 32 shows that three months post-TKR, neither the BOA nor the PBTs correlated with the OKS or the KOOS sub-scales. Nevertheless, there was a significant, positive, strong relationship between the BOA and TUG. A significant, negative, very strong relationship was observed between the BOA and the 40mFPWT. The relationship between BOA and the SCT was close to significance.

Table 32 Correlation coefficients and significance levels (in brackets) between the variables at three months post-TKR

	BOA (n = 8)	30sCT (n = 8)	TUG (n = 8)	40m fast-paced walk test (n = 8)	Stair climb test (n = 7)
30sCT (n = 8)	-0.558‡ (0.151)				
TUG (n = 8)	**0.896‡ (0.003*)**				
40m fast-paced walk test (n = 8)	**-0.963 (< 0.001*)**				
Stair climb test (n = 7)	0.853 (0.015)				
OKS (n = 8)	-0.036‡ (0.932)	-0.331‡ (0.423)	-0.112‡ (0.792)	0.171‡ (0.686)	0.667‡ (0.102)
KOOS (n = 8)					
Pain	-0.085‡ (0.842)	-0.463‡ (0.248)	0.013‡ (0.977)	0.209‡ (0.620)	0.579‡ (0.173)
ADL	-0.491‡ (0.217)	0.720‡ (0.044)	-0.605‡ (0.112)	0.509‡ (0.198)	-0.147‡ (0.753)

BOA: Belief of Osteoarthritis; 30sCT: 30-second Chair Test; TUG: Timed-up and go; n: number; OKS: Oxford Knee Score; KOOS: Knee injury and Osteoarthritis Outcome Score; ADL: Activities of Daily Living; ‡: Spearman's rho.
*: significant at $p < 0.007$

To explain the unexpected sign of the relationship between the PBTs and the PROMs, the original data were reported in Table 33. The negative, non-significant relationship between the 30sCT and the OKS, and the KOOS Pain can be explained by patient 02868, who was an outlier and was influential in determining the correlation sign; this subject had the highest OKS and KOOS Pain and performed the smallest number of sit-to-stand repetitions within the group.

The surprising positive relationship between the TUG and the KOOS Pain was observed because, except for patient 02893, most patients with the highest TUG time had a low KOOS Pain score.

The unexpected positive relationship between the SCT and the OKS and the KOOS Pain was because patients who took the shortest time to complete the test (02851, 02867, 02883), also had the lowest score both in the OKS and KOOS Pain, indicating worse perceived function-symptoms, and pain, respectively.

Table 33 Patients' PROMs and PBTs three months post-TKR

Patient ID	30sCT (n repetitions)	TUG (s)	Stair Climb Test (s)	OKS affected leg	KOOS Pain	KOOS ADL
02851	24	9.53	5.52	37	66.7	92.7
02867_R	16	10.17	5.26	35	72.2	88.2
02867_L	16	10.17	5.26	35	72.2	88.2
02868_R	9	10.97	7.24	42	91.7	86.8
02868_L	9	10.97	7.24	42	91.7	86.8
02881	11	15.17	8.28	40	80.6	91.2
02883	10	10.97	5.58	31	66.7	79.4
02893	10	16.6		8	13.9	17.7

ID: identifier; n: number; s: seconds; OKS: Oxford Knee Score; KOOS: Knee injury and Osteoarthritis Outcome Score; ADL: Activities of Daily Living; R: right; L: left.

6.3.5 Relationship between BOA, PBTs and PROMs six months post-TKR

Five patients (at baseline, mean age: 69.4 ± 7.0, height: 1.67 ± 0.06, weight: 76.1 ± 11.2, BMI: 27.2 ± 3.5; 1 female) attended the six-month follow-up data collection (mean follow-up time: 6.2 ± 0.7 months), and two had a bilateral TKR, hence, seven knees were assessed (35% of the total knees analysed pre-TKR). Three out of five patients had also attended the three-month post-TKR assessment, while two had not, one because of human face-to-face research suspension due to COVID-19, and another due to the main researcher's limited availability in collecting data at the time. These two patients had six rehabilitation sessions, both one-to-one.

During the SCT, one patient utilised the handrail while descending, and another one used the handrail both ascending and descending.

Table 34 shows that there was a significant improvement in gait function (i.e., BOA), number of repetitions in the 30sCT (median 3, 1.9 IQR, more repetitions), and TUG timings (mean 1.73 ± 0.96 s faster) pre to six months post-TKR. Moreover, patients were faster in both the 40mFPWT (mean 0.24 ± 0.20 m/s faster) and the SCT (mean 1.12 ± 2.48 s faster), but the differences from the scores pre-TKR, although close to significance, were not statistically significant. Surprisingly, the median BOA post-TKR was very low, indicating that for most patients in this group, there was a substantial improvement in gait function six months post-surgery.

Table 34 Differences in gait function, PBTs, and PROMs change pre to six months post-TKR

	Pre-TKR Median (IQR) (n = 7)	6 months post-TKR Median (IQR) (n = 7)	Significance level
Gait function			
BOA	0.76 (0.34)	0.22 (0.50)	0.018*
Patient-reported outcomes			
OKS	23 (5)	40 (6.5)	0.034*
KOOS			
Pain	36.1 (9.8)	72.2 (33.3)	0.018*
ADL	45.3 (27.4)	97.1 (21.4)	0.018*
30sCT (n repetitions)	13 (5)	15 (6)	0.042*
TUG (s)	12.45 (2.94)	10.70 (2.36)	0.043*
40m fast-paced walk test (s)	1.35 (0.28)	1.55 (0.55)	0.068
Stair climb test (s)	7.53 (1.37)	6.95 (3.31)	0.068

TKR: Total Knee Replacement; IQR: Interquartile Range; n: number; BOA: Belief of Osteoarthritis; OKS: Oxford Knee Score (0 = extreme knee-related issues, 48 = no knee-related issues); KOOS: Knee injury and Osteoarthritis Outcome Score (0 = extreme knee-related issues, 100 = no knee-related issues); ADL: Activities of Daily Living; 30sCT: 30-second Chair Test; TUG: Timed-up and go.
*: significant at $p < 0.05$

Table 35 displays no correlation between the BOA and the PBTs, or between the BOA and the PROMs.

Table 35 Correlation coefficients and significance levels (in brackets) between the variables at six months post-TKR

	BOA (n = 7)	30sCT (n = 7)	TUG (n = 7)	40m fast-paced walk test (n = 7)	Stair climb test (n = 7)
30sCT (n = 7)	-0.874 (0.010)				
TUG (n = 7)	0.637 (0.124)				
40m fast-paced walk test (n = 7)	-0.738 (0.058)				
Stair climb test (n = 7)	0.671 (0.099)				
OKS (n = 7)	0.036‡ (0.938)	0.083‡ (0.859)	0.546‡ (0.205)	0.361‡ (0.426)	-0.417‡ (0.352)
KOOS (n = 7)					
Pain	-0.090‡ (0.848)	<0.001‡ (1.000)	0.624‡ (0.134)	0.440‡ (0.323)	-0.330‡ (0.465)
ADL	-0.073‡ (0.877)	<0.001‡ (1.000)	0.630‡ (0.130)	0.444‡ (0.318)	-0.333‡ (0.465)

BOA: Belief of Osteoarthritis; 30sCT: 30-second Chair Test; TUG: Timed-up and go; n: number; OKS: Oxford Knee Score; KOOS: Knee injury and Osteoarthritis Outcome Score; ADL: Activities of Daily Living; ‡: Spearman's rho.
*: significant at $p < 0.007$

6.3.6 Relationship between BOA, PBTs and PROMs twelve+ months post-TKR

Twelve months post-TKR, six patients were assessed (at baseline, mean age: 68.7 ± 6.6, height: 1.69 ± 0.08, weight: 94.7 ± 16.8, BMI: 32.9 ± 4.9; 1 female), all of whom received a TKR on one knee only (30% of the total knees analysed pre-TKR). The mean follow-up time was 16.0 ± 2.8 months, as all the patients who were assessed at this time point, had their pre-TKR visit before the COVID-19 pandemic and in certain cases, received the TKR either before the COVID-19 pandemic in March 2020 (n = 4), or in 2021 (n = 2). Data collection on NHS patients outside NHS premises, or on patients who were at risk of developing serious complications if being infected with COVID-19, was not permitted by Cardiff University regulations for several months, and this caused a delay in the data collection at the twelve-month follow-up.

Out of six, two patients were also assessed at six months post-TKR, while three patients could not attend the previous follow-ups due to COVID-19 restrictions on human face-to-face research, and one patient could not attend the preceding data collections due to work commitments.

One participant showed substantial difficulties in negotiating stairs during the task familiarisation, and it was agreed not to perform the SCT for this reason. None of the patients who performed the SCT utilised the handrail.

Table 36 shows that there was a significant decrease in the patients' BOA value twelve months post-surgery. Moreover, the OKS, KOOS pain and ADL improved significantly. The median improvement in the 30sCT was 1.5 (3.3 IQR) repetitions, patients took a mean of 0.48 ± 0.82 s and 0.09 ± 1.33 s less to complete the TUG and the SCT, respectively. Additionally, subjects were 0.09 ± 0.09 m/s faster in the 40mFPWT. Nonetheless, although there were improvements in all the PBTs scores, these were not significant.

Table 37 shows that there were no significant correlations between the BOA and the PBTs twelve months post-TKR.

Table 36 Differences in gait function, PBTs, and PROMs change pre to twelve+ months post-TKR

	Pre-TKR Median (IQR) (n = 6)	Post-TKR Median (IQR) (n = 6)	Significance level
Gait function			
BOA	0.85 (0.14)	0.63 (0.25)	0.028*
Patient-reported outcomes			
OKS	19 (9)	42 (7.8)	0.027*
KOOS			
Pain	36.1 (5.6)	91.2 (12.7)	0.028*
ADL	46.9 (10.3)	91.2 (13.6)	0.028*
PBTs			
30sCT (n repetitions)	12 (3.5)	13 (2.3)	0.068
TUG (s)	11.35 (2.42)	10.36 (1.26)	0.345
40m fast-paced walk test (s)	1.30 (0.10)	1.42 (0.10)	0.075
	(n = 5)	(n = 5)	
Stair climb test (s)	7.08 (1.57)	6.37 (1.28)	0.500

TKR: Total Knee Replacement; IQR: Interquartile Range; n: number; BOA: Belief of Osteoarthritis; OKS: Oxford Knee Score (0 = extreme knee-related issues, 48 = no knee-related issues); KOOS: Knee injury and Osteoarthritis Outcome Score (0 = extreme knee-related issues, 100 = no knee-related issues); ADL: Activities of Daily Living; 30sCT: 30-second Chair Test; TUG: Timed-up and go.
*: significant at $p < 0.05$

Table 37 Correlation coefficients and significance levels (in brackets) between the variables at twelve months post-TKR

	BOA (n = 6)	30sCT (n = 6)	TUG (n = 6)	40m fast-paced walk test (n = 6)	Stair climb test (n = 5)
30sCT (n = 6)	0.319‡ (0.538)				
TUG (n = 6)	0.219 (0.677)				
40m fast-paced walk test (n = 6)	-0.543‡ (0.266)				
Stair climb test (n = 5)	-0.400‡ (0.505)				
OKS (n = 6)	0.429‡ (0.397)	0.899‡ (0.015)	-0.200‡ (0.704)	0.200‡ (0.704)	-0.400‡ (0.505)
KOOS (n = 6)					
Pain	0.371‡ (0.468)	0.812‡ (0.050)	-0.314‡ (0.544)	0.029‡ (0.957)	-0.300‡ (0.624)
ADL	0.145‡ (0.784)	0.647‡ (0.165)	0.232‡ (0.658)	0.406‡ (0.425)	-0.667‡ (0.219)

BOA: Belief of Osteoarthritis; 30sCT: 30-second Chair Test; TUG: Timed-up and go; n: number; OKS: Oxford Knee Score; KOOS: Knee injury and Osteoarthritis Outcome Score; ADL: Activities of Daily Living; ‡: Spearman's rho.
*: significant at $p < 0.007$

6.3.7 Summary of findings in relation to the hypotheses

Hypothesis 1

The hypothesis could not be confirmed for all time points in the study, since relationships between some of the PBTs and the BOA existed only pre- (BOA associated with 40mFPWT, SCT and trends of association with 30sCT), and three months post-TKR (BOA associated with 40mFPWT, TUG, trends of association with SCT). Nevertheless, the 40mFPWT was the PBT most strongly correlated with the BOA at both time points. Trends of associations were found between BOA and 30sCT, 40mFPWT and SCT six months post-TKR, but not one-year post-surgery.

Hypotheses 2.1 and 2.2

The results disproved the hypotheses as no correlation was found between BOA and the included PROMs at any time point during the study.

Hypotheses 3.1 and 3.2

The results disproved the hypotheses as no correlation was found between the included PROMs and PBTs at any time point during the study.

6.4 DISCUSSION

6.4.1 Cardiff classifier including trunk kinematics

One of the aims of the study was to determine whether trunk kinematics features during gait would contain important information aiding in discriminating individuals with severe knee OA from NPs within the Cardiff classifier. The current results showed that in the cohort examined, there was no evidence that trunk kinematics features were some of the main biomechanical characteristics discriminating patients from a group of NPs within the Cardiff classifier. A visual inspection of the trunk kinematics' raw waveforms revealed that patients had a slightly more flexed trunk throughout the gait cycle, and a reduced trunk ROM in the frontal plane, which were also characteristics depicted by PCA, but the differences from NPs throughout the kinematics waveforms were not greater than 3 degrees in the three anatomical planes, and the trunk kinematics PCs had a limited accuracy in distinguishing patients from NPs.

Most of the existing literature comparing trunk kinematics between people with knee OA and NPs looked at discrete values (i.e., peak angles) rather than patterns throughout the gait cycle and found very small differences (Hunt et al. 2010; Bechard et al. 2012; Creaby et al. 2012), which would not be considered clinically relevant, as they were inferior to 5 degrees (Schmid et al. 2016). Moreover, only a few studies looked at alterations in the trunk kinematics in the sagittal plane (Turcot et al. 2013; Preece et al. 2019; Preece and Alghamdi 2021). The current study seems to align with previous reports, since patients had a larger trunk flexion than NPs, and the difference between groups was minimal. (Preece and Alghamdi 2021) found that patients with varying degrees of knee OA had a mean 2.8° larger trunk flexion in the first portion of stance. (Turcot et al. 2013) found that patients with a varus knee alignment had a larger trunk flexion-extension ROM compared to NPs, but the difference, although significant, was less than 1° during stance, and this pattern was not observed in subjects with a valgus knee alignment whose trunk flexion was comparable to NPs.

A recent systematic review (Iijima et al. 2019) including nine studies and 446 people with knee OA concluded that trunk lean towards the ipsilateral limb was significantly larger in patients than NPs of similar age, with a large effect size. This is in contrast with the current report, nevertheless, when examining the studies

included in the systematic review, most of them found very small differences between people with knee OA and NPs, which were inferior to 5 degrees, on average (Hunt et al. 2010; Bechard et al. 2012; Creaby et al. 2012; Turcot et al. 2013). Moreover, there were inconsistencies in the time points of the gait cycle where differences in trunk lean were found between patients and NPs, some authors found them at 20% stance (Hunt et al. 2010), others at 75% of stance (Bechard et al. 2012), which would be in correspondence of the peak knee adduction moments, and some failed to report when the differences occurred (Creaby et al. 2012; Turcot et al. 2013). To add to the inconsistencies, earlier reports found no differences in trunk lean between patients and NPs (Preece et al. 2019), some researchers found no differences in trunk lean at the abovementioned time points when looking at discrete values, but when PCA was employed in the analysis, they detected a larger trunk lean ROM due to an increased trunk lean at toe-off (Linley et al. 2010). PCA also revealed an increased trunk lean in the frontal plane throughout stance in a more recent study comparing a small group of individuals with knee OA and controls (Federolf et al. 2013). Another point to consider is that several studies (Hunt et al. 2008; Linley et al. 2010; Bechard et al. 2012; Turcot et al. 2013) utilised a marker set that was not optimal for tracking the trunk kinematics (Armand et al. 2014) since markers were placed on the acromions, which can move independently from the trunk, other authors failed to report how they tracked the trunk (Creaby et al. 2012; Federolf et al. 2013) and therefore, their results should be treated cautiously.

Some authors suggested that these small differences in trunk kinematics may be sufficient to produce a large effect on reducing the knee external moments (all joint moments mentioned followingly are meant as external), due to the mass of the trunk. Nevertheless, previous evidence showed a decrease of only 7% in the knee adduction moment (not significant in the study) when the trunk lean towards the stance limb was experimentally increased by 4 degrees in healthy subjects (Hunt et al. 2011). The reduction in the first peak knee adduction moment was 10% when increasing the trunk lean by about 6° in a group of people affected mostly by early or moderate knee OA (Simic et al. 2012). Therefore, it may be suggested that people with knee OA in the abovementioned studies (Hunt et al. 2010; Bechard et al. 2012; Creaby et al. 2012; Turcot et al. 2013) did not increase their trunk lean enough to cause a significant change in knee moments. Furthermore, an augmented trunk lean

towards the stance limb had the opposite effect of increasing the knee adduction moment in people with a knee varus alignment (Iijima et al. 2018), suggesting that the knee alignment, more than the trunk lean, contributed to the knee adduction moment magnitude, and any trunk adaptation in presence of knee varus was not successful in reducing the knee loading, an idea supported by the results of the previous investigation (Hunt et al. 2008). It was found that the variance in the first peak knee adduction moment was mainly explained by the knee alignment (25%), and trunk lean contributed to the knee moments almost to the same degree (13%) as toe-out angle (12%) (Hunt et al. 2008).

The findings of previous research seem to suggest that trunk lean strategies vary widely between patient groups but may not be clinically relevant. Additionally, while trunk lean towards the stance limb and trunk flexion were larger in people with a radiographic, static, varus knee alignment, no differences were found in trunk kinematics between controls and patients with a knee valgus (Sagawa et al. 2013; Turcot et al. 2013). Although the participants' knee alignment information was not available in the present study, given the findings of earlier research (Sagawa et al. 2013; Turcot et al. 2013), it may be that the current group of patients consisted mainly of subjects with a knee valgus alignment, which might justify why no discriminative power was found for trunk kinematics gait features. It must be noted that the results of the current study were based on a limited, small number of participants, for reasons explained earlier, and further research on a larger cohort is recommended, including information on knee alignment, to explore whether different results may be achieved for subjects with a knee varus or valgus, and the relationship between trunk kinematics and knee moments. Nevertheless, this is the first study utilising a gait classification method to demonstrate that even if alterations may be found in trunk kinematics in people with severe knee OA when gait biomechanics of several joints are considered together, other gait features may be more relevant than trunk kinematics in distinguishing between NPs and patients at this stage of the knee disease.

6.4.2 Relationship between BOA and PBTs

Hypothesis 1

Most of the existing research compared PBTs to PROMs (Stratford et al. 2010; Mizner et al. 2011; Stevens-Lapsley et al. 2011a; Hamilton et al. 2012b; Graff et al. 2016; Givens et al. 2018; De Vroey et al. 2018; Choi et al. 2020; Onodera et al. 2020; Christensen et al. 2022), which have been suggested to measure different constructs of function (Tolk et al. 2017). Hence, the comparison to previous literature was limited, as only another study correlated summative measures of gait function, the GDI and GDI-kinetic, to PBTs, namely the TUG and five-time sit-to-stand, and only pre-TKR (Naili et al. 2017a). Due to the several, described limitations to data collection, a small sample of patients was available post-TKR, and for this reason, the post-TKR results should be considered exploratory in nature and treated with caution, especially at twelve months post-TKR, where the presence of unusual data seemed to have influenced the correlation between all outcomes. It is also worth noting that at each follow-up, the patients examined were slightly different from the preceding post-TKR assessment.

As hypothesised, the strongest relationship was found between the gait function classification (i.e., BOA) and the 40mFPWT both pre and three months post-TKR. At six months, the correlation between these measures was close to significance, showing that a better gait quality was strongly associated with a faster walking speed achievable. Moreover, a moderate, positive correlation was found between the BOA and the SCT pre-TKR. The relationship between BOA and the SCT was in the expected direction (positive) three and six months post-TKR, but the lack of correlation may have been linked to the small sample size because the relationship was close to significance. These results suggested that a better gait function was also related to less time negotiating the four steps as fast as possible. No previous study compared the 40mFPWT or the SCT to overall measures of gait function, but evidence showed that a reduced knee ROM was correlated with a slower SCT, and both restricted knee ROM and quadriceps weakness were associated with slower speeds during the 40mFPWT pre-TKR, but not one-year post-surgery (Christensen et al. 2022).

Contrary to the hypothesis, there was no significant correlation between the BOA and 30sCT at either time point of the study, although the relationship showed a trend, since it was in the expected direction and close to significance pre-TKR (-

0.525, p = 0.017), six months post-TKR (-0.874, p = 0.010), but less so three months post-TKR (-0.558, p = 0.151), and one year after surgery, suggesting a linear association between the measures. Of note, the results one-year post-TKR may have been due to unusual patterns in the data. The lack of a significant correlation between the BOA and 30sCT at other time points could be attributed to the relatively small sample size and the choice of utilising a Bonferroni correction due to multiple comparisons. The Bonferroni correction of the significance level is believed to be a very conservative approach (Greenhalgh 2019) had it not been applied, the association between the BOA and 30sCT, would have been significant, at least pre- and six months post-TKR. This finding would have been in line with the only other study, to the best of the author's knowledge, assessing the correlation between the gait quality evaluated with a summary measure of gait (i.e., GDI-kinetic) and the five-time sit-to-stand in forty patients awaiting TKR, showing a weak to moderate relationship between the measures (Naili et al. 2017a). The BOA and 30sCT measure slightly different constructs of function, the first related to walking, the second mainly associated with "changing basic body position" (WHO 2023), and lower limb strength (Coleman et al. 2020). However, Christensen et al. (2021) found a correlation between having a stronger quadriceps, and a higher peak vertical GRF, a greater knee extension moment during gait in a hundred ninety-five patients at least six months post-TKR. Features of the vertical GRF and knee flexion-extension moment during gait are recurringly appearing as the highest-ranking characteristics discriminating between people with knee OA and NPs in the Cardiff classifier (Table 29) (Biggs et al. 2019b) and in Chapter 5. This may explain why the BOA and the 30sCT may be indirectly correlated, even if not significantly, in the current study. It may be inferred that a better gait quality could potentially be related to more repetitions during the 30sCT.

It was hypothesised that since the TUG includes walking, a better gait quality (i.e., BOA) would be associated with the self-selected speed at which the TUG was completed. Surprisingly, a strong correlation between the BOA and the TUG was only found at three months post-TKR. Pre-TKR, the TUG timings did not seem to follow an obvious pattern, since some of the patients who had a highly compromised gait, took a shorter time to complete the test than patients who were classed as having a better gait quality. This was true at the six and twelve-month post-TKR assessments alike, but the TUG timings between patients having different BOA were

not considerably different, and the sample size was also much smaller than pre-TKR. It may be that the correlation between TUG and BOA at three months post-TKR was due to chance, given what was observed pre-TKR, where the group was three times bigger, and at the other time points where a similar number of patients to the three-month appointment was analysed. These findings highlight that the movement quality during gait may not be related to the timing to complete a task also including change of direction and sit-to-stand. The TUG is the only PBT which requires the subject to move at their preferred speed (OARSI 2023). It may be that patients took varying times during the transitional phases of the TUG, i.e., standing up, changing direction, and sitting down, and these may have been dependent on the subject's comfortable speed in carrying out those transitions, given their balance, and coordination, aspects that may have not been captured by the assessment of gait quality (= BOA). It is important to mention that the chair utilised during the test had a deep seat (anecdotally, this was commented on by several patients during the assessment), but by looking at the data (not reported), there did not seem to exist a relationship between subject's height and TUG timings.

The current findings were surprising if compared to previous evidence showing a strong relationship between the GDI-kinetic and TUG, but no correlation between TUG and GDI (comparing exclusively gait kinematics to a reference group), in a group of forty patients awaiting TKR (Naili et al. 2017a). Chapter 4 showed that BOA and GDI-kinetic correlate with each other moderately pre-TKR, so it may have been expected that a degree of correlation would have been found between BOA and TUG. The present results may suggest that when measuring function via the BOA and TUG, different outcomes would be obtained with the two measures. Previous evidence found that walking faster at a self-selected speed was moderately correlated to a shorter time to complete the TUG in the presence of advanced knee OA (Mehta et al. 2019). However, the methods in the study of Mehta et al. (2019) were different, as gait speed, as opposed to gait quality, were compared to the TUG, unlike the present study.

6.4.3 Relationship between BOA and PROMs

Hypotheses 2.1 and 2.2

Both the gait quality and the PROMs assessed in the current study improved significantly from pre-surgery to each follow-up post-TKR. The present results suggest that the gait quality was not associated with the levels of pain perceived pre or at different levels of recovery up to about a year post-TKR since no correlation was found between the BOA and OKS, or KOOS Pain.

The lack of association between the gait quality (i.e., BOA) and the OKS was unexpected, especially pre-TKR, given the larger sample size, and contradicted hypothesis 2.1 and earlier findings of our research group (Biggs et al. 2019a). A strong, negative association was found between these measures in a different set of patients undergoing TKR (n = 41), and twenty-two subjects one-year post-TKR, with comparable baseline OKS (mean: 20.5) and characteristics (mean age 68.4, height 1.67, BMI 32.5, 53.7% females) to the current cohort, except for the higher proportion of females (Biggs et al. 2019a). However, it must be noted that in the study of Biggs et al. (2019a), all patients were classed as having a severely affected gait pre-surgery (i.e., a high BOA) and none of them was classified as NPs (i.e., a BOA that was lower than the BNP), as opposed to the current study, where two patients had a very low BOA pre-TKR (classified as NPs), and concurrently had a low OKS, indicating a poor self-reported function. Additionally, in the present patient group, individuals who had a high BOA did not always have the lowest OKS both pre and post-TKR, and this explains the lack of correlation observed between the measures, and the contrasting findings from Biggs et al., (2019).

The lack of correlation between gait quality, or BOA, and the KOOS ADL/Pain pre- and post-TKR was in contrast with hypothesis 2.2. Furthermore, the results did not show a trend towards an association between these measures. This may suggest that patients' perception of their function may be independent of the actual way they move and be influenced by other factors, such as anxiety and depression (Baker et al. 2012; Sanchez-Santos et al. 2018), or pain levels (Boonstra et al. 2008; Stevens-Lapsley et al. 2011a; Hamilton et al. 2012b). Pre-surgery, the results were in line with what was found by our research group in a separate set of twelve patients listed for a TKR, assessed via the Cardiff classifier (Biggs 2016), and by Naili et al. (2017), where patients had similar KOOS Pain and ADL scores to the current study, and where the GDI was used as a measure of gait quality in patients

pre-TKR (n = 40). It was surprising that the pain levels were not correlated to the gait quality, but this may be an indication that being pain a subjective experience, it may not follow the same trajectory of gait quality, for instance, previous research showed that lower PROMs were associated with pain catastrophising (Riddle et al. 2022). There is very limited evidence comparing gait biomechanics to PROMs pre-TKR, and another issue is the use of varying PROMs. Earlier research revealed that temporospatial gait parameters were highly correlated to the Knee Society Score-physical function (Bolink et al. 2015b), but the lift-up force when ascending a step was not (Jacobs and Christensen 2009). Moreover, the GDI-kinetic had a weak to moderate correlation with the KOOS ADL, but not the KOOK Pain (Naili et al. 2017a). Being the GDI-kinetic a summative measure of gait kinetic function, it was surprising that no association between the gait quality (= BOA) and KOOS ADL was found. It was difficult to justify the contrast with the results of Naili et al. (2017) since the only difference from the current study was a higher proportion of females but it is unclear how this could have influenced the results. It is suggested that future research may compare the GDI and GDI-kinetic scores to the KOOS sub-scores of the current cohort to verify if a correlation would be found between the KOOS ADL and GDI-kinetic as in the study of (Naili et al. 2017a). The present findings, compared to previous evidence (Naili et al. 2017a; Biggs et al. 2019a), showed that PROMs may be very variable in relation to the patient's overall gait function, whereby in certain groups, gait quality and PROMs may be associated, and in others may not be.

Post-TKR, in contrast with the current results, Biggs (2016) observed a negative, moderate relationship between the BOA and the KOOS Pain and ADL in eighteen patients about one-year post-TKR. The author (Biggs 2016) failed to report the KOOS ADL and Pain scores, therefore it was not possible to determine if any differences between the patient groups may have been present to justify such finding. In the existing literature, the correlation between gait measures and PROMs post-TKR seems to be dependent on the PROM and the gait variable analysed. Earlier findings at three months post-TKR were contrasting and showed that a moderate correlation was found between self-selected gait speed and the Knee Society Function Score, but no association was found with the WOMAC (Senden et al. 2011). Conversely, in another study (Bonnefoy-Mazure et al. 2017), weak correlations were found between the WOMAC and temporospatial parameters (i.e.

cadence and gait speed), or knee flexion during gait. Only one report comparing PROMs with gait biomechanics six months post-TKR was found, the authors (Kurihara et al. 2021) discovered that having larger peak hip external flexion and extension moments, knee flexion moment and knee power absorption (i.e., eccentric activity) was related to having a lower Japanese Knee Outcome Measure (i.e., a better score). However, these findings were difficult to contextualise as patients were not compared to a reference group, and it was impossible to understand whether having larger moments meant being more similar to NPs or not. Twelve months post-TKR, weak and moderate correlations were found between the WOMAC-physical function sub-score, pain (Visual Analogue Scale) and gait speed, respectively, and a weak relationship was found between the peak knee flexion and WOMAC-physical function during gait (Bonnefoy-Mazure et al. 2017). Some of these findings were corroborated by other authors, who also observed a moderate relationship between gait temporospatial parameters and the Knee Society Function Score (Bolink et al. 2015b). The present patient group was very small post-surgery, and results were highly affected by the presence of outliers or unusual data, therefore, the different findings to (Biggs 2016) could most likely be due to this factor, while the diverging findings from previous literature (Senden et al. 2011; Bolink et al. 2015b; Bonnefoy-Mazure et al. 2017; Kurihara et al. 2021) may also be due to the use of different measures of subjective and objective function utilised, because temporospatial parameters were utilised (Senden et al. 2011; Bonnefoy-Mazure et al. 2017), as opposed to gait biomechanics, or only individual gait variables were compared to PROMs (Kurihara et al. 2021) rather than an overall measure of gait quality.

6.4.4 Relationship between PBTs and PROMs

Hypotheses 3.1 and 3.2

The results of the current study contrasted with the hypotheses since no correlation was found between the PBTs and the OKS or the KOOS Pain and KOOS ADL at any of the time points of the study. This finding was especially unexpected pre-TKR since the patient group was larger than the groups analysed post-TKR, and less susceptible to outliers. Moreover, the patients included in the present study had generally comparable PBTs scores at baseline to earlier investigations; in the case of the SCT, previous studies used stairs with a varied number of steps, therefore, the

score was divided by the number of steps and multiplied by eight, for comparison with the present investigation; pre-surgery, the current patients took slightly longer to complete the TUG (12.7 s) compared to previous studies (range 9.8-11.6 s) (Mizner et al. 2011; Naili et al. 2017a), similar scores (n repetitions = 10) were found for the 30sCT (range 9-11) (Tolk et al. 2017; Christensen et al. 2022), the SCT (7.53 s), (range 4.66-7.60 s) (Kennedy et al. 2005; Mizner et al. 2011; Tolk et al. 2017), and 40mFPWT (1.28 m/s), (range 1.05-1.30 m/s) (Kennedy et al. 2005; Tolk et al. 2017). Limited evidence was available for the three-month time point (Stratford et al. 2010; Stevens-Lapsley et al. 2011a; Choi et al. 2020). The scores at three and six months were slightly higher than previous studies for the TUG (Stratford et al. 2010; Stevens-Lapsley et al. 2011a; Schache et al. 2019; Choi et al. 2020). The patients in the current study were faster in the SCT (5.58 s) than previous findings (7.59 s) (Choi et al. 2020) three months post-TKR, but had similar scores (6.95 s) to Schache et al. (2019) (7.00 s) six months post-TKR; no information was found concerning the 40m walk test and 30sCT three months post-TKR. One year post-surgery, the current patients were slower than in previous literature in the TUG (Mizner et al. 2011; Stevens-Lapsley et al. 2011a; Graff et al. 2016), and SCT (Mizner et al. 2011; Christensen et al. 2022), but had similar scores in the 30sCT to (Christensen et al. 2022), and were faster than the patients in (Tolk et al. 2017; Christensen et al. 2022) in the 40mFPWT.

Existing literature found a poor association between PROMs and PBTs both pre (Mizner et al. 2011; Hamilton et al. 2012b; Tolk et al. 2017; Givens et al. 2018; Onodera et al. 2020; Christensen et al. 2022; Sharma et al. 2023) and at different stages post-TKR (Stratford et al. 2010; Mizner et al. 2011; Hamilton et al. 2012b; Graff et al. 2016; Choi et al. 2020; Christensen et al. 2022). The correlation was mainly weak between PROMs and a battery of functional performance tests (Hamilton et al. 2012b), the 40mFPWT (Tolk et al. 2017), the SCT (Mizner et al. 2011; Tolk et al. 2017; Onodera et al. 2020; Christensen et al. 2022; Sharma et al. 2023), the TUG (Stratford et al. 2010; Mizner et al. 2011; Graff et al. 2016; Givens et al. 2018; Choi et al. 2020; Onodera et al. 2020; Christensen et al. 2022), and the 30sCT (Tolk et al. 2017; Christensen et al. 2022). This suggests that PBTs and PROMs measure different constructs (Mizner et al. 2011; Bolink et al. 2015b), which is highly likely since PROMs explore a combination of aspects related to the disease, covering symptoms, pain, the distance that can be walked before pain occurs (as in

the OKS), or the grade of difficulty in performing a variety of ADLs (e.g., in the KOOS ADL, shopping, putting on socks, lying in bed, getting inside/out the car, etc.), and there may be only a few questions that would be related to the PBTs, for example, there are only five questions out of seventeen in the KOOS ADL which cover the same tasks analysed by the PBTs.

When looking at the sample size of the abovementioned studies, it was noted that it was much greater than the current one when correlations were found between PBTs and PROMs, including 65 (Givens et al. 2018), 68 (Christensen et al. 2022), 85 (Tolk et al. 2017), 100 (Mizner et al. 2011), 153 (Onodera et al. 2020), 183 (Hamilton et al. 2012b), and 397 participants with knee OA (Sharma et al. 2023). The only exception was the study of (Graff et al. 2016), where only twenty-four patients twelve months post-TKR were included, and a moderate relationship was found between the OKS and TUG, KOOS Pain and TUG, while a strong association was observed between KOOA ADL and TUG. The patients walked as fast as possible during the TUG (Graff et al. 2016), as opposed to their comfortable speed, as suggested by the OARSI guidance (OARSI 2023), this is the only aspect that may justify the difference from the current report (where patients walked at their selected speed), considering that the exclusion criteria, anthropometrics, demographics and statistical tests in the study (Graff et al. 2016) were similar to the present investigation. For instance, (Naili et al. 2017a) enrolled 40 participants and did not find a correlation between the TUG or five-time sit-to-stand with any of the KOOS sub-scales pre-TKR; similarly, (Stevens-Lapsley et al. 2011a) included 39 patients and did not find a significant correlation between the change in the KOOS and change in TUG or SCT between pre and three, six or twelve months post-TKR.

Therefore, it may be suggested that relatively large sample sizes may be required to observe a significant relationship between PROMs and PBTs, and the results of the current study may be due to a type II error, considering the small sample size (Portney and Watkins 2014), especially post-TKR. The present study, although exploratory, and revealing no association between PROMs and PBTs, somewhat aligns with previous research, in that it suggests that PBTs and PROMs seem to measure different aspects of function and they should be used in conjunction.

This study had several limitations that should be accounted for when considering the results, especially the outputs post-TKR. The main limitation was the small number of patients at the follow-ups reduced the power of the analysis and the correlations between variables were very sensitive to the presence of outliers. Another limitation was the different patients analysed at each time point, this was inevitable due to the reasons discussed below, and no analysis was carried out to compare scores between follow-ups for this motive. During the study, attempts were made to increase the number of participants. Minor amendments to the research protocol were made to accommodate for new measures during data collection to guarantee the safety of the participants and staff during the uncertainties of the COVID-19 pandemic, and to start data collection as soon as possible. Most of the patients already enrolled in the study attended their follow-up, but the restrictions in collecting face-to-face data on human participants limited the data that could be collected on the nine patients who enrolled before COVID-19 happened. The recruitment of new participants was very difficult and was not successful until February 2022. Equally, difficulties were encountered in enrolling NPs as a limited number of subjects responded to the recruitment messages on several platforms (Yammer, Twitter, emails to local groups in the targeted age, and word of mouth). This resulted in an NP group that was not age-matched to the patients, and this may have resulted in underestimating the patient's gait function, as demonstrated in Chapter 5. The correlation analysis was limited in that it does not imply the existence of dependent and independent variables and therefore, it could not be stated that having a better gait quality and function increased the PBTs scores. A linear regression analysis would have been able to determine the direction of the relationship between each PBT and the BOA, and this could be the direction of future work on a larger sample.

6.5 CONCLUSION

This was the first study employing trunk kinematics in the sagittal, frontal and transverse plane, together with other gait features, within the Cardiff classifier. Trunk kinematics features were not accurate in discriminating patients pre-TKR from NPs during gait, and gait features related to GRF and other joints' biomechanics were more crucial in determining the classification of function. This finding may need to be confirmed by further studies on larger samples, but it suggested that the accuracy of

the classification in previous and future studies utilising the Cardiff classifier should not be impacted, had trunk kinematics features not been included.

For the first time, all the core three PBTs suggested by OARSI with the addition of the optional TUG were compared to an objective measure of gait quality including both gait kinetics and kinematics, namely, the BOA of the Cardiff classifier, in people with knee OA and at three follow-ups post-TKR. The current findings suggest that there was a correlation or trends of association between the gait quality and the 40mFPWT, the SCT, the 30sCT before and after TKR, and potentially the TUG three months post-TKR. Although PBTs do not inform on the movement quality during the task, impaired biomechanics and the presence of several compensations may be reflected in poorer PBTs scores. These results support the use of the abovementioned PBTs in clinical practice. In the absence of technologies to measure an individual's biomechanics, clinicians may utilise the PBTs and couple the score with annotations on the observed movement quality and compensations, which could then be compared at various stages post-TKR.

The lack of correlation between the BOA and the PROMs and between PBTs and PROMs pre and at three stages post-TKR supports the idea that objective function may not play a major role in the patient's perceived function and other factors may be involved. This adds to the previous literature suggesting that objective measures of physical function and PROMs may measure different constructs of function and, therefore, they should be utilised together.

Chapter 7: The application of the Cardiff classifier to the assessment of a dynamic knee-loaded activity *in vivo*

7.1 INTRODUCTION

The current work developed from a collaboration between our research group and the research group involving the University of Canberra (Canberra, Australia), the University of New South Wales (Sydney, Australia), and the Australian National University (Canberra, Australia), whose focus was analysing knee osteo and arthrokinematics (identified as 'knee kinematics' in this chapter) pre- and post-TKR *in vivo* measured using dynamic single-plane fluoroscopy and image registration, as part of their recent randomised controlled trial titled "A prospective imaging study of cruciate-retaining, cruciate-substituting and rotating platform total knee replacement in osteoarthritis and healthy ageing: a randomised control trial" (PICKLeS study). A visit from Professor Mark Pickering to our department in January 2020, where the Professor illustrated his research group activity, and the author presented the Cardiff classifier, initiated a mutual interest in combining the knee kinematics *in vivo* measured via dynamic imaging, with the Cardiff classifier ability to assess objective function in people with severe knee OA and post-TKR, by combining several input variables.

While 3D gait analysis with motion capture technologies offers the advantage of measuring the biomechanics of several joints simultaneously, some of its limitations are the soft tissue artifacts and assuming that a fixed relationship exists between skin and underlying bones during dynamic tasks. For these reasons, only joint rotations are extracted following a gait assessment, as 3D gait analysis is not suitable to measure joint translations, for which other technologies are most appropriate, such as computed tomography (CT), magnetic resonance imaging (MRI) or fluoroscopy. The combination of 3D bone models of the joint in question, obtained via CT or MRI, with 2D images obtained via fluoroscopy, called image registration, ultimately allows assessing the joint osteokinematics (i.e., the relative rotation of the bones) and arthrokinematics (i.e., the relative translations of the joint surfaces) in 3D and *in vivo* with six degrees of freedom. Single-plane fluoroscopy of the knee allows recording images in one plane of motion, most often the sagittal plane. However, measurements of out-of-plane translations (i.e., orthogonal to the

plane of the image intensifier) are susceptible to higher degrees of inaccuracy (Scarvell et al. 2010; Tsai et al. 2010; Acker et al. 2011). Some research groups were able to overcome this limitation by utilising bi-plane fluoroscopy, allowing them to record knee kinematics from different perspectives, with an overlapping view of the joint (Yue et al. 2011; Li et al. 2013; Kefala et al. 2017).

Although fluoroscopy is widely used clinically to guide surgical procedures requiring real-time feedback, image registration is mainly utilised in research due to the timing and expertise needed to process the data. Fluoroscopy entails exposure to low-dose ionising radiation (i.e., X-rays), and the ability to inspect only one joint at a time with a limited capturing area. Nevertheless, this technique combined with image registration offers the advantage of observing the joint kinematics *in vivo* with a high level of accuracy (except for out-of-plane translations when utilising single-plane fluoroscopy, as previously mentioned). This information has been extremely valuable in understanding the kinematics of a native knee (Freeman and Pinskerova 2005; Li et al. 2013; Kefala et al. 2017; Galvin et al. 2018), the differences between various TKR designs (Okamoto et al. 2014; Hamai et al. 2015; Murakami et al. 2018; Lynch et al. 2021), and in showing that although certain motions are similar to the natural knee, knee implants do not allow to fully reproduce the knee kinematics of a native knee not affected by pathology (Yue et al. 2011; Mills et al. 2023).

Most of the studies look at each of the knee rotations and translations data separately and report knee kinematics as a function of knee flexion, describing translations or rotations every 10 degrees of knee flexion, to then compare the results at each data point between groups (Fantozzi et al. 2006; Nishio et al. 2014). Such approach aims to identify patterns in the translations and rotations of the knee. It may be argued that, as done in the studies in Chapters 4-5-6, and as demonstrated by the study of Lynch et al. (2020), utilising principal component analysis (PCA) on the *in vivo* knee kinematics waveforms, could be a valuable technique to achieve similar results, showing a pattern of variation between groups. The combination of PCA and the Cardiff classifier would allow to evaluate the knee function as a whole, rather than looking at patterns of rotations and translations separately, and aid in understanding the overall degree of recovery of knee kinematics *in vivo* post-TKR.

To the best of the author's knowledge, no studies combined knee kinematics *in vivo* in a single measure of knee function. The assessment of the knee kinematics

in vivo during step-up-down has been previously carried out in our research group, but only in small cohorts of healthy volunteers (three to five) (Whatling 2009; Williams 2018), and on a unique population of twenty-seven knees who received a mal-aligned TKR (Williams et al. 2020). In these studies, the authors described the single knee osteo and arthrokinematics using discrete values or ranges of movement (Whatling 2009; Williams 2018). For the current work, the small number of healthy participants from the previous studies at Cardiff University (Whatling 2009; Williams 2018) would have made it difficult to explore the application of the Cardiff classifier to *in vivo* knee kinematics using this dataset. Furthermore, the unique cohort of patients available from the work of Williams et al. (2020) would have made it challenging to relate the outputs to the general population who received a TKR.

The collaboration with Professor Mark Pickering and Dr Joseph Lynch who were involved in the PICKLeS study, offered the opportunity to access a sufficiently large dataset of NPs, and patients with end-stage knee OA pre- and two years post-TKR which could be utilised for the aims of this investigation. The *in vivo* assessment of knee kinematics for the PICKLeS study included standing from sitting, kneeling and step-up-down. For this exploratory study, only step-up-down data was made available for analysis, due to other researchers already being involved in analysing or having carried out earlier research on kneeling and standing from sitting for other projects. Step-up can be compared to stair ascending, a common everyday life activity that patients usually find challenging in the presence of knee OA (Hensor et al. 2015). Following the application of the Cardiff classifier to a sub-set of the PICKLeS study data on knee kinematics, the main outcome variable utilised in this study was the resulting BOA, and it was meant as a representation of “knee function”, therefore, the two terms will be used interchangeably from now onwards.

Aim 1: to explore the application of the Cardiff classifier to *in vivo* knee kinematics measured using image registration and dynamic single plane fluoroscopy during a step-up movement, pre- and post-TKR.

The following research question was formulated: can the Cardiff classifier discriminate between patients with end-stage knee OA and NPs when utilising *in vivo* knee kinematics data during a step-up?

Hypothesis 1.1: post-TKR, patients’ knee function improves but is not comparable to the BOA of NPs.

As seen in the Chapter 6, several researchers were interested in understanding the relationship between what patients can do objectively during PBTs, and what they perceive they can do (i.e., PROMs). Similarly, there is a growing interest in determining whether certain patterns of *in vivo* knee kinematics may be associated with PROMs, but the literature available is limited to a few studies (Fantozzi et al. 2006; Nishio et al. 2014; Warth et al. 2017; Galvin 2019; Van Onsem et al. 2020). A recent investigation including fifty-six participants from the PICKLeS study found that the degree of knee flexion during kneeling explained 51.4% of the OKS variance in people with end-stage knee OA, where a larger knee flexion was associated with a higher OKS (Galvin 2019). Previous investigations looked at differences in PROMs between groups achieving different discrete kinematics (Nishio et al. 2014; Van Onsem et al. 2020) or correlated a discrete value of knee kinematics to a PROM score (Fantozzi et al. 2006; Kage et al. 2021). While this could be valuable, it requires several statistical comparisons and it may be argued that a single, discrete knee kinematics value may not capture the overall knee function. The novel application of the Cardiff classifier to *in vivo* knee kinematics would help to combine the knee rotations and translation, offering a more comprehensive assessment of the knee kinematics of people with severe knee OA and post-TKR. Moreover, the BOA value obtained via the classifier could be utilised for further statistical analysis and to aid in understanding the relationship between overall knee kinematics and PROMs.

Aim 2: to determine whether the BOA pre- and post-TKR is correlated with PROMs. It was difficult to hypothesise on the strength of the correlations due to the lack of previous similar research. Based on the earlier findings from our colleagues at the University of Canberra (Galvin 2019), the following hypotheses were formulated.
Hypothesis 2.1: pre-TKR, there is a significant association between OKS and BOA
Hypothesis 2.2: post-TKR, there is a significant association between OKS and BOA
Null hypothesis 2.3: post-TKR, there is not a significant association between BOA and satisfaction levels. This hypothesis originated from the demonstration that satisfaction is highly influenced by factors that are not related to function, as discussed in Chapter 2 (Literature Review).

Hypothesis 2.4: post-TKR, there is a significant correlation between the BOA and the patient's perception of "the knee feeling like a normal knee".

7.2 METHODS

The data for the current longitudinal, retrospective study was part of a wider randomised controlled trial, the PICKLeS study, approved by the Australian National University Ethics Committee (2017/354) and the ACT Health Human Research Ethics Committee (ETH.4.11.071).

Before the author became involved in the research, all data had been collected and some of it had been processed (more details in the sections to follow). The author's contribution to the project consisted in:

- Performing image registration for part of the step-up-down data (pre-TKR patients)
- Writing a MATLAB code (MathWorks Inc., Natick, Massachusetts, US) to process and extract the kinematic data to allow utilising it with the existing MATLAB code performing PCA (in preparation to utilise the Cardiff classifier)
- Using the Cardiff classifier to analyse knee kinematics *in vivo* pre- and post-TKR
- Examining how the *in vivo* knee function during step-up, evaluated via the Cardiff classifier, related to PROMs.

7.2.1 Participants

Participants with severe knee OA were recruited from one orthopaedic surgeon patient's waiting list for a TKR. A convenience sample of NPs was recruited from the local community via posters, brochures, word of mouth and direct contact with the researchers involved with the PICKLeS study. People who showed an interest in participating were contacted by phone. During the first contact, the potential participants were screened for inclusion and exclusion criteria (Table 38) and given details about the study. If the potential participants were eligible and agreed to take part in the study, they were sent a letter or an email containing the study information sheet. Participants were given three days to a week to read the information sheet and were contacted again by phone to allow clarifying any questions about the study and to offer an appointment provided they agreed to take part in the study. On the assessment day, participants were offered to discuss any additional questions about the study with the researcher. If the participants showed

to have understood the study requirements and were satisfied to proceed, the consent form was signed. Written informed consent was obtained from all participants before data collection.

Table 38 Participants' inclusion and exclusion criteria

	Inclusion criteria	Exclusion criteria
Severe knee OA participants	• Between 40-80 years old • Presence of knee osteoarthritis (confirmed through knee x-rays review) • In a waiting list to receive a TKR • Suitability to receive either a cruciate-retaining fixed bearing, posterior-stabilised fixed bearing or cruciate-retaining mobile bearing knee implant • Ability to return for a 12- and 24-month post-surgery follow-up	• Pregnancy • Issues with understanding and unable to give informed consent • Psychosocial factors preventing giving consent or preventing completing the requirements of the study, including an activated "enduring power of attorney" • Body Mass Index > 38 (extremely obese) • Presence of knee OA solely in the knee lateral compartment • Maximum knee flexion < 90˚ • Knee fixed flexion contracture ≥ 10˚ • University of California, Los Angeles activity scale score ≤ 2 (= wholly inactive, dependent on others or mostly inactive/severely restricted to the minimum of activities of daily living) • Listed for knee replacement revision surgery • The reason for the total knee replacement is a pathological fracture • Metastasis
Non-pathological volunteers	• Over the age of 18 years • One pain-free knee	• Pregnancy • Issues with understanding and unable to give informed consent • Psychosocial factors preventing giving consent or preventing completing the requirements of the study, including an activated "enduring power of attorney" • Metastasis • History of knee injury or osteoarthritis

The PICKLeS study required at least eighty NPs and sixty people with severe knee OA and post-TKR. Patients were randomly allocated to receive one of three knee designs with the same femoral geometry, cemented cruciate-retaining rotating platform, cruciate-retaining fixed-bearing or posterior-stabilised fixed-bearing implant. Patients were blinded to which implant they received. The random allocation allowed to have similar proportions of males and females in each group. For the current exploratory study, it was initially decided to include only the twenty knee OA patients

who received a cruciate-retaining, fixed-bearing implant design and attended both the pre- and two-year post-TKR assessment, and an equal number of randomly selected NPs. However, due to the poor quality of the fluoroscopic images, it was not possible to complete the data processing (i.e., image registration) for nine knee OA patients pre-TKR. These patients moved relatively fast during the recorded trials, producing motion artifacts that made it impossible to identify the bone's silhouette during image registration. Additionally, one patient was lost at follow-up (due to needing revision surgery). This caused the sample size to decrease drastically (n = 12) and therefore, six patients with a different implant design were randomly selected and included in the analysis, for a total of eighteen patients pre- and post-TKR. As previously stated, the current study aimed to evaluate the application of the Cardiff classifier to the knee kinematics of step-up, and the correlation between the classification and PROMs, and not to explore the differences in knee kinematics in varying knee implant designs. It was deemed appropriate to include patients who received a different knee implant, considering that a Cochrane Review demonstrated that the use of fixed or mobile bearing implants seems to produce similar results in terms of quality of life, functional and clinical outcomes (Hofstede et al. 2015).

7.2.2 Data collection protocol

Data collection took place between February 2012 and November 2018 at Canberra Hospital (Canberra, Australia). Knee imaging was performed at the Medical Imaging department whilst the collection of subjective outcome measures and PBTs data (not utilised in the current study and, therefore, not described further) was completed at the Trauma and Orthopaedic Research Unit. NPs completed only one data collection session. Patients' data were collected at three time points, within one month pre-surgery, 12 months, and 24 months post-TKR. Each session lasted about 90 minutes, including, in the following order, the fluoroscopic assessment, Computed Tomography (CT) scans (only pre-surgery), clinical examination (knee ROM, knee anterior-posterior and mediolateral stability, knee extension lag, knee malalignment and flexion contracture – all tests necessary to fill out the American Knee Society Score), PBTs and PROMs. For the aim of this study, only pre and 24 months post-TKR patients' data were analysed, where 24 months data was not available, the 12 months post-TKR was utilised.

Several PROMs were collected at each assessment for all participants, including OKS, the University of California and Los Angeles Activity Scale (a single-item score assessing patient's activity level and independence), the American Knee Society Score, Pain Visual Analogue Scale (VAS), Assessment of quality of life, Functional Comorbidity Index. Only the relevant PROMs for this study will be presented. Regarding the OKS, the details were described in section 3.5.2. In brief, the OKS assesses knee-related pain and function with twelve questions, the total score ranges from zero, indicating extreme knee-related issues, to forty-eight, no issues. Patients were required to complete a VAS satisfaction at follow-up, and a VAS scale to measure "How much does your knee feel like a normal knee?" (Figure 34). The VAS is a 100 mm horizontal or vertical line, where each end identifies the extreme of the assessed factor (pain, satisfaction, etc.).

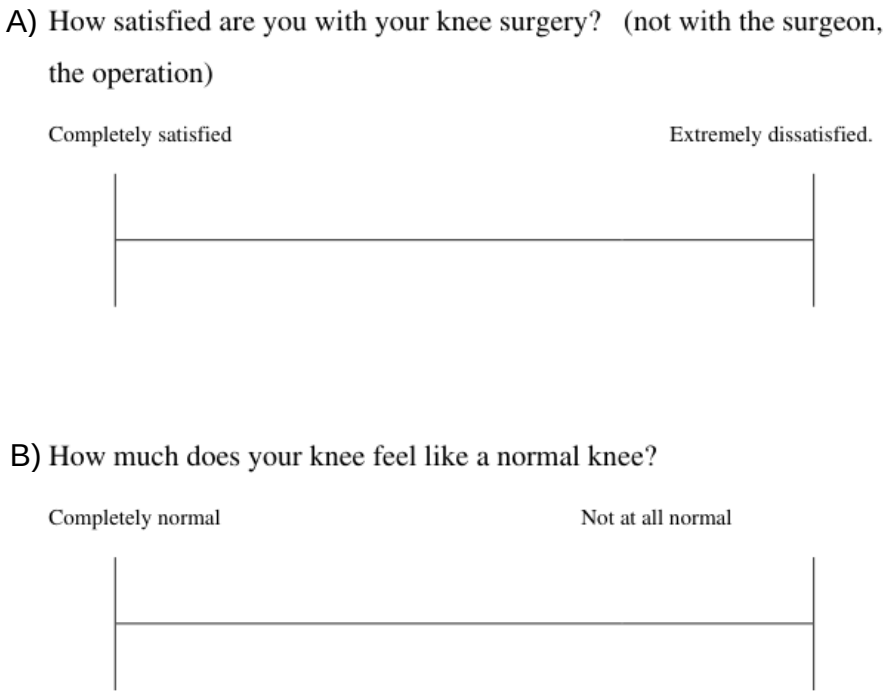


Figure 34 A) VAS satisfaction; B) VAS capturing the patient's perception about their knee.

For instance, in the VAS Pain, the left end represents "No pain", while the right end represents "The worse pain imaginable". The subject marks the line on the point they believe best captures how they feel about their current situation. The score is the distance in mm, measured with a ruler, from the left end to the mark

(range 0 – 100 mm). The VAS Pain is valid (Joyce et al. 1975), with good reliability in people with rheumatic conditions and is responsive to change (Hawker et al. 2011).

7.2.3 Single-plane fluoroscopy assessment

The AXIOM-Artis MP single plane fluoroscope (Siemens, Munich, Germany) (Figure 35) was utilised to assess the knee kinematics during standing from sitting, kneeling and step-up-down. For this study, only step-up was considered. A fluoroscope allows the recording of a series of X-rays of a joint's dynamic task in vivo in two dimensions (2D). The fluoroscope recorded at 30 Hz, producing 1024 x 1024 pixel images with 12bits/pixel. The X-ray source and image intensifier (screen size: 280 mm) were 1200 mm apart. All participants wore a lead apron on the upper body to protect them from radiation for the duration of the fluoroscopic assessment. Participants started with their knee between the X-ray source and the image intensifier and their foot resting on a 250 mm high step, which is a common rise utilised in various previous studies (Moonot et al. 2010; Tibesku et al. 2011; Kuroyanagi et al. 2012).

The subject's upper body was not constrained, and handrails were available for light support in front of the participant, who was instructed to use it only for balancing purposes and not to aid the stepping-up-down. Subjects were asked to maintain the contralateral leg extended and away from the X-ray emission area. The lateral aspect of the knee was positioned on the side of the image intensifier and images were recorded in the knee sagittal plane. Participants were asked to keep the foot on the step straight and familiarised with the step-up-down task before recording the trial, with the aim to minimise radiation exposure during the following data recording (maximum three recordings). Subjects were recorded during one only repetition of the step-up to full knee extension and step-down until the contralateral foot was again in full contact with the floor and were paced by the main researcher ("start, two, three, down, two, three") to standardise the participant's movement speed during the task and to limit blurring of the images recorded. Only the affected knee was examined in the patients, while a randomly selected side was considered in NPs.

The fluoroscopy system was equipped with a curved panel detector (i.e., image intensifier), which naturally produces a pincushion distortion of the images

(the image is “pinched” in the centre as if a pin had been pushed through it, and the distortion is more pronounced at the edges of the image). Another source of image distortion is the sigmoidal distortion, caused by the Earth’s magnetic field deviating the electrons in the image intensifier (Gutírrez et al. 2008). To rectify these, a calibration frame made of two acrylic Perspex® (Perspex Distribution Ltd, UK) sheets (400 mm x 400 mm) 200 mm apart was utilised. Each of the sheets had 1 mm tantalum beads embedded in a regular grid, 20 mm from one another. At the end of the session, a recording of the calibration frame was performed to aid correction of the image distortion during data processing at a later stage.

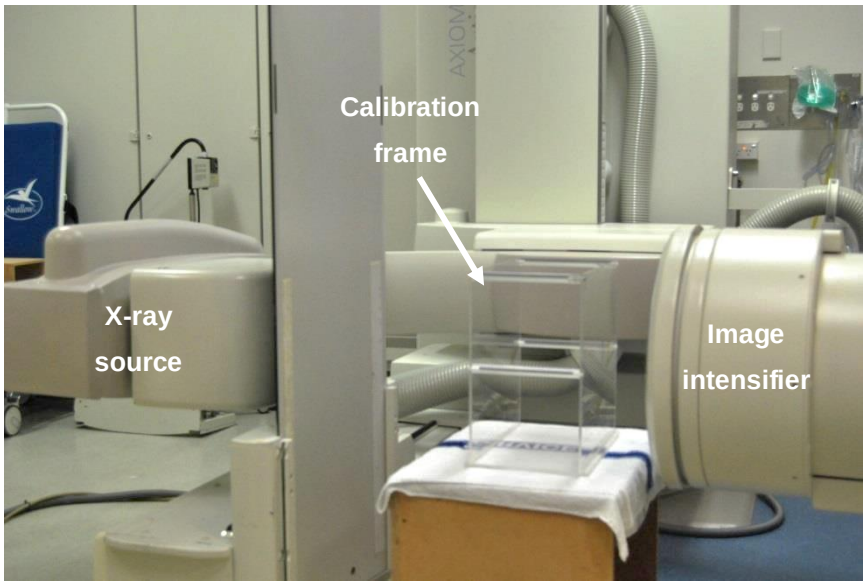


Figure 35 The AXIOM-Artis MP with the calibration frame. X-ray source: X-ray generator, powering the X-ray tube, which contains a filament that once heated, produces electrons that are accelerated towards a tungsten component which releases X-rays, the Collimator determines the size of the X-ray beam. Image intensifier: transforms X-rays into light, which is recorded by a camera, where the light is transformed into an electric signal sent to a monitor (not shown in the picture).

7.2.4 CT scanner data acquisition

To reconstruct the six degrees of freedom of the native knee kinematics during data processing, 2D fluoroscopic data, recording the joint kinematics in a dynamic task, had to be combined with 3D bone scans produced in a static condition, which were obtained via CT scans. CT allows to record a series of images

that can be reconstructed (via image segmentation, discussed in section 7.2.5) to create a 3D model of the bone, and it is the gold standard for this purpose (Häller et al. 2021), offering clear images of the cortical bone, surrounding and underlying structures (Johnston et al. 2009). Each participant lay in a supine position inside a Toshiba Aquilion 16 spiral CT scanner (Toshiba Medical Systems, Tokyo, Japan) with their knee surrounded by the calibration frame described earlier (Figure 36).

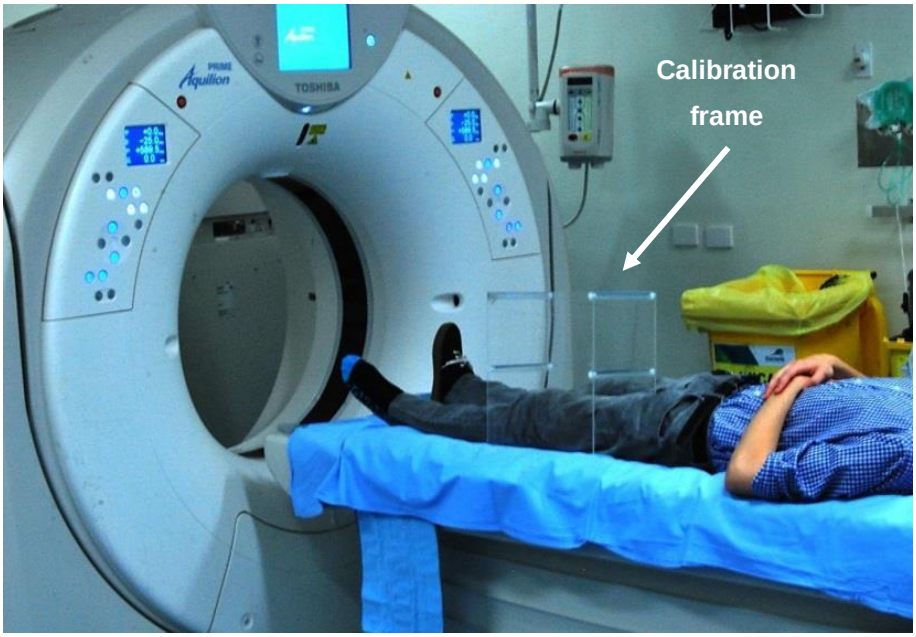


Figure 36 The Toshiba Aquilion 16 spiral CT scanner, a participant with the calibration frame over the examined knee.

The CT scanner rotated around the subject and recorded a series of 3D X-ray images starting from approximately 150 mm below the knee joint line and terminating 150 mm above it (about 20 seconds recording). This allowed to produce a 3D scan of the knee joint bony structures. The scan slices were 1 mm thick and had a 512 x 512 voxels resolution with spatial dimensions 0.625 x 0.625 x 0.5 mm^3 and 16 bits/pixel.

Due to the imaging technologies employed, participants were exposed to ionising radiation. The total radiation dose for the three data collection sessions (pre and two follow-ups post-TKR) including both fluoroscopy and CT scans was less than 8 millisieverts (mSv), which is roughly three times the dose of annual radiation a person is exposed to in the UK (2.7 mSv) (UK Health Security Agency 2011). The Director of Medical Imaging at Canberra Hospital evaluated the appropriate radiation

dose for the study, based on the Australian Radiation Protection and Nuclear Safety Agency guidelines.

7.2.5 Data processing

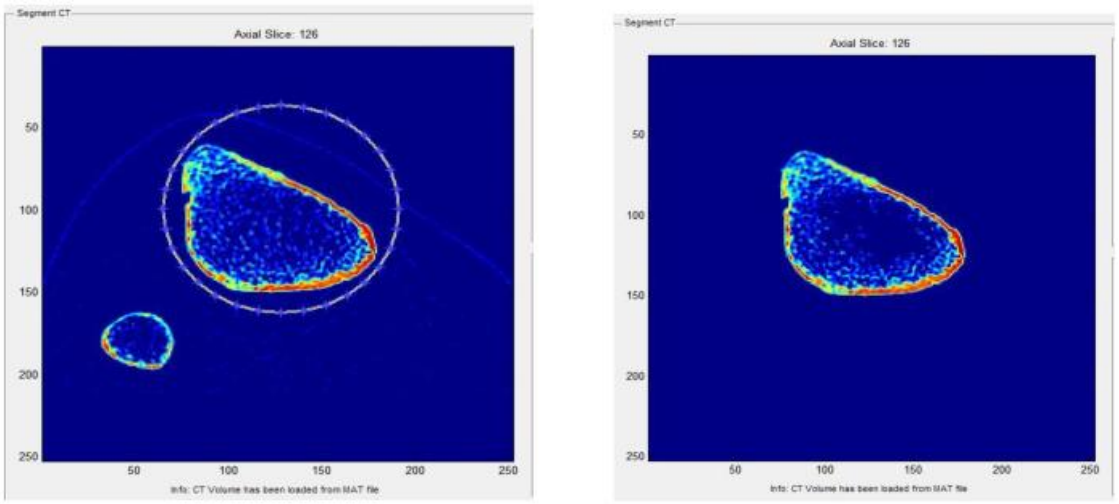
Previous work at the University of Canberra (Galvin 2019; Galvin et al. 2019) and the Australian National University (Lynch et al. 2019; Lynch et al. 2020; Lynch et al. 2021) based on data from the PICKLeS randomised controlled trial, utilised Orthovis (MATLAB, MathWorks Inc., Natick, Massachusetts, US), a bespoke package allowing to perform image segmentation, registration and the calculation of six degrees of freedom joint kinematics. Orthovis had been extensively utilised to analyse data during deep kneeling (recorded utilising the same equipment mentioned in earlier paragraphs) by Dr Catherine Galvin (University of Canberra), and Dr Joseph Lynch (Australian National University). Therefore, in continuity with the earlier work mentioned above, Orthovis was utilised by the author for data processing for the current project following the training provided directly by Professor Mark Pickering (University of New South Wales), who created Orthovis. The accuracy of Orthovis was found to be higher in the sagittal plane (i.e., the plane in which the 2D images were recorded) and it was 0.3 degrees for rotations, 0.2 mm for anterior-posterior and superior-inferior (i.e., compression-distraction) translations, while the accuracy was lower out-of-plane rotations and translation, and corresponding to 0.5 degrees for rotation, and 0.9 mm mediolateral translation (Scarvell et al. 2010; Akter et al. 2014).

Segmentation

Image segmentation had already been completed for all participants at the time of the current study, with no need for further input by the author. Segmentation involved identifying the contour of the bone for each of the images (i.e., slices) that had been acquired via CT scans, and eliminating the surrounding soft tissues and artifacts.

The CT scan was loaded into Orthovis, and the area of interest on the CT scan image was selected by overlaying a rectangle over it, in the three planes, to select the bone volume to be utilised for the segmentation. Manual segmentation was performed within Orthovis, where an overall segmentation threshold to define

lower and upper grey scale boundaries was defined, and pixel connectivity limits were set to eliminate volumes not needed for further analysis (i.e., the soft tissues surrounding the bone, calcifications, and other bones). Then, a polynomial (with the option to modify the number and size of the nodes) was fit around the bone silhouette, allowing it to eliminate the tissues outside it (Figure 37). The segmentation was then refined manually with a digital eraser. All slices were combined to create the bone model. The femur, tibia and fibula together, and the tibia without the fibula were segmented with this approach.



A B

Figure 37 Screenshots of Orthovis interface during segmentation of the tibia. A) Tibia and fibula, the polynomial surrounding the tibia. B) Final segmentation.

Image registration and kinematics calculations

Image registration is the process of fitting the 3D bone segment obtained via image segmentation of CT scans (or a TKR implant model) over the 2D images obtained via fluoroscopy, and it was carried out with the following procedure:

1. *Creation of a Digitally Reconstructed Radiograph (DRR)*: this was a virtual radiograph of the segmented bone model obtained in the previous step, providing the 3D bone segment that could be registered on the 2D fluoroscopy images during step 4 below. Creating a DRR was not necessary for post-TKR assessments as Computer-Aided Design (CAD) models were utilised for the image registration
2. *Correction of the fluoroscopy image distortion*: the recording containing the calibration frame was loaded (Figure 38A), and a grid was manually overlaid

to find the best match between the tantalum beads on the distorted image and the circles in the grid (Figure 38B) (the beads' real position relative to each other was known by the software); the difference between the known position of the beads and that of the grid was utilised automatically by Orthovis to calculate the coefficients of a polynomial function to correct the image distortion

3. *Definition of the bone/implant models coordinate systems*: the CT scan was displayed in the frontal, sagittal and transverse planes, and the user had to manually identify anatomical landmarks on the bone segment or implant, which would allow determining axes to create the segment right-handed coordinate system. The axes were defined based on the recommendations of Grood and Suntay (1983), described in Figure 39, and Figure 40, respectively. The TKR CAD models were provided by the manufacturer, the coordinate system was defined according to the work of Grood and Suntay (1983), and described in Figure 41 and Figure 42, respectively.

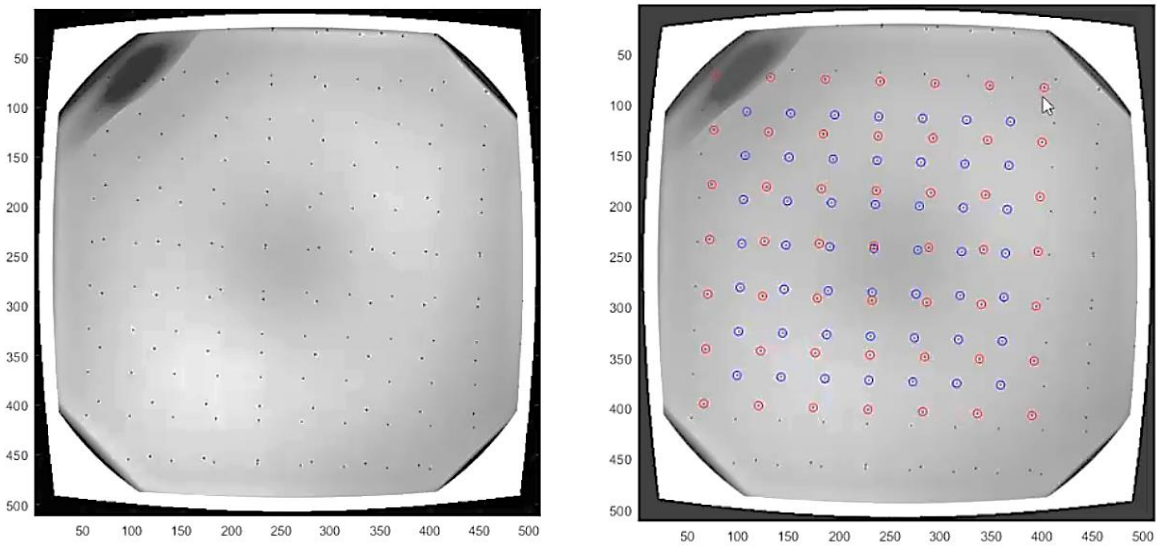


Figure 38 Screenshot of Orthovis interface during distortion correction. A) Calibration frame recording. B) Grid over the tantalum beads for image distortion correction.

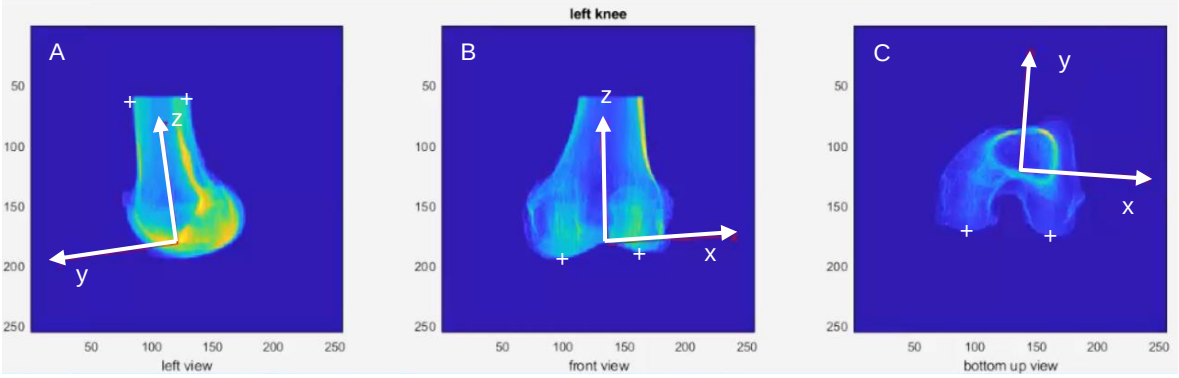


Figure 39 Femur coordinate system (left knee). Origin: sagittal plane (A): distal point of the femoral shaft; frontal plane (B): proximal point of the intercondylar notch.

+: the landmarks are manually selected by the user.

z-axis (superior-inferior): a vector aligned with the mechanical axis of the femur (between the intercondylar notch and the femoral head) defined by selecting the intercondylar notch and the most distal points on the condyles ("+" in B), the most distal point on the femoral shaft, and the midpoint between the proximal cortical bone anterior and posterior limits ("+" in A)

y-axis (anterior-posterior): the cross product between the z-axis, and the most posterior points on the condyles ("+" in C: transverse plane, view from the bottom-up)

x-axis (mediolateral): cross product between the z and y axes, pointing laterally in a left knee, and medially in a right knee (right-handed coordinate system).

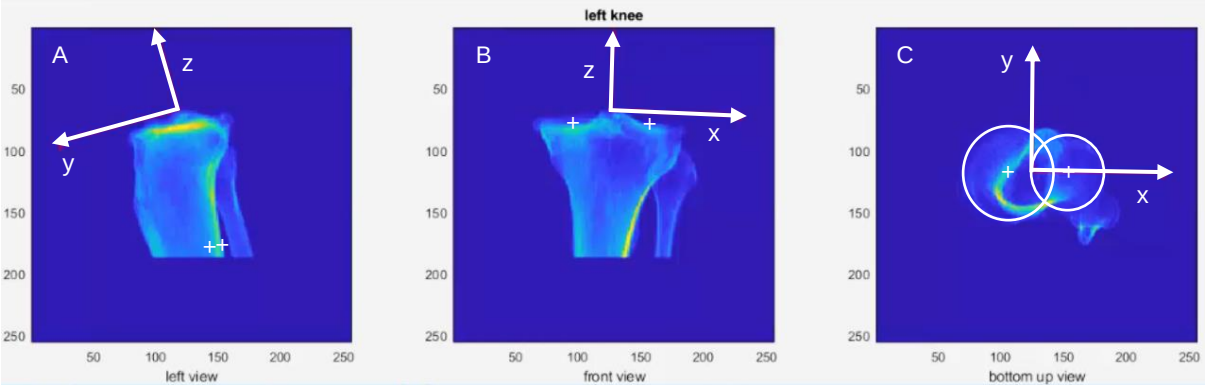


Figure 40 Tibia coordinate system (left knee). Origin: sagittal plane (A): proximal eminence of tibial spines; frontal plane (B): mid-point of the tibial spines; transverse plane (C): midpoint of the tibial plateaus' centres.

+: the landmarks manually selected by the user.

z-axis (superior-inferior): a vector coinciding with the tibial mechanical axis (between the tibial spine eminences, and the centre of the ankle), defined by selecting the intercondylar eminence (A), the midpoint between the tibial spines (B), and the cortical bone limits posterior to the tibia ("+" in A).

y-axis (anterior-posterior): cross product between the z-axis and the line connecting the tibial plateaus centres ("+" in B, C).

x-axis (medio-lateral): cross product between the z and y axes, pointing laterally in a left knee, and medially in a right knee (right-handed coordinate system).

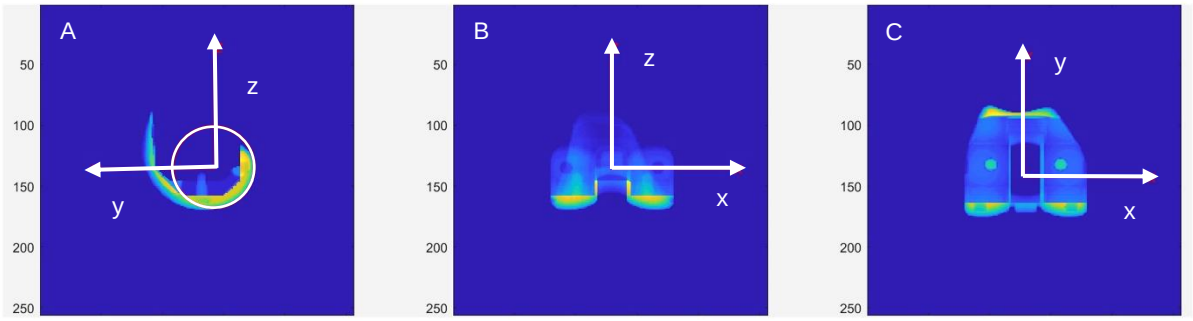


Figure 41 Coordinate system for the femoral implant component (right knee, *cruciate retaining*). Origin: sagittal plane (A): centre of a circle fitting the condyles; frontal plane (B): mid-point of the anterior intercondylar notch; transverse plane (C): mid-point between the anterior and posterior aspects of the intercondylar space on the implant.

x-axis (medio-lateral): by fitting a cylinder to the posterior and distal aspects of both femoral condyles, the x-axis was the axis of the cylinder. Pointing laterally in a left knee, and medially in a right knee (right-handed coordinate system).

z-axis (superior-inferior): perpendicular to the transverse flat surface of the femoral component and pointing proximally (selected by identifying the midpoint of the intercondylar notch.

y-axis (anterior-posterior): cross product of the x and z-axes.

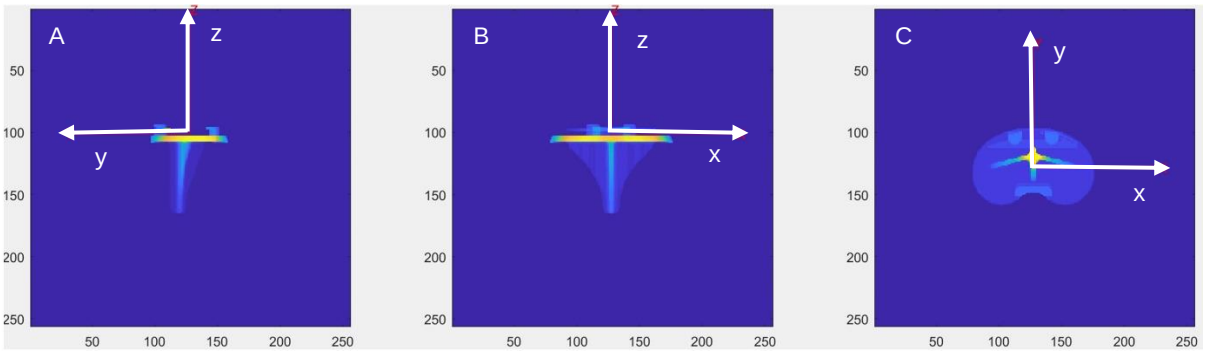


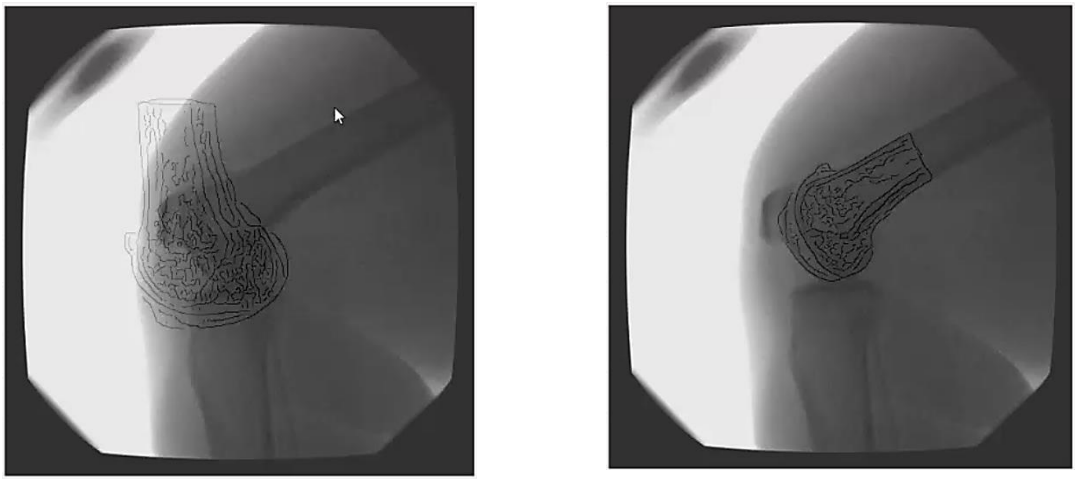
Figure 42 Coordinate system for the tibial implant component (right knee, *cruciate retaining*). Origin: sagittal and frontal planes (A, B): mid-point of the tibial plate; transverse plane (C): mid-point between the anterior and posterior aspects of the tibial plate.

x-axis (mediolateral): vector parallel to a line across the tibial baseplate (B), bisecting the anterior and posterior baseplate (C). Pointing laterally in a left knee, and medially in a right knee (right-handed coordinate system).

z-axis (superior-inferior): coincident with the axis of the tibial stem (also perpendicular to the tibial tray) and pointed superiorly.

y-axis (anterior-posterior): cross product between x and z axes.

4. *Image registration*: the fluoroscopy recording was loaded into Orthovis, where it could be visually inspected for data quality verification. The first frame with good image quality (i.e., clear contours of the bones) at the beginning of the recording was utilised to start the registration. Image registration had to be completed for the femur, before registering the tibia. The DRR was loaded on the first fluoroscopy frame and was manually manipulated (i.e., rotated and translated over the 2D image) by the user to best match the bone silhouette to the fluoroscopy image, and this first frame was registered (Figure 43).



A B

Figure 43 Screenshot of Orthovis interface showing the DRR loaded on the 2D image (A), and the DRR aligned with the 2D image (B).

Then, the automatic registration was started, and no further input was required by the user. This was an iterative process registering 3D bone/implant models to 2D images frame by frame until the last frame of the fluoroscopy recording (20 seconds to a minute to register each frame). In certain instances (i.e., blurred or dark 2D images), the registration was not adequate, with a mismatch between the 3D model and 2D images, which could be manually corrected. The automatic registration was completed via the following standard steps by Orthovis:

- Geometric transform: a rigid transformation (i.e., only allowing translations and rotations, e.g., not shear) fitted the DRR to the fluoroscopy by rotating and translating the DRR

- Similarity measure: based on an edge detecting algorithm finding features such as edges in the projection of the 3D bone model and the 2D image; it calculated the change needed to obtain the best alignment between the DRR 3D position at the current frame, and the 2D fluoroscopy image
- Optimisation: it aligned the edges of the DRR to those of the 2D image

3D kinematics were calculated with the Grood and Suntay (1983) convention, following the ISB recommendations (Wu and Cavanagh 1995), whereby flexion-extension occurs about the medio-lateral x-axis of the proximal segment (femur), internal-external rotation occurs relative to the distal-proximal z-axis of the distal segment (tibia), while adduction-abduction occurs about a floating axis which is the cross product between the femur x-axis and the tibia z-axis. Translations were described as the movement of the femur coordinate system's origin relative to the tibia's coordinates system origin. The tibia was the reference system; therefore, knee kinematics were described as the femur rotations and translations relative to the tibia. Flexion, adduction, internal rotation, anterior translation and superior translation (i.e., knee distraction) had a positive sign for both the left and right knee. Table 39 shows the variables extracted via Orthovis.

Once the image registration of both the femur and tibia was completed, it was possible to visually verify the reconstructed kinematics in a 3D animation of the task and to visualise the graphs describing the joint rotations and translations. This allowed to check for potential errors during the registration process that could be rectified. The knee kinematics data were exported in an Excel Spreadsheet.

Step 1 of the abovementioned procedure had already been completed for all participants' data and no further input was needed from the author, while the next steps were taken by the author for data processing. The author only processed pre-TKR data, whereas post-TKR and NPs data had already been registered at the time of the current project.

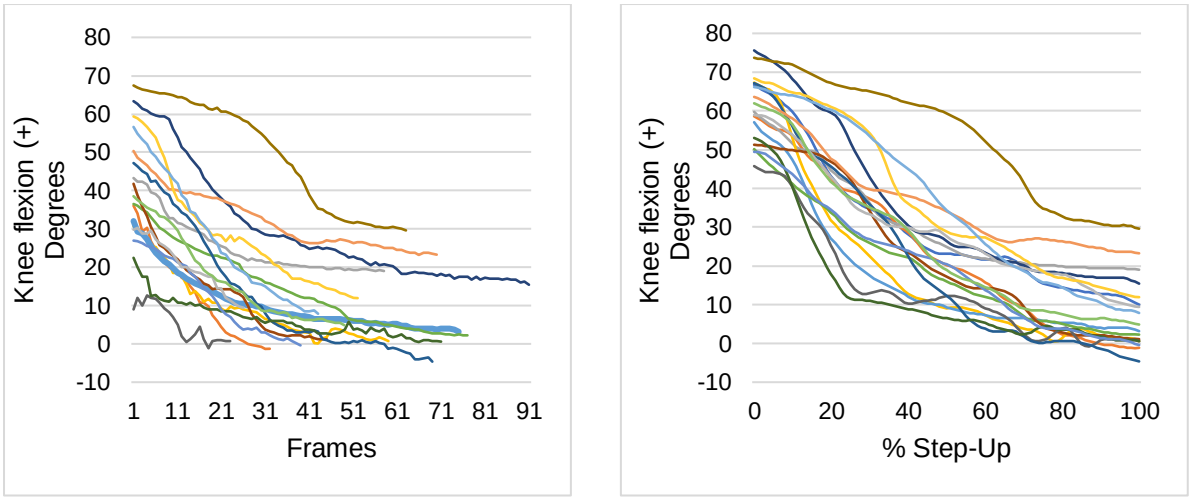
Determining phases of the activity

The task needed to be divided into phases to allow further data processing. In order to do this, an algorithm previously written in "R" (R Core Team, 2018) by Dr Catherine Galvin was utilised. In brief, the code calculated the first derivative of knee

flexion against time (i.e., angular velocity), and determined the frame whereby the angular velocity was zero (a change in the curve direction); these frames corresponded to the start or end of the phases. Five phases were identified during the step-up-down task, including phase 1 (zero knee angular velocity, starting position with one foot on the step), phase 2 (stepping-up), phase 3 (knee extension, zero angular velocity), phase 4 (stepping-down), phase 5 (zero knee angular velocity, foot on the step, weight on both lower limbs). The output was an Excel spreadsheet containing all patients in one column, and their phases in the next column, with the knee kinematics in separate succeeding columns. This meant that before additional analysis could be carried out, the data for each of the participants had to be extracted.

Data filtering and normalisation

The author developed a code in MATLAB (MathWorks Inc., Natick, Massachusetts, US) (Appendix F), which separated the data for each participant, extracted only the relevant phases (phase 2, stepping up) and relative kinematics for this study, filtered and normalised the data to % of step-up. Followingly, the description of the procedure utilised in the code. Joint kinematics derived via fluoroscopic assessment is normally a combination of the real rotations and translations and signal interference due to measurement or data processing inaccuracy. Joint movements tend to be smooth and continuous and the presence of measurement errors due to low-quality fluoroscopic images which lead to image registration imprecisions produced an output that had a noise which is not a representation of the real joint motion (Figure 44A). When considering a kinematic waveform, the high-frequency content tends to be linked to signal noise, while the low-frequency content is mostly related to the joint kinematics (Bifulco et al. 2012). The signal interference can be decreased using a low pass filtering technique, which eliminates frequencies over a certain threshold. Therefore, a fourth-order, zero-phase, Butterworth filter, which is largely utilised when analysing joint kinematics (Sinclair et al. 2013) with a cut-off frequency of 4 Hz was utilised to smooth the knee kinematics data. The task was then time-normalised to 101 time points, from 0 to 100% of the step-up task (Figure 44B). The outputs for each joint rotation and translation were saved in separate Excel Sheets.



A B

Figure 44 Example of patients' knee flexion-extension angles pre-TKR. A) Raw data; B) Filtered and task-normalised data

Table 39 Knee variables analysed in the current study and included in the Cardiff Classifier

Knee kinematics during step-up	**Calculated by Orthovis**	**Filtered and normalised for further analysis**	**Utilised within the Cardiff Classifier**
Flexion-extension angle	✓	✓	✓
Adduction-abduction angle	✓	✓	✓
Internal-external rotation angle	✓	✓	✓
Compression-distraction	✓	✓	✓
Medio-lateral translation	✓	✗	✗
Anterior-posterior translation	✓	✓	✓
Tibia contact points on the femur	✗	✗	✗

7.2.6 Classification of joint function

PCA and the Cardiff classifier were utilised to reduce data dimensionality and classify the data, respectively, as described in section 3.7. In brief, PCA was performed on the flexion-extension, adduction-abduction, internal-external rotation, anterior-posterior, and compression-distraction translation waveforms to obtain PC scores for each participant. Medio-lateral translations were not included in the analysis due to the higher degree of error in this movement (Scarvell et al. 2010; Akter et al. 2014). PC reconstruction was carried out to interpret the biomechanical meaning of each component, as described in section 3.7. The first three PCs for

each of the abovementioned variables and each participant were retained for further analysis, for a total of fifteen knee kinematics features. Therefore, for each participant (NPs and pre-TKR), PC scores for the first three PCs of the variables shown in Table 39 were utilised to train the Cardiff classifier in discriminating between OA and NP knee kinematics. All fifteen features were utilised to classify participants pre- and post-TKR and to obtain the Belief values for each participant (i.e., BOA, BNP, Uncertainty). The accuracy of the classifier was verified with leave-one-out cross-validation. The out-of-sample classification simplex plot was produced, and the out-of-sample classification accuracy was calculated as the proportion of out-of-sample subjects out of the total whose classification matched their class. The sensitivity (Eq. 27) and specificity (Eq. 28) of the classification based on the full dataset were reported.

7.2.7 Statistical analyses

All analyses were performed in IBM SPSS Statistics (version 27, IBM, Armonk, NY, US). Normality distribution was determined via the Shapiro Wilk's test (p-value > 0.05: parametric distribution) and inspection of Q-Q plots. For all statistical tests, the significance level was $\alpha = 0.05$, unless otherwise specified. To compare baseline characteristics between patients and NPs, continuous variables were examined with an independent t-test or its non-parametric equivalent (Mann-Whitney U test), ordinal variables (i.e., OKS) were examined with the Mann-Whitney U test, and categorical variables (i.e. sex) with Pearson's Chi-Square test.

Aim 1 – Application of the Cardiff classifier to *in vivo* knee kinematics during a step-up

To verify if differences in the patients' BOA, BNP and Uncertainty values pre to post-TKR were significant, a paired samples t-test or its non-parametric equivalent, the Wilcoxon signed-rank test, was run. To compare the Belief values post-TKR to those of NPs, an independent samples t-test or its non-parametric equivalent (Mann-Whitney U test) was utilised. A Bonferroni correction was applied to the significance level to account for multiple comparisons (Belief values pre- VS, post-TKR, Belief values post-TKR VS. NPs = 2 comparisons); considering $\alpha = 0.05$

for each statistical test before the adjustment, the Bonferroni-adjusted $\alpha = 0.05/2 = 0.025$ (Sedgwick 2014).

Aim 2 - Correlations between knee function and PROMs

A correlation analysis was conducted to determine the relationship between the BOA and OKS, BOA and post-surgery satisfaction, and BOA and the patient's perception of the knee feeling like a normal knee. Pearson's correlation was utilised for normally distributed data, and its non-parametric equivalent (Spearman's correlation) was used for ordinal (i.e., OKS) or non-normally distributed data.

The strength of the correlation, where significant, was interpreted as suggested by (Dancey and Reidy 2011), where r is the correlation coefficient (indicating both r_s, Pearson's correlation coefficient, and ρ: Spearman's correlation coefficient):

- $|r| < 0.3$ meant a weak correlation
- $0.4 < |r| < 0.6$ was regarded as a moderate correlation
- $|r| > 0.7$ was interpreted as a strong correlation.

7.3 RESULTS

7.3.1 Participants characteristics

Table 40 shows the participants' baseline characteristics. Sex distribution, age, and height were comparable between patients and NPs. However, patients had significantly higher weight, BMI, and pain levels (VAS Pain) than NPs. Moreover, patients' pain and function measured via the OKS were significantly lower (= worse) than NPs. In the NP group, 42.1% of participants (n = 8) had their right knee assessed. Most patients received a TKR on the left side (n = 10). 66.6% (n = 12) of patients had a cruciate-retaining fixed-bearing knee implant, 16.7% (n = 3) of patients received a cruciate-retaining, mobile-bearing prosthesis, and the remaining 16.7% (n = 3) had a posterior-sacrificing knee implant.

Table 40 Participants' characteristics at baseline

Participants characteristics	Non-pathological volunteers (n = 19)	Patients pre-TKR (n = 18)	Significance level
Females, n (% within group)	9 (50)	9 (50)	0.877
Age (years), mean (SD), *range*	70.1 (6.3) *60 - 80*	67.3 (8.8) *51 - 82*	0.275
Height (m), mean (SD)	1.71 (0.10)	1.71 (0.89)	0.867
Weight (kg), mean (SD)	73.3 (16.2)	97.2 (17.5)	< 0.001*
BMI (kg/m^2), mean (SD)	25.1 (5.0)	33.1 (5.6)	< 0.001*
OKS, median (IQR)	47 (2)	25 (5)	< 0.001‡*
VAS Pain (mm), median (IQR)	0 (3)	61 (37)	< 0.001‡*

TKR: Total Knee Replacement; n: number; IQR: interquartile range; SD: standard deviation; m: meters; kg: kilograms; BMI: Body Mass Index; OKS: Oxford Knee Score; VAS: Visual Analogue Scale; ‡: non-parametric test;
*: statistically significant at $p < 0.05$.

7.3.2 Application of the Cardiff classifier to *in vivo* knee kinematics during a step-up

The variables entered in the Cardiff classifier allowed to discriminate patients pre-surgery from NPs, with an 83.8% out-of-sample accuracy (Figure 45). The level of uncertainty in the overall classification was relatively high (Uncertainty > 0.30), especially for NPs (Table 42).

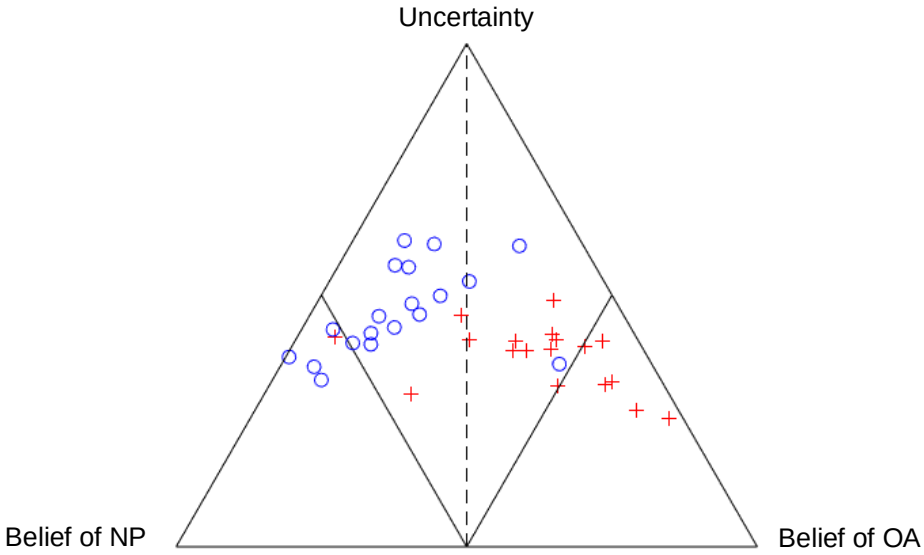


Figure 45 Out-of-sample classification accuracy simplex plot; the blue circles represent the non-pathological volunteers, the red crosses the patients pre-surgery. The black, dashed line indicates the boundary between the non-pathological and osteoarthritic area of the plot. The solid lines within the plot, indicate a Belief of osteoarthritis (BOA) and Belief o*f* non-pathological (BNP) of 0.50 (the boundary between dominant and non-dominant BOA or BNP, respectively).

In the classifier classification based on the whole dataset, the sensitivity was 89%, and the specificity was 84%.

When considering the knee kinematics features best discriminating patients pre-TKR from NPs, only compression-distraction PC1 had an accuracy of over 80%. Only two knee kinematics features (internal-external rotation PC1 and flexion-extension PC2), and one arthrokinematics feature (anterior-posterior shift PC2) had an accuracy of 70% or over (Table 41). Some of the knee kinematics features were of difficult interpretation, due to the very low squared factor loadings, and the small differences in the single component reconstructions and the waveforms of the participants with low and high principal component scores. Figure 46 shows the ensemble averages of the knee kinematics for patients pre- and post-TKR and NPs.

Table 41 Interpretation and accuracy of the knee kinematics features in discriminating patients pre-TKR from NPs during a step-up activity

Principal components	Interpretation of low PC score	Variance represented (%)	Rank based on accuracy	Accuracy (%)
Compression-distraction PC1	Patients had a larger compression throughout the task	87	1	84
Internal-external rotation PC1	NPs had a more externally rotated femur relative to the tibia throughout the task	87	2	78
Flexion-extension angle PC2	NPs had a larger knee extension at the end of the task (from about 70%)	15	3	76
Anterior-posterior shift PC2	Patients had a larger anterior translation in the first half of the task	17	4	70
Anterior-posterior shift PC3	NPs had less anterior translation in the first 20% of the task	5	5	65
Adduction-abduction PC1	NPs had less knee varus throughout the task	94	6	62
Compression-distraction PC2	Patients had a reduced range of compression	8	7	62
Adduction-abduction PC2	Difficult to interpret	2	8	59
Adduction-abduction PC3	Difficult to interpret	2	9	59
Flexion-extension angle PC3	NPs had a larger knee flexion at the start of the task and around 50% of the task	5	10	57
Flexion-extension angle PC1	Patients had a less flexed knee throughout the task	75	11	57
Internal-external rotation PC3	Difficult to interpret	2	12	57
Internal-external rotation PC2	NPs had a larger internal rotation of the femur relative to the tibia at the end of the task	8	13	51
Compression-distraction PC3	Very few differences between groups, difficult to interpret	3	14	51
Anterior-posterior shift PC1	NPs had less anterior translation throughout the task	74	15	49

PC: principal component; NPs: non-pathological subjects

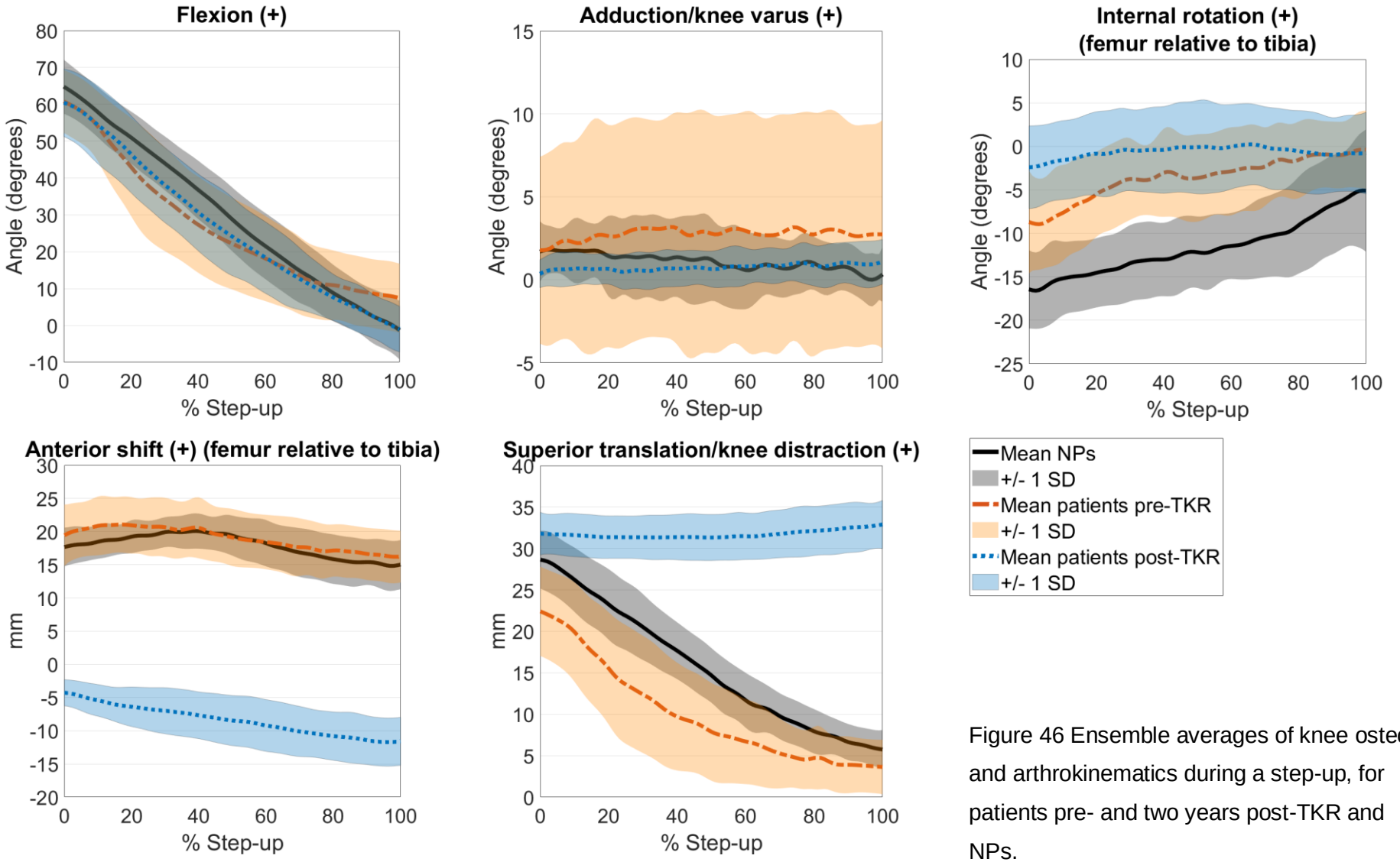


Figure 46 Ensemble averages of knee osteo and arthrokinematics during a step-up, for patients pre- and two years post-TKR and NPs.

Table 42 Comparison of the knee function and OKS between patients pre and post-TKR and comparison of the Belief values post-TKR against NPs

	Pre-TKR (n = 18) Mean (SD)	Post-TKR (n = 18) Mean (SD)	NPs (n = 19) Mean (SD)	Significance level pre to post-TKR	Significance level post-TKR to NPs
Knee objective function					
BOA	0.50 (0.13)	0.36 (0.11)	0.10 (0.08)	0.002‡*	< 0.001‡*
BNP	0.16 (0.11)	0.39 (0.09)	0.41 (0.14)	< 0.001*	0.612
Uncertainty	0.37 (0.07)	0.28 (0.04)	0.46 (0.14)	0.001*	< 0.001‡*
Subjective function					
OKS	25 (5)	43 (8)	-	0.002*	-

NP: non-pathological volunteer; n: number; SD: standard deviation; TKR: Total Knee Replacement; BOA: Belief of Osteoarthritis; BNP: Belief of Non-Pathological; OKS: Oxford Knee Score (0: extreme issues, 48: no issues); ‡: non-parametric test, results reported as median (interquartile range).
*: statistically significant at $p < 0.025$.

Patients attended their follow-up at 24.5 (1.8 IQR) months, on a median. Two patients did not attend the two-year post-TKR appointment, and therefore, the one-year post-TKR knee kinematics was utilised. 88.9% (n = 16) of patients had a decrease in their BOA, indicating an improvement in their knee function (Figure 47).

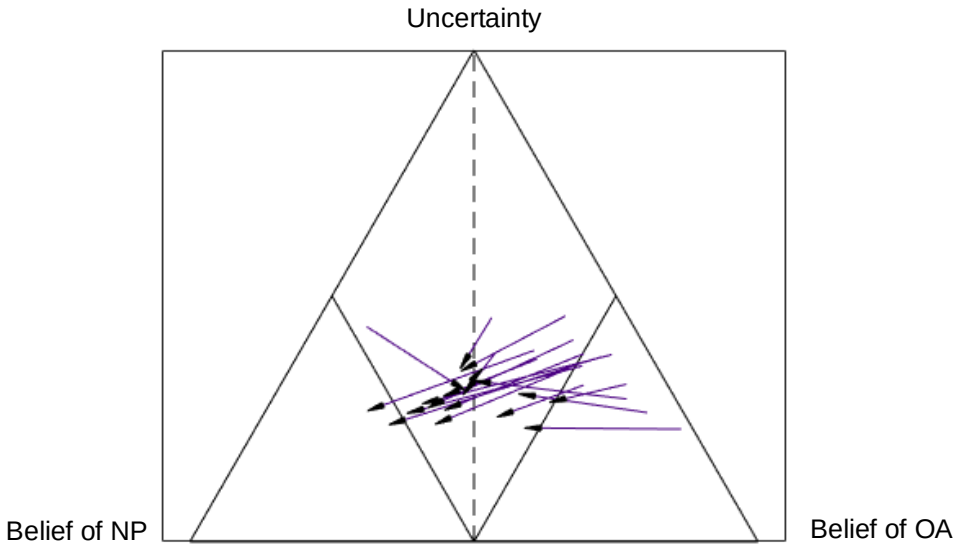


Figure 47 Patients' change in knee function pre- to two years post-TKR (purple arrows). The black, dashed line indicates the boundary between the non-pathological and osteoarthritic area of the plot. The solid lines within the plot, indicate a Belief of osteoarthritis (BOA) and Belief o*f* non-pathological (BNP) of 0.50 (the boundary between dominant and non-dominant BOA or BNP, respectively).

Table 42 shows that patients had a significant increase in the BNP and a significant decrease in the BOA and Uncertainty post-TKR. The patients' OKS score was significantly higher post-TKR, indicating an improvement in self-perceived pain and function. When assessed against NPs, patients post-TKR had a significantly higher BOA, a comparable BNP, and a significantly smaller Uncertainty. NPs had a high level of uncertainty, compared to patients pre- and post-TKR.

7.3.3 Correlation between knee kinematics and PROMs

Pre-surgery, there was no correlation between the patients' perceived knee pain and function (OKS) and their objectively measured knee function (BOA), during a step-up movement, as shown in Table 43. Similarly, post-TKR, no association between the objective knee function and the OKS was found.

Post-TKR, the mean satisfaction was 6.8 mm (14.4) (range 0 – 57) on a VAS scale (0 completely satisfied, 100 extremely dissatisfied), meaning that most patients were highly satisfied with the TKR; only two patients had a VAS satisfaction score higher than 6 and were less satisfied than the remaining patients in the group (31 and 57 mm on the VAS). Therefore, considering the satisfaction scores, only one patient (5.6%) out of eighteen was dissatisfied (VAS pain score: 57, closer to dissatisfaction on the VAS scale).

When asked "How much does your knee feel like a normal knee?", the patients' mean score was 19.2 (SD: 23.7) (range 0 – 88), on a VAS scale from 0 to 100, where 0 represented "completely normal", meaning that most patients felt their knee was not dissimilar to a normal knee. Nevertheless, no correlation was found between BOA and satisfaction, nor the perception of the knee feeling like a normal knee.

Table 43 Correlation coefficients (significance level) for the relationship between objective knee function and PROMs pre- and post-TKR

	BOA pre-TKR (n = 18)	BOA post-TKR (n = 18)
Pre-TKR		
OKS	-0.123‡ (0.627)	
Post-TKR		
OKS		-0.453‡ (0.059)
Surgery satisfaction		0.227 (0.365)
How much does the knee feel like a normal knee?		-0.097 (0.701)

BOA: Belief of Osteoarthritis; TKR: Total Knee Replacement; n: number; OKS: Oxford Knee Score; ‡: Spearman's rho
*: statistically significant at $p < 0.05$

7.4 DISCUSSION

7.4.1 Application of the Cardiff classifier to *in vivo* knee kinematics during a step-up

One of the aims of this study was to explore the application of the Cardiff classifier to the assessment of knee kinematics *in vivo* measured via dynamic single-plane fluoroscopy during a loaded activity, namely, step-up. It was found that by using fifteen features of joint rotations in three planes and translations in the sagittal plane, the classifier was able to discriminate between subjects with severe knee OA and NPs with an 83.8% out-of-sample accuracy.

Only four PCs had relatively high accuracy (≥ 70%) in discriminating patients with severe knee OA from NPs, the most accurate feature (87% accuracy) was compression-distraction PC1, describing patients' larger tibiofemoral compression throughout the task, indicating joint space narrowing. This was not surprising, due to these patients having severe knee OA, at which stage, joint space narrowing is a typical characteristic of the knee joint, often seen in standing radiographs (Guermazi et al. 2008), in *in vivo* knee kinematics during a step-up (Saari et al. 2005a), and linked to cartilage damage and menisci extrusion (Bloecker et al. 2013). The second most accurate feature was internal-external rotation PC1, indicating that although the pattern of movement was similar between patients and NPs, patients had a more internally rotated femur relative to the tibia than NPs throughout the step-up. This finding was in line with a recent systematic review highlighting that the lateral femoral condyle has a reduced posterior translation in several ADLs, including step-up (Scarvell et al. 2018), thus indicating a more internally rotated femur over the tibia in people with severe knee OA. Moreover, previous reports showed that with an increase in knee extension during a step-up, there was a gradual increase in femur internal rotation in both NPs and people with severe OA (Saari et al. 2005a; Yue et al. 2011), indicating the presence of a screw-home mechanism (Freeman and Pinskerova 2005; Varadarajan et al. 2009), as in the present investigation. Flexion extension angle PC2 had 76% accuracy in discriminating between patients and NPs, and it described patients' restricted knee extension at the end of the step-up, suggestive of potential knee flexion deformity. This finding was not unexpected, since people with severe knee OA are often unable to fully extend their knees (Scarvell et al. 2007; Pua et al. 2019). Anterior-posterior shift PC2 had a 70%

accuracy in distinguishing between knee OA and NPs knee kinematics and identifies patients having a larger anterior translation of the femur over the tibia in the first half of step-up (from flexion to extension). An earlier report observed a similar pattern during a leg press (Scarvell et al. 2007), and Scarvell et al. (2018) showed that the lateral condyles were more anteriorly positioned in people with knee OA compared to NPs during knee flexion. However, it must be noted that the anterior translation in the current study was not calculated separately for medial and lateral condyles as Orthovis did not allow to compute the femoral contact points on the tibia, and therefore, it is not clear whether the overall anterior translation was possibly driven by the lateral condyle arthrokinematics. The current results contrasted with what was found by Saari et al. (2005) who looked at the overall anterior-posterior translation of the femur relative to the tibia and observed that people with severe knee OA had a more posteriorly positioned femur throughout the step-up. Nevertheless, the authors (Saari et al. 2005a), included only participants with medial knee OA and utilised a different technique to calculate joint kinematics and this may justify the discrepancies from the present study. A lack of a standard method to define the joint coordinate system when measuring *in vivo* knee kinematics has been an ongoing limiting factor in comparing the results from different reports (Scarvell et al. 2018).

Hypothesis 1.2: patients' BOA improves but is not comparable to the BOA of NPs post-TKR

The Cardiff classifier revealed that while *in vivo* knee kinematics were closer to that of healthy knees two years post-TKR, they were not comparable to that of NPs, in accordance with the hypothesis. The current investigation supports the findings from previous researchers, where the knee limited internal-external rotation and posterior translation pre-TKR were maintained post-TKR (Kitagawa et al. 2010). Similarly, Yue et al. (2011) observed a reduced external rotation of the femur during flexion, and a paradoxical anterior translation of the femur relative to the tibia at the start of knee flexion following cruciate-retaining TKR. Rehman et al. (2022), corroborated the above observations, as cruciate-retaining knees showed an unnatural posterior translation during knee extension while performing a step-up.

The classifier displayed a higher level of uncertainty both pre- and post-TKR, and a lower level of discriminative accuracy, if compared to previous studies

employing this technique in biomechanics data originating from 3D gait analysis (Watling 2013; Metcalfe et al. 2017; Biggs et al. 2019a; Biggs et al. 2019b; Biggs et al. 2021). Due to the small number of variables available for the classification (three rotations and two translations), all the first three PCs for each of the variables were retained to train, test the classifier and then classify subjects as having NP or OA knee kinematics characteristics (a total of 15 PCs or knee kinematics features). Previous research found that not all first three PCs for a variable may be relevant to aid in discriminating patients from NPs (Biggs 2016; Biggs et al. 2019b). In the current investigation, eleven variables out of fifteen (73.3%) had a discriminatory accuracy of < 70%, therefore, some of the present PCs utilised in the classifier, may have not been relevant in aiding in distinguishing patients pre-TKR and NPs. Hence, it may be suggested that the information included in the classifier may not have been enough to reach higher levels of accuracy in classifying patients pre-TKR from NPs. When considering the results of the Chapter 4 and 5, and (Biggs et al. 2019b), at least 90% of the eighteen highest-ranking features discriminating between patients pre-TKR and NPs had an accuracy >70%.

The present results could indicate that more input variables may be beneficial for a more accurate classification of *in vivo* knee function in patients with severe knee, using the Cardiff classifier. When examining the tibiofemoral contact points between medial and lateral knee compartments in healthy knees, it was shown that the lateral femoral condyle translates posteriorly to a larger degree than the medial one when moving into knee flexion (Grieco et al. 2018). Differences in tibiofemoral contact point patterns were found between native and osteoarthritic knees (Saari et al. 2005a; Scarvell et al. 2007; Yue et al. 2011). Orthovis did not allow for the calculation of the femoral contact points on the tibia, and different software may be utilised in the future to obtain these measures. It is recommended that future work should explore whether including knee tibiofemoral contact points may increase the classification accuracy of the classifier. Moreover, medio-lateral translations were not considered due to the higher level of inaccuracy of these measurements, and future research could explore the relevance of including these variables in the classifier. For this purpose, the use of a bi-plane fluoroscopy system, currently utilised at the MSKBRF facility, would be advisable, to obtain data of higher accuracy in the medio-lateral translations.

7.4.2 Correlations between knee function and PROMs

Hypothesis 2.1, 2.2: pre- and post-TKR, there is a significant association between OKS and BOA

Another aim of the study was to explore whether the knee function pre and post-TKR measured *in vivo* using dynamic imaging would be associated with PROMs. In contrast with the hypotheses, no correlation was found pre- or post-TKR between the OKS and the knee function (i.e., BOA). Although not significant, the current results showed an inverse relationship between OKS and BOA, where a higher BOA (more compromised knee function) may have been related to a lower OKS (more severe pain and affected function). Future research should expand on the current findings, including a larger sample of patients.

Null hypothesis 2.3: post-TKR, there is not a significant association between BOA and satisfaction levels

In the current study, 94.4% of patients were satisfied with the surgery, which is a slightly higher rate than what was found in a systematic review with meta-analysis, where satisfaction post-TKR ranged between 71.7 and 92.5% (Khatib et al. 2015). As hypothesised, no correlation was found between knee function and satisfaction. Satisfaction is a complex concept to measure, and it was demonstrated that it does not depend only on the objective function achieved post-TKR, but it can be influenced by several factors, to name a few, expectations and pain in other joints (Scott et al. 2010), socio-economic aspects (Barrack et al. 2014), hospital experience (Hamilton et al. 2013).

Hypothesis 2.4: post-TKR, there is a significant correlation between the BOA and the perception of “the knee feeling like a normal knee”

Surprisingly, and opposed to what was hypothesised, there was no relationship between the patients perceiving that their knee felt like a normal knee, and the knee osteo and arthrokinematics. This means that there must have been other features influencing the patient’s perception, either related to the joint kinematics and that were not included in the classifier (i.e., medio-lateral translations, patellofemoral kinematics), or potentially associated with other aspects of a healthy joint, such as muscles strength, proprioception, the presence of compensating

strategies and pain in the other joints of the lower limb. Chapter 6 showed that multiple joints were affected by compensatory mechanisms pre- and post-TKR, and the overall gait function did not correlate with PROMs. The findings of the current study seem to add to the evidence demonstrating that objective function, indicating what patients can actually do when compared to NPs, may not be the only element contributing to what patients perceive they can do or how they feel about the replaced knee. Moreover, PROMs assess ADLs which are normally bipedal activities, and therefore, issues on the contralateral side of the operated knee, which are often present both pre and post-TKR (Metcalfe 2014; Metcalfe et al. 2017), may also influence the PROM score.

Only a handful of studies explored the relationship between knee-specific PROMs and knee kinematics *in vivo* (Fantozzi et al. 2006; Galvin 2019), while others looked at differences in PROMs between patients with certain knee kinematics patterns (Nishio et al. 2014; Warth et al. 2017), or explored knee kinematics characteristics between patients with different PROMs (Van Onsem et al. 2020). Galvin (2019) examined a large group of fifty-six patients listed for a TKR and found that *in vivo* maximum knee flexion during kneeling was predictive of the OKS, with higher flexion predicting a higher OKS (i.e., less pain and better perceived function). Some of the patients examined in the report (Galvin 2019) were in common with the ones included in the current study, however, the task assessed was different and possibly more demanding than a step-up-down movement, given the higher degrees of knee flexion required in kneeling, and considering patients with severe knee OA, usually have limitations in knee flexion (Brown et al. 2009b; Pua et al. 2013; Standifird et al. 2016).

Fantozzi et al. (2006) analysed the *in vivo* knee kinematics in twenty-three subjects between two and four years post-TKR during step-up and down and found that a more pronounced posterior position of the femur on the tibia at peak knee flexion during the task was moderately associated with a higher (better) Knee Society Score (KSS) “Knee” and “Function” sub-scores, but only in patients receiving a cruciate-retaining implant. The KSS is a combination of patient-reported details and clinical assessment outcomes, where the “Knee” subs-core evaluates pain, flexion contracture, passive knee flexion ROM, knee alignment and stability assessed by the clinician, while the “Function” sub-score addresses the distance the

patient can walk, the use of walking aids, the description of stair negotiation strategies (e.g., use of a handrail). Of note, one of the inclusion criteria was the absence of pain at the time of assessment, which may limit the external validity of the findings since some people, as in the current study, continue experiencing pain post-TKR. When dividing patients by those showing a medial pivot intraoperatively (in an open kinetic chain movement) from those who did not, some researchers found significant differences in KSS scores (Nishio et al. 2014), while others did not (Warth et al. 2017). In addition, Van Onsem et al. (2020) did not observe any *in vivo* anterior-posterior translation differences during a knee flexion-extension (executed via an open kinetic chain movement) between patients with excellent or low KOOS or KSS-satisfaction who received a posterior-stabilised implant. However, the authors (Van Onsem et al. 2020) found that those with better KOOS scores and higher satisfaction had a less anterior translation of both femoral condyles between 0-30° of knee flexion in a closed kinetic chain movement and a larger posterior translation of the lateral condyle between 60-90° of knee flexion. In healthy knees, the medial condyle moves posteriorly by about 2 mm during flexion, less than the lateral condyle which moves posteriorly by roughly 20 mm, indicating a gradual external rotation of the femur relative to the tibia when moving into flexion (Zingde and Slamin 2017). Hence, the authors (Van Onsem et al. 2020) suggested that knee anterior-posterior translations that are similar to those of a native knee would yield better PROMs.

Contrary to the present study, all the above reports focussed on one single aspect of knee kinematics *in vivo*. This could be an advantage as it may highlight the knee kinematics characteristic that seemed to be influential in determining PROMs, and that could be targeted by improving the knee implant designs. Nevertheless, the movement of the knee entails a combination of translations and rotations, and the advantage of the current study was to analyse them together and evaluate how the overall knee kinematics and not only, for instance, anterior-posterior translations, medial pivot or peak knee flexion, may be related to what patients perceive in terms of function.

The results of this study should be considered given its limitations. Although there was an attempt in standardising the step-up task by using a step of the same height for all participants, the position of the foot and hip were not standardised, and

this may have caused variations in the knee kinematics. Moreover, participants started the task with varying amounts of knee flexion. This may have been due to their knee ROM limitations or due to having different heights. Nevertheless, no significant correlation was found between the knee flexion angle at the start of the task and the patient's height, pre- (r_s = -0.438, p = 0.068) or post-TKR (ρ = -0.026, p = 0.919). It is suggested that the current findings were not influenced by participants' different heights.

The sample size was relatively small and the lack of correlation between PROMs and the knee function may have been due to this factor, or due to the lack of more relevant knee kinematics features within the classifier that may have potentially led to a stronger relationship with PROMs.

The task was normalised to % of the step-up, but several studies reported the translations and rotations as a function of knee flexion angle. The choice made in the current study was an attempt to minimise data loss. Some subjects started the knee flexion at 46 degrees pre-TKR and 35 degrees post-TKR. If all data were to be cropped to the smallest knee flexion (and knee extension) available in the dataset, there would have been data loss. Retaining the data in its original form without normalising it would have made it difficult to run PCA in a way that would have been easily interpretable.

This study analysed *in vivo* knee kinematics only during a step-up activity, and it may be that more challenging activities evaluating the knee kinematics at larger degrees of knee flexion, such as deep kneeling, squatting or stair descending could offer more relevant information to classify patients as having an OA or NP knee function.

7.5 CONCLUSION

For the first time, the Cardiff classifier was utilised to evaluate the *in vivo* knee osteo and arthrokinematics of patients pre- and post-TKR during a step-up movement measured using dynamic single-plane fluoroscopy. The promising results show that the classifier was able to discriminate between subjects with severe knee OA and NPs with relatively high accuracy and demonstrated that the knee kinematics of people receiving various knee implants, although improving, was not comparable to those of NPs two years post-TKR. Considering the recent addition of

a bespoke bi-plane fluoroscopy system at the MSKBRF facility, and the plans to collect longitudinal data on people pre- and post-TKR/HTO, the findings of the present study may offer the basis to develop the Cardiff classifier further. The recommendation is that more input variables should be utilised in the classifier, such as tibial contact points and medio-lateral translations, to improve the accuracy of the classification. Future research should explore whether more biomechanically demanding tasks may be more relevant in aiding the classification of *in vivo* knee function.

To the best of the author's knowledge, this was the first investigation exploring the correlation between PROMs and a summative measure of *in vivo* knee osteo and arthrokinematics during a step-up, in patients with severe knee OA and two years post-TKR. The current findings demonstrated that knee function was not correlated with the OKS pre- or post-TKR, or the patient's satisfaction or the perception of the operated knee feeling like a normal knee two years after the surgery. The existing literature suggests that certain knee rotations or translations may be related to clinical and patient-reported measures like the KSS, or PROMs such as the KOOS and OKS. The present study may suggest that the overall *in vivo* knee function alone may not be the primary driver in determining what patients perceive they can do or the limitations they have during everyday life tasks pre and two years post-TKR.

Chapter 8: Conclusions and future work

8.1 SUMMARY CONCLUSIONS

The body of work in this thesis contributed to advancing and refining the application of the Cardiff classifier to subjects with severe knee OA and post-TKR, demonstrated the advantage of using the Classifier over similar measures of gait biomechanics, identified factors predictive of the biomechanics post-TKR, and the relationship between biomechanics, patient-reported outcomes and physical performance pre- and post-TKR.

Conclusions regarding the studies within the thesis were addressed within each of the chapters. Therefore, the present chapter will present a summary of the most prominent findings for each thesis objective, the contributions to knowledge, the general limitations, and recommendations for future research.

Objective 1: *To compare the measurement abilities of the Cardiff classifier to the Gait Deviation Index and the Gait Deviation Index-kinetics pre and one-year post-TKR, and to identify the key gait features discriminating knee OA from NP biomechanics.*

There is limited research describing the responsiveness of the classifier in patients pre- to one-year post-TKR; moreover, no studies compared the classifier to other summative measures of gait function to evaluate its performance. Chapter 4 showed that in two groups of patients scheduled for TKR, from different countries (UK and Sweden), but with similar baseline characteristics, the classifier had a large responsiveness in measuring the change in gait biomechanics one year post-TKR. This finding added to the limited knowledge regarding the classifier's responsiveness to change, which corresponded to one earlier study produced by our research group.

A novel finding introduced by the current work was that, in both the Cardiff and Karolinska patients, the classifier's responsiveness to change was greater than that obtained via the GDI and GDI-kinetic, two summative measures of gait function often used in the OA biomechanics research field. Therefore, it is recommended that when assessing the overall gait function change pre- to post-TKR, the classifier may be a more appropriate measure than the other indexes examined in this study, possibly due to combining both kinetics and kinematics and GRF gait features.

The classifier had never been compared to other measures of gait function before. The novel contribution of the current study was showing that pre-TKR, the classifier outputs correlated with both the GDI and GDI-kinetic in both cohorts, suggesting that kinetics and kinematics may be affected to a similar degree at this stage. Conversely, the two gait indexes and their change pre- to post-TKR showed poor or mixed agreement with the gait function post-TKR or its change, measured via the classifier. It must be noted that while patients were compared to a group of age-matched NPs in the Karolinska cohort, in the Cardiff cohort they were compared to a group of mixed-age NPs, who were much younger than the patients. The fact that similar responsiveness to change of the classifier was obtained in the two cohorts, despite the different characteristics of the reference groups, aided in confirming the measurement properties of the classifier.

It is recommended that GDI, GDI-kinetics and classifier may be used interchangeably to measure the objective gait function pre-TKR, as the outputs should be similar between measures. However, different outcomes should be expected if deciding to use the GDI/GDI-kinetic or the classifier to measure gait change or gait function post-TKR. The advantage of the classifier was its ability to show which gait features were most discriminative between patients pre-TKR and NPs, and this information could be useful in developing targeted rehabilitation interventions. By comparing the outputs of the two classifiers developed for each cohort (Cardiff and Karolinska), it was found that over half of the gait features discriminating severe OA from NPs were similar between cohorts. The results also highlighted that the Karolinska patients had a slightly better gait function than the Cardiff ones both pre- and one-year post-TKR. However, it was not clear if these differences were due to the Cardiff cohort's reference group being significantly younger than the patients, as opposed to the age-matched controls in the Karolinska group. Most previous studies adopting the classifier compared OA patients to an NP group of much younger age, and only two studies compared patients pre-surgery to NPs of similar age. However, the effect of using mixed-age NPs or NPs of similar age to the patients within the classifier had never been explored, and this was addressed in Chapter 5.

Objective 2: *To determine the appropriateness of using a mixed-age reference group to assess the gait biomechanics of people with severe knee OA within the Cardiff classifier, and to identify pre-surgical and surgical factors predictive of the gait biomechanics and gait biomechanics changes one-year post-TKR.*

The preliminary work in Chapter 5 found that classifying patients against a much younger mixed-age reference group via the classifier, produced higher BOA values in patients both pre- and post-TKR, therefore, underestimating patients' objective function, as hypothesised. This was a novel finding, which has implications for future research assessing gait function in people pre- and post-TKR; although when evaluating the change in gait function pre to one-year post-TKR, a mixed-age reference group may be utilised within the classifier, it is suggested that a reference group of similar age to the patients should be employed when assessing the absolute gait function pre- and one-year post-TKR. In this way, patients' function would be compared to the biomechanics they would be expected to have at their present age, had they not had knee OA, rather than being compared to biomechanics expected when they were younger.

The study explored the predictive role of various factors over the change in gait function and the gait function one-year post-TKR. Gait biomechanics pre-TKR, baseline anthropometrics, demographics, comorbidities, surgical factors, issues in other joints, knee ROM change pre to post-TKR, and PROMs were not associated with the change in gait function (i.e., BOA change) post-TKR. It might be suggested that those with a lower baseline BOA score who should theoretically have a greater recovery potential with TKR surgery, do not necessarily achieve a larger gait improvement than those with a higher baseline objective function. It is recommended that a larger sample size should be used to repeat a similar investigation, as there was high variability in how much gait function improved across patients. It must be noted that there may be other factors involved in how much a patient's gait biomechanics improves, and that should be investigated in future research, such as rehabilitation, level of physical activity, duration of symptoms before the surgery, presence of depression and anxiety which may influence the engagement with rehabilitation and physical activity.

The most prominent finding was that BMI, sex, and gait biomechanics pre-TKR explained 56% of the variance of the gait biomechanics post-TKR, indicating

that those with a better gait pre-TKR, also had better gait function one year post-surgery, irrespective of their sex and BMI. A patient sub-group analysis also showed that a greater knee ROM pre-TKR was associated with a better gait function post-TKR. Considering that delaying the TKR decreases the quality-adjusted life years, it is not cost-effective, and given the results of the current study, it is not advisable to wait for the biomechanics of the lower limb to be extremely affected by knee OA to undergo TKR, as this may result in worse gait biomechanics post-TKR. Given the relatively small sample size in the current study, it is recommended that further research on a larger sample of patients (including more females) and NPs is carried out to corroborate the current findings on predicting TKR gait biomechanics outcomes.

Objective 3: *To investigate the merit of adding gait trunk kinematics to aid in classifying NPs and individuals with severe knee OA via the Cardiff classifier, and to explore the interrelationship between the gait biomechanics, performance-based tests, and patient-reported outcomes before and after TKR surgery.*

There is limited research exploring differences in trunk kinematics between OA patients and NPs. For the first time it was shown that when using the classifier, gait trunk kinematics in the frontal, sagittal and transverse planes were not relevant in aiding in the discrimination of gait biomechanics of NPs and subjects with late-stage OA. It is suggested the assessment of gait function via the classifier, should not be affected by the lack of trunk kinematics data.

PBTs measurement properties have been mainly assessed against PROMs, but there is limited evidence comparing PBTs to other objective measures of function. The study showed that there was a correlation, or trends of association, between gait biomechanics and the core PBTs suggested by OARSI (40mfast-paced walk test, stair climb test and 30s chair test), pre-, three and six months post-TKR. PBTs measure the time to accomplish a task or the number of repetitions of sit-to-stand within a timeframe and do not inform on movement quality, but the current novel findings showed that the presence of more compromised gait biomechanics may be associated with a poorer performance during the PBTs both pre-, three and six months post-TKR. These preliminary results offer additional evidence in support of using the abovementioned PBTs in clinical practice, although the results relative to

the post-TKR time points should be treated with caution, due to the small sample size, therefore, further investigation is needed on a larger sample.

By utilising a different group of patients than the ones employed previously in our research group, it was demonstrated that no correlation, nor a trend of association, could be found between PROMs and gait biomechanics pre-or three, six and approximately twelve months post-TKR. This was contradictory to earlier findings within our research group, suggesting that worse gait biomechanics were associated with poorer PROMs pre- and one-year post-TKR (Biggs et al. 2019a; Whatling et al. 2022). The additional lack of correlation between PROMs and PBTs in this study reinforced the existing evidence that objective measures of function and PROMs measure different constructs and should be utilised together in evaluating OA or TKR outcomes.

Objective 4: *To explore the ability of the Cardiff classifier to discriminate between NP and OA subjects using in vivo knee kinematics during a loaded activity (measured via dynamic single-plane fluoroscopy and image registration), and to investigate the association between knee kinematics and patient-reported outcomes pre-and post-TKR.*

This work aided in expanding the application of the Cardiff classifier beyond the assessment of gait biomechanics. Employing the classifier to assess *in vivo* knee kinematics during a step-up motion showed, for the first time, a good level of accuracy in discriminating between severe knee OA and NP knee function. There was a higher level of uncertainty in the classification when compared to the usual uncertainty level observed when utilising the classifier on gait biomechanics. It is suggested that more input variables may be needed to improve the accuracy of the classifier when using dynamic imaging of the kinematics of the knee, such as tibial contact points, and knee mediolateral translations. Another recommendation is that more demanding tasks should be explored via fluoroscopy and, consequently, the classifier.

In line with previous literature, it was demonstrated that *in vivo* knee kinematics, while improving two years post-TKR, were not comparable to those of a reference group. The novelty of the current work consisted of summarising and interpreting the key knee kinematics features via PCA and assessing the knee

kinematics in its entirety, rather than by single variables and discrete kinematics, offering a more comprehensive evaluation of the knee function.

Further novel findings were that no correlation was found between knee kinematics and OKS pre-or post-TKR, or satisfaction, nor the patient's perception of their knee "feeling like a normal knee" two years post-TKR. These results add to what was observed in Chapter 6, and previous research, suggesting that objective function, in this case, the knee kinematics, may measure different constructs of physical function, compared to PROMs. The novelty of this piece of work lies in comparing a summative measure of knee function to PROMs, rather than performance-based tests or discrete kinematics to PROMs, which is what can be found in the existing literature.

8.2 THESIS LIMITATIONS

Several limitations were identified and discussed within the relevant chapters, but there were overall limitations to the current PhD research, discussed followingly.

Data from different sources

Chapters 4 and 6 include retrospective analyses of data collected outside our research facilities and by different research groups. This means that there was little control over how data were collected, it was unknown what Standard Operating Procedures were in place for data collection and processing, and there may have been inter-operator differences if the data was collected by several individuals. These factors may have affected the data quality. Nevertheless, research protocols were in place, to ensure data collection was carried out consistently, the researchers involved in data sharing had checked the data quality before sharing it with the author, and this would have limited the errors associated with the data.

Inter-operator differences in data collection and processing

As mentioned in Chapter 3, data for the TKR project have been collected across several years and by different researchers. This may have introduced inter-operator variability in the way data was collected and processed. Previous research (Della Croce et al. 2005) found that inter-operator errors in localising anatomical landmarks range between 11.5 and 24.8 mm, depending on the anatomical

landmark, with larger errors on the posterior superior iliac spines, and smaller errors on the metatarsal heads of the foot. To minimise errors, researchers involved in the TKR study were trained by experienced researchers on locating anatomical landmarks and were formally authorised to collect data once the training had been completed successfully. Furthermore, Visual 3D pipelines with set processing steps were developed to minimise the operator's input and errors (i.e., missing essential data processing steps). Additionally, the current author inspected the data that had been processed before the start of the current PhD to assess data quality, errors were rectified and in the presence of poor data quality, subjects were excluded from further analysis.

Lack of age-matched NPs

In Chapters 4 and 6, NPs were much younger than patients pre-TKR in the Cardiff cohort. Ideally, the reference group should have similar characteristics to the patients, to limit the influence of demographic and anthropometric factors on the results. There have been ongoing challenges in recruiting NP volunteers who matched the TKR study inclusion criteria throughout the duration of the research. Often, people over the age of sixty who were interested in participating had knee or lower limb pain or had had a knee injury, which constituted exclusion criteria. During the current PhD project, one of the main aims was to recruit NPs of similar age and sex to the OA patients, but the challenges of the COVID-19 pandemic limited the ability to achieve this goal.

The Cardiff NP group data available at the time of data analysis for Chapters 4 and 6, included a very limited number of NPs with of similar age and sex to the patients. Therefore, to avoid entering a very small reference group within the Cardiff classifier, and potentially obtaining a poor model, the choice of including mixed-age NPs was made. This approach was justified by previous work on the classifier carried out within our research group, which produced a classification with an accuracy of > 90% in discriminating OA patients from NPs, when using a mixed-age NP group, much younger than patients, on average. Discrepancies between the Cardiff and Karolinska classifiers were found in Chapter 4, such as slightly different highest-ranking gait features between the classifiers, a lower BOA pre- and post-TKR in the Karolinska patients compared to the Cardiff ones, which may have been

due to the Cardiff patients being compared to a mixed-age NP group. While this was a limitation, it highlighted potential risks of errors deriving from comparing patients' biomechanics to that of a much younger cohort, which had never been observed or addressed before. Therefore, to understand the implications of using a mixed-age NP group, as opposed to an NP group of similar age to the patients, further research was carried out in Chapter 5, which produced implications for the work within the chapter itself, and for future research. For Chapter 5, it was possible to merge some of the pre-existing NP group data (collected before the current PhD research) with some of the NP data collected for the current PhD project, to produce an NP group of similar age to the OA patients, and of sufficient size to be utilised within the classifier. However, this reference group could not be utilised within Chapter 6, due to the aims of this chapter, which required to have the availability of trunk kinematics data, collected only since the start of the current PhD research.

8.3 KEY CONCLUSIONS

The current research included different cohorts across the studies and produced several novel contributions to knowledge. The most prominent contributions and recommendations are summarized below and were grouped under "advancing the classifier method", and "clinical".

Advancing the classifier method

- The present novel research showed that the classifier is more responsive in measuring gait function change than the GDI and GDI-kinetic. GDI/GDI-kinetic and classifier outputs correlate pre-TKR and might be utilised interchangeably to assess gait function, while it is not advised to do so to evaluate gait function or its change one-year post-TKR as these measures do not always correlate in different patient groups
- Patients appeared to have a worse gait function (i.e., higher BOA), both pre- and post-TKR when compared to a much younger NP group. It is suggested that a reference group of similar age to the patients should be employed within the classifier when assessing the absolute gait function pre- and one-year post-TKR. However, when evaluating the change in gait function pre- to one-year post-TKR, a mixed-age reference group may be utilised within the

classifier, as this would produce comparable results to utilising an NP group of similar age to the patients

- Trunk kinematics did not discriminate well between patients with severe knee OA and NPs within the classifier, other gait characteristics were shown to have higher discriminative accuracy. It is suggested that not utilising trunk kinematics in assessing function pre- and post-TKR via the classifier should not impact the evaluation of patients' gait function
- The application of the classifier to assess knee in vivo kinematics during step-up was explored pre- and post-TKR. The accuracy of the classification was good, but the high level of classification uncertainty suggested that more input variables, such as tibia contact point and mediolateral translations, should be employed in future similar studies

Clinical contributions

- Gait function was assessed pre- and one-year post-TKR, and it was found that having a worse function pre-TKR was predictive of a poorer function post-TKR (even when controlling for sex and BMI). These results suggest that to have a better gait function post-TKR, it is not advised to wait for the gait biomechanics to deteriorate severely before undergoing a TKR. Further research on a larger sample should be conducted to confirm the findings
- There was a correlation or trends of association between the gait quality and the core PBTs (whose use was suggested by OARSI) before and after TKR. Although PBTs do not inform on the movement quality during the task, impaired biomechanics and the presence of several compensations may be reflected in poorer PBTs scores. These results may support the use of the abovementioned PBTs in clinical practice, but further research on a larger sample is needed
- The current work showed that PROMs were poorly correlated with PBTs, gait function, or knee function during a step-up. This adds to the current literature and reinforces the recommendations from earlier investigations suggesting that subjective and objective measures of function should be combined to obtain a complete understanding of patient outcomes following TKR

8.4 RECOMMENDATIONS FOR FUTURE WORK

Reliability, Minimal Detectable Change (MDC), and Minimal Important Change (MIC) of the Cardiff classifier

For any measurement tool utilised in clinical research, the outcome must be consistent, therefore having sufficient reliability. To date, no research has been carried out on the Cardiff classifier reliability. Future investigations should collect repeated gait data through two sessions a few days apart, on non-pathological volunteers and patients pre- and post-TKR. It is recognised that it may be challenging to obtain consistent gait data on patients pre-TKR due to the variable nature of knee OA symptoms. This dataset would allow calculation of MDC, defining the minimal change score above which there is a “true” change.

In the current thesis, the internal responsiveness to change of the Cardiff classifier was assessed via a distribution-based method (i.e., effect size), due to the absence of an external anchor measure that would have determined the patient’s perception of their function being improved or worsened pre- to post-TKR. Future research should collect the patient’s view on their perceived functional improvement with an anchor question such as (but not limited to) “How much do you think your overall movement quality has changed from pre- to post-surgery?” and a Likert set of answers ranging from “it got significantly worse”, including a “no change” option, up to “it got significantly better”. It would be important to find an anchor question that reflects the description of “gait function”. This would allow to determine the MIC for the Cardiff classifier score and to test the classifier’s internal responsiveness with an anchor-based approach, which would make the interpretation of the classifier measurement properties clinically relevant.

Sensitivity and specificity of the Cardiff classifier

The Cardiff classifier has been mainly employed in groups of individuals affected by knee medial compartment OA In earlier studies), or end-stage knee OA (as in the current thesis) showing that it is an accurate tool for discriminating patients from NPs. Nevertheless, its sensitivity and specificity in distinguishing patients at different stages of the OA disease have never been tested. This information would contribute to investigating and possibly provide data for refining the measuring properties of the classifier. The collaboration with Dr Naili and Professor Holsgaard-

Larsen, started conversations on future research utilising a cohort of patients with mild radiographic knee OA (KL grade 1-2) whose data were collected with the same methods utilised by Dr Naili who has availability of a cohort of people with end-stage OA (KL grade 3-4) and NPs (utilised in this thesis). The availability of data collected with the same methods would make it possible to merge the patient groups within the classifier. A previous study on the abovementioned cohort (Naili et al. 2019a) showed that the GDI and GDI-kinetic did not have sufficient ability to discriminate between patients at different stages of the disease. A Receiver Operator Characteristics Curve analysis could be performed to determine the classifier's sensitivity and specificity in classifying patients at different stages of the disease. It would be expected that patients with more severe knee OA would have a higher BOA than subjects with mild knee OA.

Assessing gait function change throughout the recovery post-TKR

The Cardiff classifier has been utilised to assess the change in gait function pre- to six months post-TKR, and more often pre- to about twelve months post-surgery. Limited literature described the gait function evolution from pre- to three, six, twelve months post-TKR using the classifier (Jones et al. 2008; Whatling et al. 2022), and included a limited sample size (i.e., n = 20 or less), it was demonstrated that the classifier was responsive in measuring gait function changes across time points. In the current thesis, it was not possible to assess the longitudinal change in gait function between different time points in Chapter 6, due to different patients being available to attend each follow-up. Within our research group, there is a small dataset of patients (n = 26) who attended all four appointments (pre-, three, six, and twelve months post-TKR). The gait function could be assessed via the classifier at each time point, and the output would provide information on how the gait function evolves post-TKR. This evidence would be useful for clinicians during the discussion with patients regarding what they should expect post-surgery in terms of objective function change.

External validation of the prediction model and using wearable technologies

Chapter 5 developed a prediction model for the overall gait function one-year post-TKR. The sample size was relatively small and future studies including a larger

cohort should be carried out. Moreover, the external validity of the current model should be tested on a different group of patients for validation purposes.

It remains challenging to define what would be the optimal timing to perform TKR to maximise the biomechanical outcomes, and pain while minimising the revision rates. This would require collecting longitudinal data on biomechanics, PROMs, and revisions, on thousands of patients, which is not feasible when motion capture technology is involved. The research in digital technology is becoming more popular in recent years and as Emery et al. (2019) suggested, the employment of wearable technologies may help evaluate gait biomechanics before and after surgery. Future research could investigate lower limb biomechanics with portable devices that could be used in the clinic, such as inertial measurement units, which could inform on joint angles, and portable force plates or pressure walkways, which collect GRF data. The data collected with these technologies could form the basis for the assessment of gait function via the Cardiff classifier.

Utilising the Cardiff classifier on *in vivo* knee kinematics on a larger cohort of individuals

The work in Chapter 7 of the current thesis was exploratory and only included a relatively small sample of participants despite the availability of a much larger sample within the PICKLeS cohort. Future research could utilise the full dataset within the Cardiff classifier, including tibiofemoral contact points during a step-up, and potentially explore the classification of knee function during kneeling, for which data is available at the University of Canberra.

References

Acker, S. et al. 2011. Accuracy of single-plane fluoroscopy in determining relative position and orientation of total knee replacement components. *Journal of Biomechanics* 44(4), pp. 784–787. doi: 10.1016/J.JBIOMECH.2010.10.033.

Akter, M., Lambert, A.J., Pickering, M.R., Scarvell, J.M. and Smith, P.N. 2014. 3D CT to 2D low dose single-plane fluoroscopy registration algorithm for in-vivo knee motion analysis. *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014*, pp. 5121–5124. doi: 10.1109/EMBC.2014.6944777.

Alghadir, A., Anwer, S. and Brismée, J.M. 2015. The reliability and minimal detectable change of Timed Up and Go test in individuals with grade 1 - 3 knee osteoarthritis. *BMC Musculoskeletal Disorders* 16(1), pp. 1–7. Available at: https://link.springer.com/articles/10.1186/s12891-015-0637-8 [Accessed: 14 May 2023].

Algina, J. and Olejnik, S. 2010. Sample Size Tables for Correlation Analysis with Applications in Partial Correlation and Multiple Regression Analysis. *Multivariate Behavioural Research* 38(3), pp. 309–323. Available at: https://www.tandfonline.com/doi/abs/10.1207/S15327906MBR3803_02 [Accessed: 20 April 2023].

Armand, S., Sangeux, M. and Baker, R. 2014. Optimal markers' placement on the thorax for clinical gait analysis. *Gait & Posture* 39(1), pp. 147–153. Available at: https://www.sciencedirect.com/science/article/pii/S0966636213002993#bib0070 [Accessed: 11 September 2019].

Arner, J.W. et al. 2016. The Effects of Anterior Cruciate Ligament Deficiency on the Meniscus and Articular Cartilage. *Orthopaedic Journal of Sports Medicine* 4(4), p. 232596711663989. doi: 10.1177/2325967116639895.

Arthritis Research UK. 2019. *Tower Hamlets: Prevalence of osteoarthritis in England and local authorities*. Available at: https://www.arthritisresearchuk.org/internal-resources/data-and-statistics/bulletins/osteoarthritis/london/tower-hamlets.aspx [Accessed: 26 February 2019].

ASA. 2020. *ASA Physical Status Classification System 2014*. Available at: https://www.asahq.org/standards-and-guidelines/asa-physical-status-classification-system [Accessed: 6 January 2023].

Astephen, J.L., Deluzio, K.J., Caldwell, G.E. and Dunbar, M.J. 2008a. Biomechanical changes at the hip, knee, and ankle joints during gait are associated with knee osteoarthritis severity. *Journal of Orthopaedic Research* 26(3), pp. 332–341. Available at: http://doi.wiley.com/10.1002/jor.20496 [Accessed: 29 April 2021].

Astephen, J.L., Deluzio, K.J., Caldwell, G.E., Dunbar, M.J. and Hubley-Kozey, C.L. 2008b. Gait and neuromuscular pattern changes are associated with differences in knee osteoarthritis severity levels. *Journal of Biomechanics* 41(4), pp. 868–876. doi: 10.1016/J.JBIOMECH.2007.10.016.

Astephen Wilson, J.L. 2012. Challenges in dealing with walking speed in knee osteoarthritis gait analyses. *Clinical Biomechanics* 27(3), pp. 210–212. doi: 10.1016/J.CLINBIOMECH.2011.09.009.

Astephen Wilson, J.L., Dunbar, M.J. and Hubley-Kozey, C.L. 2015. Knee joint biomechanics and neuromuscular control during gait before and after total knee arthroplasty are sex-specific. *J Arthroplasty* 30(1), pp. 118–125. doi: 10.1016/j.arth.2014.07.028.

Bade, M.J. et al. 2020. Movement pattern biofeedback training after total knee arthroplasty: Randomized clinical trial protocol. *Contemporary Clinical Trials* 91, p. 105973. doi: 10.1016/J.CCT.2020.105973.

Bade, M.J., Kittelson, J.M., Kohrt, W.M. and Stevens-Lapsley, J.E. 2014. Predicting Functional Performance and Range of Motion Outcomes After Total Knee Arthroplasty. *Am J Phys Med Rehabil* 93(7), pp. 579–585. Available at: /pmc/articles/PMC4350663/ [Accessed: 9 April 2022].

Baert, I.A.C., Jonkers, I., Staes, F., Luyten, F.P., Truijen, S. and Verschueren, S.M.P. 2013. Gait characteristics and lower limb muscle strength in women with early and established knee osteoarthritis. *Clinical Biomechanics* 28(1), pp. 40–47. doi: 10.1016/J.CLINBIOMECH.2012.10.007.

Baker, P.N., Deehan, D.J., Lees, D., Jameson, S., Avery, P.J., Gregg, P.J. and Reed, M.R. 2012. The effect of surgical factors on early patient-reported outcome

measures (PROMS) following total knee replacement. *Journal of Bone and Joint Surgery - Series B* 94 B(8), pp. 1058–1066. doi: 10.1302/0301-620X.94B8.28786/LETTERTOEDITOR.

Baker, P.N., Van Der Meulen, J.H., Lewsey, J., Gregg, P.J., Orthopaedic, C. and Fellow, R. 2007. The role of pain and function in determining patient satisfaction after total knee replacement DATA FROM THE NATIONAL JOINT REGISTRY FOR ENGLAND AND WALES. 89(7). Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1020.5412&rep=rep1&type=pdf [Accessed: 12 May 2019].

Baker, R. 2001. Pelvic angles: a mathematically rigorous definition which is consistent with a conventional clinical understanding of the terms. *Gait & Posture* 13(1), pp. 1–6. doi: https://doi.org/10.1016/S0966-6362(00)00083-7.

Bannuru, R.R., Dasi, U.R. and McAlindon, T.E. 2010. Reassessing the role of acetaminophen in osteoarthritis: systematic review and meta-analysis. *Osteoarthritis and Cartilage* 18, p. S250. doi: 10.1016/S1063-4584(10)60585-7.

Barrack, R.L. et al. 2014. Impact of Socioeconomic Factors on Outcome of Total Knee Arthroplasty. *Clinical Orthopaedics and Related Research* 472(1), pp. 86–97. Available at: https://link.springer.com/content/pdf/10.1007%2Fs11999-013-3002-y.pdf [Accessed: 12 May 2019].

Beach, A., Regazzola, G., Neri, T., Verheul, R. and Parker, D. 2019. The effect of knee prosthesis design on tibiofemoral biomechanics during extension tasks following total knee arthroplasty. *The Knee* 26(5), pp. 1010–1019. doi: 10.1016/J.KNEE.2019.07.008.

Bechard, D.J., Birmingham, T.B., Zecevic, A.A., Jones, I.C., Giffin, J.R. and Jenkyn, T.R. 2012. Toe-out, lateral trunk lean, and pelvic obliquity during prolonged walking in patients with medial compartment knee osteoarthritis and healthy controls. *Arthritis Care & Research* 64(4), pp. 525–532. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/acr.21584 [Accessed: 4 June 2023].

Bell, A.L., Pedersen, D.R. and Brand, R.A. 1990. A comparison of the accuracy of several hip center location prediction methods. *Journal of Biomechanics* 23(6), pp. 617–621. Available at:

https://www.sciencedirect.com/science/article/pii/0021929090900547 [Accessed: 9 March 2023].

Bellamy, N. et al. 1997. Recommendations for a Core Set of Outcome Measures for Future Phase III Clinical Trials in Knee, Hip, and Hand Osteoarthritis. Consensus Development at OMERACT III. *The Journal of Rheumatology* 24(4), pp. 799–802.

Bennell, K., Dobson, F. and Hinman, R. 2011. Measures of physical performance assessments: Self-Paced Walk Test (SPWT), Stair Climb Test (SCT), Six-Minute Walk Test (6MWT), Chair Stand Test (CST), Timed Up & Go (TUG), Sock Test, Lift and Carry Test (LCT), and Car Task. *Arthritis Care and Research* 63(SUPPL. 11). doi: 10.1002/ACR.20538.

Berliner, J.L., Brodke, D.J., Chan, V., SooHoo, N.F. and Bozic, K.J. 2017. Can Preoperative Patient-reported Outcome Measures Be Used to Predict Meaningful Improvement in Function After TKA? *Clinical Orthopaedics and Related Research* 475(1), pp. 149–157. Available at: https://link.springer.com/article/10.1007/s11999-016-4770-y [Accessed: 1 May 2023].

Bifulco, P., Cesarelli, M., Cerciello, T. and Romano, M. 2012. A continuous description of intervertebral motion by means of spline interpolation of kinematic data extracted by videofluoroscopy. *Journal of Biomechanics* 45(4), pp. 634–641. doi: 10.1016/J.JBIOMECH.2011.12.022.

Biggs, P., Holsgaard-Larsen, A., Holt, C.A. and Naili, J.E. 2021. Gait function improvements, using Cardiff Classifier, are related to patient-reported function and pain following hip arthroplasty. *J Orthop Res* 40(5), pp. 1182–1193. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/jor.25149 [Accessed: 14 October 2021].

Biggs, P.R. 2016. *Classification of Biomechanical Changes in Gait Following Total Knee Replacement: An Objective, Multi-feature Analysis*. Cardiff University.

Biggs, P.R., Whatling, G.M., Wilson, C. and Holt, C.A. 2019a. Correlations between patient-perceived outcome and objectively-measured biomechanical change following Total Knee Replacement. *Gait and Posture* 70, pp. 65–70. doi: 10.1016/j.gaitpost.2019.02.028.

Biggs, P.R., Whatling, G.M., Wilson, C., Metcalfe, A.J. and Holt, C.A. 2019b. Which osteoarthritic gait features recover following total knee replacement surgery? *PLoS ONE* 14(1). Available at: /pmc/articles/PMC6347391/?report=abstract [Accessed: 14 January 2021].

Bigoni, M. et al. 2013. Acute and late changes in intraarticular cytokine levels following anterior cruciate ligament injury. *Journal of Orthopaedic Research* 31(2), pp. 315–321. doi: 10.1002/jor.22208.

Birmingham, T.B. et al. 2017. Changes in biomechanical risk factors for knee osteoarthritis and their association with 5-year clinically important improvement after limb realignment surgery. *Osteoarthritis and Cartilage* 25(12), pp. 1999–2006. doi: 10.1016/J.JOCA.2017.08.017.

Blagojevic, M., Jinks, C., Jeffery, A. and Jordan, K.P. 2010. Risk factors for onset of osteoarthritis of the knee in older adults: a systematic review and meta-analysis. *Osteoarthritis and Cartilage* 18(1), pp. 24–33. Available at: http://www.sciencedirect.com/science/article/pii/S1063458409002258.

Bloecker, K. et al. 2013. Tibial coverage, meniscus position, size and damage in knees discordant for joint space narrowing – data from the Osteoarthritis Initiative. *Osteoarthritis and Cartilage* 21(3), pp. 419–427. doi: 10.1016/J.JOCA.2012.11.015.

Bolink, S.A.A.N., Grimm, B. and Heyligers, I.C. 2015a. Patient-reported outcome measures versus inertial performance-based outcome measures: A prospective study in patients undergoing primary total knee arthroplasty. *The Knee* 22(6), pp. 618–623. doi: 10.1016/J.KNEE.2015.04.002.

Bolink, S.A.A.N., Grimm, B. and Heyligers, I.C. 2015b. Patient-reported outcome measures versus inertial performance-based outcome measures: A prospective study in patients undergoing primary total knee arthroplasty. *The Knee* 22(6), pp. 618–623. doi: 10.1016/J.KNEE.2015.04.002.

Bonnefoy-Mazure, A., Armand, S., Sagawa, Y., Suvà, D., Miozzari, H. and Turcot, K. 2017. Knee Kinematic and Clinical Outcomes Evolution Before, 3 Months, and 1 Year After Total Knee Arthroplasty. *The Journal of arthroplasty* 32(3), pp. 793–800. Available at: http://www.ncbi.nlm.nih.gov/pubmed/28007371 [Accessed: 13 May 2019].

Boonstra, M.C., De Waal Malefijt, M.C. and Verdonschot, N. 2008. How to quantify knee function after total knee arthroplasty? *Knee* 15(5), pp. 390–395. doi: 10.1016/j.knee.2008.05.006.

Bowd, J. 2022. *DOES GAIT RETRAINING HAVE THE POTENTIAL TO SLOW OA DEVELOPMENT AND PROLONG THE BENEFITS OF KNEE REALIGNMENT SURGERY?* Cardiff: Cardiff University.

Boyan, B.D. et al. 2013. Hormonal modulation of connective tissue homeostasis and sex differences in risk for osteoarthritis of the knee. *Biology of Sex Differences* 4(1), pp. 1–10. Available at: https://link.springer.com/articles/10.1186/2042-6410-4-3 [Accessed: 15 May 2023].

Boyer, K.A., Johnson, R.T., Banks, J.J., Jewell, C. and Hafer, J.F. 2017a. Systematic review and meta-analysis of gait mechanics in young and older adults. *Experimental Gerontology* 95, pp. 63–70. doi: 10.1016/J.EXGER.2017.05.005.

Boyer, K.A., Johnson, R.T., Banks, J.J., Jewell, C. and Hafer, J.F. 2017b. Systematic review and meta-analysis of gait mechanics in young and older adults. *Experimental Gerontology* 95, pp. 63–70. doi: 10.1016/J.EXGER.2017.05.005.

Boyer, K.A., Johnson, R.T., Banks, J.J., Jewell, C. and Hafer, J.F. 2017c. Systematic review and meta-analysis of gait mechanics in young and older adults. *Experimental Gerontology* 95, pp. 63–70. doi: 10.1016/J.EXGER.2017.05.005.

Brandon, S.C.E., Graham, R.B., Almosnino, S., Sadler, E.M., Stevenson, J.M. and Deluzio, K.J. 2013. Interpreting principal components in biomechanics: Representative extremes and single component reconstruction. *Journal of Electromyography and Kinesiology* 23(6), pp. 1304–1310. doi: 10.1016/J.JELEKIN.2013.09.010.

Braun, H.J. and Gold, G.E. 2012. Diagnosis of osteoarthritis: Imaging. *Bone* 51(2), pp. 278–288. doi: 10.1016/J.BONE.2011.11.019.

Brenneman, E.C. and Maly, M.R. 2018. Identifying changes in gait waveforms following a strengthening intervention for women with knee osteoarthritis using principal components analysis. *Gait & Posture* 59, pp. 286–291. doi: 10.1016/J.GAITPOST.2017.07.006.

Brittain, R. et al. 2022. *National Joint Registry - 19th Annual Report 2022*. Available at: www.njrcentre.org.uk [Accessed: 1 May 2023].

Brouwer, R.W., Raaij, T.M. van, Jakma, T.T.S.C., Verhagen, A.P., Verhaar, J.A. and Bierma-Zeinstra, S.M. 2005. Braces and orthoses for treating osteoarthritis of the knee. *Cochrane Database of Systematic Reviews* (1). doi: 10.1002/14651858.CD004020.PUB2.

Brown, K., Kachelman, J., Topp, R., Quesada, P.M., Nyland, J., Malkani, A. and Swank, A.M. 2009a. Predictors of functional task performance among patients scheduled for total knee arthroplasty. *Journal of Strength and Conditioning Research* 23(2), pp. 436–443. Available at: https://journals.lww.com/nsca-jscr/Fulltext/2009/03000/Predictors_of_Functional_Task_Performance_Among.13.aspx [Accessed: 20 December 2022].

Brown, K., Kachelman, J., Topp, R., Quesada, P.M., Nyland, J., Malkani, A. and Swank, A.M. 2009b. Predictors of functional task performance among patients scheduled for total knee arthroplasty. *Journal of Strength and Conditioning Research* 23(2), pp. 436–443. Available at: https://journals.lww.com/nsca-jscr/Fulltext/2009/03000/Predictors_of_Functional_Task_Performance_Among.13.aspx [Accessed: 1 May 2023].

Burmeister, E. and Aitken, L.M. 2012. Sample size: How many is enough? *Australian Critical Care* 25(4), pp. 271–274. doi: 10.1016/J.AUCC.2012.07.002.

Bytyqi, D., Shabani, B., Lustig, S., Cheze, L., Karahoda Gjurgjeala, N. and Neyret, P. 2014. Gait knee kinematic alterations in medial osteoarthritis: Three dimensional assessment. *International Orthopaedics* 38(6), pp. 1191–1198. Available at: https://link.springer.com/article/10.1007/s00264-014-2312-3 [Accessed: 16 January 2024].

Capin, J.J., Bade, M.J., Jennings, J.M., Snyder-Mackler, L. and Stevens-Lapsley, J.E. 2022. Total Knee Arthroplasty Assessments Should Include Strength and Performance-Based Functional Tests to Complement Range-of-Motion and Patient-Reported Outcome Measures. *Physical Therapy* 102(6), pp. 1–10. Available at: https://academic.oup.com/ptj/article/102/6/pzac033/6556168 [Accessed: 1 May 2023].

Cappozzo, A., Catani, F., Leardini, A., Benedetti, M.G. and Della Croce, U. 1996a. Position and orientation in space of bones during movement: experimental artefacts. *Clinical Biomechanics* 11(2), pp. 90–100. Available at: http://www.sciencedirect.com/science/article/pii/0268003395000461.

Cappozzo, A., Catani, F., Leardini, A., Benedetti, M.G. and Della Croce, U. 1996b. Position and orientation in space of bones during movement: experimental artefacts. *Clinical Biomechanics* 11(2), pp. 90–100. doi: 10.1016/0268-0033(95)00046-1.

Cardiff University. 2020. *Research participants data protection notice*. Available at: https://www.cardiff.ac.uk/public-information/policies-and-procedures/data-protection/research-participants-data-protection-notice.

Chau, T. 2001. A review of analytical techniques for gait data. Part 1: fuzzy, statistical and fractal methods. *Gait & Posture* 13(1), pp. 49–66. doi: 10.1016/S0966-6362(00)00094-1.

Chehab, E.F., Andriacchi, T.P. and Favre, J. 2017. Speed, age, sex, and body mass index provide a rigorous basis for comparing the kinematic and kinetic profiles of the lower extremity during walking. *Journal of Biomechanics* 58, pp. 11–20. doi: 10.1016/J.JBIOMECH.2017.04.014.

Chehab, E.F., Favre, J., Erhart-Hledik, J.C. and Andriacchi, T.P. 2014. Baseline knee adduction and flexion moments during walking are both associated with 5 year cartilage changes in patients with medial knee osteoarthritis. *Osteoarthritis and Cartilage* 22(11), pp. 1833–1839. doi: 10.1016/J.JOCA.2014.08.009.

Choi, J.H., Kim, B.R., Kim, S.R., Nam, K.W., Lee, S.Y., Kim, W. Bin and Kim, Y.J. 2020. Physical Performance Correlates with Self-Reported Physical Function and Quality of Life in Patients at 3 Months after Total Knee Arthroplasty. *Annals of geriatric medicine and research* 24(2), p. 99. Available at: /pmc/articles/PMC7370793/ [Accessed: 17 June 2022].

Christensen, J.C., Blackburn, B., Browning, B., Wilbur, C., Trinity, J.D., Gililland, J.M. and Pelt, C.E. 2022. Patient-reported outcomes measurement information system physical function and knee injury and osteoarthritis outcome score relationship on performance measures in people undergoing total knee arthroplasty. *Disability and*

Rehabilitation, pp. 1–9. Available at: https://www.tandfonline.com/doi/abs/10.1080/09638288.2022.2134934.

Christensen, J.C., Capin, J.J., Hinrichs, L.A., Aljehani, M., Stevens-Lapsley, J.E. and Zeni, J.A. 2021a. Gait mechanics are influenced by quadriceps strength, age, and sex after total knee arthroplasty. *Journal of Orthopaedic Research*® 39(7), pp. 1523–1532. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/jor.24878 [Accessed: 7 November 2022].

Christensen, J.C., Pelt, C.E., Bo Foreman, K., LaStayo, P.C., Anderson, A.E., Gililland, J.M. and Mizner, R.L. 2021b. Longitudinal study of knee load avoidant movement behavior after total knee arthroplasty with recommendations for future retraining interventions. *The Knee* 30, pp. 90–99. doi: 10.1016/J.KNEE.2021.03.014.

Clement, N.D., Afzal, I., Demetriou, C., Deehan, D.J., Field, R.E. and Kader, D.F. 2020. The preoperative oxford knee score is an independent predictor of achieving a postoperative ceiling score after total knee arthroplasty. *Bone and Joint Journal* 102 B(11), pp. 1519–1526. doi: 10.1302/0301-620X.102B11.BJJ-2019-1327.R1/LETTERTOEDITOR.

Cohen, J. 1988. Statistical Power Analysis for the Behavioral Sciences. *Statistical Power Analysis for the Behavioral Sciences*. doi: 10.4324/9780203771587.

Coleman, G., Dobson, F., Hinman, R.S., Bennell, K. and White, D.K. 2020. Measures of Physical Performance. *Arthritis Care & Research* 72(S10), pp. 452–485. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/acr.24373 [Accessed: 14 May 2023].

Collins, N.J., Misra, D., Felson, D.T., Crossley, K.M. and Roos, E.M. 2011. Measures of knee function: International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form, Knee Injury and Osteoarthritis Outcome Score (KOOS), Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form (KOOS-PS), Knee Ou. *Arthritis Care & Research* 63(S11), pp. S208–S228. doi: 10.1002/acr.20632.

Collins, N.J., Prinsen, C.A.C., Christensen, R., Bartels, E.M., Terwee, C.B. and Roos, E.M. 2016. Knee Injury and Osteoarthritis Outcome Score (KOOS): systematic

review and meta-analysis of measurement properties. *Osteoarthritis and Cartilage* 24(8), pp. 1317–1329. doi: 10.1016/J.JOCA.2016.03.010.

Collins, N.J. and Roos, E.M. 2012. Patient-reported outcomes for total hip and knee arthroplasty: commonly used instruments and attributes of a “good” measure. *Clinics in geriatric medicine* 28(3), pp. 367–394. doi: 10.1016/J.CGER.2012.05.007.

Costello, K.E. et al. 2021. Ground reaction force patterns in knees with and without radiographic osteoarthritis and pain: descriptive analyses of a large cohort (the Multicenter Osteoarthritis Study). *Osteoarthritis and Cartilage* 29(8), pp. 1138–1146. doi: 10.1016/J.JOCA.2021.03.009.

Craig, D.G.N., Bates, C.M., Davidson, J.S., Martin, K.G., Hayes, P.C. and Simpson, K.J. 2012. Staggered overdose pattern and delay to hospital presentation are associated with adverse outcomes following paracetamol-induced hepatotoxicity. *British Journal of Clinical Pharmacology* 73(2), pp. 285–294. doi: 10.1111/j.1365-2125.2011.04067.x.

Creaby, M.W., Bennell, K.L. and Hunt, M.A. 2012. Gait Differs Between Unilateral and Bilateral Knee Osteoarthritis. *Archives of Physical Medicine and Rehabilitation* 93(5), pp. 822–827. doi: 10.1016/J.APMR.2011.11.029.

Della Croce, U., Leardini, A., Chiari, L. and Cappozzo, A. 2005. Human movement analysis using stereophotogrammetry: Part 4: assessment of anatomical landmark misplacement and its effects on joint kinematics. *Gait & Posture* 21(2), pp. 226–237. Available at: http://www.sciencedirect.com/science/article/pii/S0966636204000785.

Cross, M. et al. 2014. The global burden of hip and knee osteoarthritis: estimates from the global burden of disease 2010 study. *Annals of the Rheumatic Diseasesheumatic diseases* 73(7), pp. 1323–30. doi: 10.1136/annrheumdis-2013-204763.

Culliford, D., Maskell, J., Judge, A., Cooper, C., Prieto-Alhambra, D. and Arden, N.K. 2015. Future projections of total hip and knee arthroplasty in the UK: results from the UK Clinical Practice Research Datalink. *Osteoarthritis and Cartilage* 23(4), pp. 594–600. doi: 10.1016/J.JOCA.2014.12.022.

Cury, A.C., Pinto, R.Z., Madaleno, F.O. and Resende, R.A. 2021. Do older adults present altered pelvic and trunk movement pattern during gait? A systematic review

with meta-analysis and GRADE recommendations. *Brazilian Journal of Physical Therapy* 25(5), pp. 484–499. doi: 10.1016/J.BJPT.2021.01.003.

Cyteval, C. 2016. Imaging of knee implants and related complications. *Diagnostic and Interventional Imaging* 97(7–8), pp. 809–821. doi: 10.1016/J.DIII.2016.02.015.

Dakin, H., Gray, A., Fitzpatrick, R., MacLennan, G., Murray, D., Kat, T. and Group, T. 2012. Rationing of total knee replacement: a cost-effectiveness analysis on a large trial data set. *BMJ Open* 2(1), p. e000332. Available at: https://bmjopen.bmj.com/content/2/1/e000332 [Accessed: 13 June 2022].

Dakin, H.A., Eibich, P., Gray, A., Smith, J., Barker, K.L., Beard, D. and Price, A.J. 2020. Who gets referred for knee or hip replacement? A theoretical model of the potential impact of evidence-based referral thresholds using data from a retrospective review of clinic records from an English musculoskeletal referral hub. *BMJ open* 10(7). Available at: https://pubmed.ncbi.nlm.nih.gov/32620710/ [Accessed: 3 March 2023].

Dancey, C.P. and Reidy, John. 2011. *Statistics without maths for psychology*. Prentice Hall/Pearson.

D'Apuzzo, M.R., Novicoff, W.M. and Browne, J.A. 2015. The John Insall Award: Morbid Obesity Independently Impacts Complications, Mortality, and Resource Use After TKA. *Clinical Orthopaedics and Related Research* 473(1), pp. 57–63. Available at: https://link.springer.com/article/10.1007/s11999-014-3668-9 [Accessed: 5 May 2023].

Davis, R.B., Õunpuu, S., Tyburski, D. and Gage, J.R. 1991. A gait analysis data collection and reduction technique. *Human Movement Science* 10(5), pp. 575–587. doi: 10.1016/0167-9457(91)90046-Z.

Dawson, J., Fitzpatrick, R., Murray, D. and Carr, A. 1998. Questionnaire on the perceptions of patients about total knee replacement. *The Journal of Bone and Joint Surgery. British volume* 80-B(1), pp. 63–69. Available at: http://online.boneandjoint.org.uk/doi/10.1302/0301-620X.80B1.0800063 [Accessed: 12 May 2019].

Deluzio, K.J. and Astephen, J.L. 2007. Biomechanical features of gait waveform data associated with knee osteoarthritis. An application of principal component analysis. *Gait and Posture* 25(1), pp. 86–93. doi: 10.1016/j.gaitpost.2006.01.007.

Deluzio, K.J., Wyss, U.P., Costigan, P.A., Sorbie, C. and Zee, B. 1999. Gait assessment in unicompartmental knee arthroplasty patients: Principal component modelling of gait waveforms and clinical status. *Human Movement Science* 18(5), pp. 701–711. doi: 10.1016/S0167-9457(99)00030-5.

Deluzio, K.J., Wyss, U.P., Zee, B., Costigan, P.A. and Sorbie, C. 1997. Principal component models of knee kinematics and kinetics: Normal vs. pathological gait patterns. *Human Movement Science* 16(2–3), pp. 201–217. doi: 10.1016/S0167-9457(96)00051-6.

Dempster, A.P. 1968. A generalization of Bayesian inference (with discussion). *J Roy Stat Soc, Series B* 30(2), pp. 205–247.

Dempster, W.T. 1955. *Space requirements of the seated operator, geometrical, kinematic, and mechanical aspects of the body with special reference to the limbs*. Available at: https://apps.dtic.mil/sti/pdfs/AD0087892.pdf [Accessed: 9 March 2023].

Deshpande, B.R. et al. 2016. Number of Persons With Symptomatic Knee Osteoarthritis in the US: Impact of Race and Ethnicity, Age, Sex, and Obesity. *Arthritis Care & Research* 68(12), pp. 1743–1750. doi: 10.1002/acr.22897.

Dettoni, F., Bonasia, D.E., Castoldi, F., Bruzzone, M., Blonna, D. and Rossi, R. 2010. HIGH TIBIAL OSTEOTOMY VERSUS UNICOMPARTMENTAL KNEE ARTHROPLASTY FOR MEDIAL COMPARTMENT ARTHROSIS OF THE KNEE: A REVIEW OF THE LITERATURE. *The Iowa Orthopaedic Journal* 30, pp. 131–140. Available at: /pmc/articles/PMC2958284/ [Accessed: 11 January 2024].

Devasenapathy, N., Maddison, R., Malhotra, R., Zodepy, S., Sharma, S. and Belavy, D.L. 2019. Preoperative Quadriceps Muscle Strength and Functional Ability Predict Performance-Based Outcomes 6 Months After Total Knee Arthroplasty: A Systematic Review. *Physical Therapy* 99(1), pp. 46–61. Available at: https://academic.oup.com/ptj/article/99/1/46/5134174.

DeVita, P. and Hortobagyi, T. 2000. Age causes a redistribution of joint torques and powers during gait. *Journal of Applied Physiology* 88(5), pp. 1804–1811. Available

at: https://journals.physiology.org/doi/10.1152/jappl.2000.88.5.1804 [Accessed: 23 November 2022].

Diss, C.E. 2001. The reliability of kinetic and kinematic variables used to analyse normal running gait. *Gait & Posture* 14(2), pp. 98–103. doi: https://doi.org/10.1016/S0966-6362(01)00125-4.

Dobson, F. et al. 2013. OARSI recommended performance-based tests to assess physical function in people diagnosed with hip or knee osteoarthritis. *Osteoarthritis and Cartilage* 21, pp. 1042–1052. Available at: http://dx.doi.org/10.1016/j.joca.2013.05.002 [Accessed: 12 May 2019].

Dobson, F. et al. 2017. Reliability and measurement error of the Osteoarthritis Research Society International (OARSI) recommended performance-based tests of physical function in people with hip and knee osteoarthritis. *Osteoarthritis and Cartilage* 25(11), pp. 1792–1796. doi: 10.1016/J.JOCA.2017.06.006.

Dobson, F., Hinman, R.S., Hall, M., Terwee, C.B., Roos, E.M. and Bennell, K.L. 2012. Measurement properties of performance-based measures to assess physical function in hip and knee osteoarthritis: a systematic review. *Osteoarthritis and Cartilage* 20(12), pp. 1548–1562. Available at: http://www.sciencedirect.com/science/article/pii/S1063458412009387.

Donà, G., Preatoni, E., Cobelli, C., Rodano, R. and Harrison, A.J. 2009. Application of functional principal component analysis in race walking: An emerging methodology. *Sports Biomechanics* 8(4), pp. 284–301. doi: 10.1080/14763140903414425.

Duffell, L.D., Jordan, S.J., Cobb, J.P. and McGregor, A.H. 2017. Gait adaptations with aging in healthy participants and people with knee-joint osteoarthritis. *Gait & Posture* 57, pp. 246–251. doi: 10.1016/J.GAITPOST.2017.06.015.

Eckhard, L., Munir, S., Wood, D., Talbot, S., Brighton, R., Walter, B. and Baré, J. 2021. The ceiling effects of patient reported outcome measures for total knee arthroplasty. *Orthopaedics and Traumatology: Surgery and Research* 107(3), p. 102758. doi: 10.1016/j.otsr.2020.102758.

Emery, C.A. et al. 2019. Establishing outcome measures in early knee osteoarthritis. *Nature Reviews Rheumatology* 15(7), pp. 438–448. Available at: www.nature.com/nrrheum [Accessed: 25 February 2021].

Esbjörnsson, A.C., Rozumalski, A., Iversen, M.D., Schwartz, M.H., Wretenberg, P. and Broström, E.W. 2014. Quantifying gait deviations in individuals with rheumatoid arthritis using the Gait Deviation Index. *Scandinavian Journal of Rheumatology* 43(2), pp. 124–131. Available at: https://www.tandfonline.com/doi/abs/10.3109/03009742.2013.822095 [Accessed: 6 July 2021].

Esrafilian, A., Karimi, M.T., Amiri, P. and Fatoye, F. 2013. Performance of subjects with knee osteoarthritis during walking: Differential parameters. *Rheumatology International* 33(7), pp. 1753–1761. Available at: https://link.springer.com/article/10.1007/s00296-012-2639-2 [Accessed: 28 April 2023].

European Parliament and Council of the European Union. 2018. *General Data Protection Regulation (GDPR)*. Available at: https://gdpr.eu/tag/gdpr/ [Accessed: 20 April 2020].

Evans, J.T., Walker, R.W., Evans, J.P., Blom, A.W., Sayers, A. and Whitehouse, M.R. 2019. How long does a knee replacement last? A systematic review and meta-analysis of case series and national registry reports with more than 15 years of follow-up. *The Lancet* 393(10172), pp. 655–663. doi: 10.1016/S0140-6736(18)32531-5.

Fantozzi, S., Catani, F., Ensini, A., Leardini, A. and Giannini, S. 2006. Femoral rollback of cruciate-retaining and posterior-stabilized total knee replacements: In vivo fluoroscopic analysis during activities of daily living. *Journal of Orthopaedic Research* 24(12), pp. 2222–2229. doi: 10.1002/JOR.20306.

Favre, J., Erhart-Hledik, J.C. and Andriacchi, T.P. 2014. Age-related differences in sagittal-plane knee function at heel-strike of walking are increased in osteoarthritic patients. *Osteoarthritis and Cartilage* 22(3), pp. 464–471. doi: 10.1016/J.JOCA.2013.12.014.

Favre, J. and Jolles, B.M. 2016. Gait analysis of patients with knee osteoarthritis highlights a pathological mechanical pathway and provides a basis for therapeutic interventions. *EFORT Open Reviews* 1(10), p. 368. Available at: /pmc/articles/PMC5367582/ [Accessed: 28 April 2023].

Federolf, P.A., Boyer, K.A. and Andriacchi, T.P. 2013. Application of principal component analysis in clinical gait research: Identification of systematic differences between healthy and medial knee-osteoarthritic gait. *Journal of Biomechanics* 46(13), pp. 2173–2178. doi: 10.1016/J.JBIOMECH.2013.06.032.

Fejer, R. and Ruhe, A. 2012. What is the prevalence of musculoskeletal problems in the elderly population in developed countries? A systematic critical literature review. *Chiropractic and Manual Therapies* 20(1), pp. 1–52. Available at: https://link.springer.com/articles/10.1186/2045-709X-20-31 [Accessed: 18 January 2024].

Felson, D.T. et al. 2000. Osteoarthritis: New Insights. Part 1: The Disease and Its Risk Factors. *Annals of Internal Medicine* 133(8), p. 635. doi: 10.7326/0003-4819-133-8-200010170-00016.

Felson, D.T., Niu, J., Clancy, M., Sack, B., Aliabadi, P. and Zhang, Y. 2007. Effect of recreational physical activities on the development of knee osteoarthritis in older adults of different weights: The Framingham Study. *Arthritis & Rheumatism* 57(1), pp. 6–12. doi: 10.1002/art.22464.

Fitzpatrick, R., Dawson, J., Fitzpatrick, R., Murray, D. and Carr, A. 1998. Questionnaire on the perceptions of patients about total knee replacement. *The Journal of Bone & Joint Surgery British Volume* 80-B(1), pp. 63–69. doi: 10.1302/0301-620X.80B1.0800063.

Flegal, K.M., Carroll, M.D., Ogden, C.L. and Curtin, L.R. 2010. Prevalence and Trends in Obesity Among US Adults, 1999-2008. *JAMA* 303(3), p. 235. doi: 10.1001/jama.2009.2014.

Fortin, P.R. et al. 1999. Outcomes of Total Hip and Knee Replacement Preoperative Functional Status Predicts Outcomes at Six Months After Surgery. *Arthritis & Rheumatism* 42(8), pp. 1722–1728. Available at:

https://onlinelibrary.wiley.com/doi/pdf/10.1002/1529-0131(199908)42:8%3C1722::AID-ANR22%3E3.0.CO;2-R [Accessed: 3 June 2019].

Fortin, P.R. et al. 2002. Timing of total joint replacement affects clinical outcomes among patients with osteoarthritis of the hip or knee. *Arthritis & Rheumatism* 46(12), pp. 3327–3330. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/art.10631 [Accessed: 13 June 2022].

Freeman, M.A.R. and Pinskerova, V. 2005. The movement of the normal tibio-femoral joint. *Journal of Biomechanics* 38(2), pp. 197–208. doi: 10.1016/J.JBIOMECH.2004.02.006.

Galvin, C. 2019. *Healthy and Osteoarthritic Knee Kinematics. A 3D/2D Image Registration Study of Kneeling.* University of Canberra.

Galvin, C.R., Perriman, D.M., Lynch, J.T., Pickering, M.R., Newman, P., Smith, P.N. and Scarvell, J.M. 2019. Age has a minimal effect on knee kinematics: A cross-sectional 3D/2D image-registration study of kneeling. *The Knee* 26(5), pp. 988–1002. doi: 10.1016/J.KNEE.2019.07.012.

Galvin, C.R., Perriman, D.M., Newman, P.M., Lynch, J.T., Smith, P.N. and Scarvell, J.M. 2018. Squatting, lunging and kneeling provided similar kinematic profiles in healthy knees—A systematic review and meta-analysis of the literature on deep knee flexion kinematics. *The Knee* 25(4), pp. 514–530. doi: 10.1016/J.KNEE.2018.04.015.

Gandhi, R., Tsvetkov, D., Davey, J.R. and Mahomed, N.N. 2009a. Survival and clinical function of cemented and uncemented prostheses in total knee replacement. *The Journal of Bone and Joint Surgery. British volume* 91-B(7), pp. 889–895. doi: 10.1302/0301-620X.91B7.21702.

Gandhi, R., Tsvetkov, D., Davey, J.R., Syed, K.A. and Mahomed, N.N. 2009b. Relationship between self-reported and performance-based tests in a hip and knee joint replacement population. *Clinical Rheumatology* 28(3), pp. 253–257. Available at: http://link.springer.com/10.1007/s10067-008-1021-y [Accessed: 12 May 2019].

George, J., Gautam, D., Devasenapathy, N. and Malhotra, R. 2021. Is It Worth Delaying Total Knee Replacement as Late as Possible? A Cost-Effectiveness

Analysis Using a Markov Model in the Indian Setting. *Value in Health Regional Issues* 24, pp. 173–180. doi: 10.1016/J.VHRI.2020.12.009.

George, J., Piuzzi, N.S., Ng, M., Sodhi, N., Khlopas, A.A. and Mont, M.A. 2017. Association Between Body Mass Index and Thirty-Day Complications After Total Knee Arthroplasty. *The Journal of Arthroplasty* 33(3), pp. 865–871. Available at: https://europepmc.org/article/med/29107493 [Accessed: 5 May 2023].

Gerig, G., Welti, D., Guttmann, C.R.G., Colchester, A.C.F. and Székely, G. 2000. Exploring the discrimination power of the time domain for segmentation and characterization of active lesions in serial MR data. *Medical Image Analysis* 4(1), pp. 31–42. doi: 10.1016/S1361-8415(00)00005-0.

Giesinger, J.M. et al. 2018. Patient-reported outcome metrics following total knee arthroplasty are influenced differently by patients' body mass index. *Knee Surg Sports Traumatol Arthrosc* 26, pp. 3257–3264. Available at: https://doi.org/10.1007/s00167-018-4853-2 [Accessed: 12 May 2019].

Giesinger, J.M., Hamilton, D.F., Jost, B., Behrend, H. and Giesinger, K. 2015. WOMAC, EQ-5D and Knee Society Score Thresholds for Treatment Success After Total Knee Arthroplasty. *The Journal of Arthroplasty* 30(12), pp. 2154–2158. doi: 10.1016/J.ARTH.2015.06.012.

Giesinger, K., Hamilton, D.F. and Giesinger, J.M. 2014. Comparative responsiveness of outcome measures for total knee arthroplasty. *Osteoarthritis and Cartilage* (22), pp. 184–189. Available at: http://dx.doi.org/10.1016/j.joca.2013.11.001 [Accessed: 12 May 2019].

Gill, S. and McBurney, H. 2008. Reliability of performance-based measures in people awaiting joint replacement surgery of the hip or knee. *Physiotherapy Research International* 13(3), pp. 141–152. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/pri.411 [Accessed: 13 May 2023].

Givens, D.L., Eskildsen, S., Taylor, K.E., Faldowski, R.A. and Del Gaizo, D.J. 2018. Timed Up and Go test is predictive of Patient-Reported Outcomes Measurement Information System physical function in patients awaiting total knee arthroplasty. *Arthroplasty Today* 4(4), pp. 505–509. doi: 10.1016/j.artd.2018.07.010.

Goh, G.S., Baker, C.M., Tarabichi, S., Clark, S.C., Austin, M.S. and Lonner, J.H. 2022. The Paradox of Patient-Reported Outcome Measures: Should We Prioritize “Feeling Better” or “Feeling Good” After Total Knee Arthroplasty? *The Journal of Arthroplasty*. doi: 10.1016/J.ARTH.2022.04.017.

Gorton, G.E., Hebert, D.A. and Gannotti, M.E. 2009. Assessment of the kinematic variability among 12 motion analysis laboratories. *Gait and Posture* 29(3), pp. 398–402. doi: 10.1016/J.GAITPOST.2008.10.060.

Graff, C., Hohmann, E., Bryant, A.L. and Tetsworth, K. 2016. Subjective and objective outcome measures after total knee replacement: is there a correlation? *ANZ Journal of Surgery* 86(11), pp. 921–925. Available at: http://doi.wiley.com/10.1111/ans.13708 [Accessed: 12 May 2019].

Greenhalgh, T. 2019. *How to read a paper: the basics of evidence-based medicine and healthcare*. Sixth edit. Oxford: John Wiley & Sons Ltd.

Grieco, T.F., Sharma, A., Dessinger, G.M., Cates, H.E. and Komistek, R.D. 2018. In Vivo Kinematic Comparison of a Bicruciate Stabilized Total Knee Arthroplasty and the Normal Knee Using Fluoroscopy. *The Journal of Arthroplasty* 33(2), pp. 565–571. Available at: http://www.arthroplastyjournal.org/article/S0883540317308379/fulltext [Accessed: 31 August 2021].

Grood, E.S. and Suntay, W.J. 1983. A Joint Coordinate System for the Clinical Description of Three-Dimensional Motions: Application to the Knee. *Journal of Biomechanical Engineering* 105(2), pp. 136–144. Available at: https://dx.doi.org/10.1115/1.3138397 [Accessed: 28 June 2023].

Guermazi, A., Burstein, D., Conaghan, P., Eckstein, F., Hellio Le Graverand-Gastineau, M.P., Keen, H. and Roemer, F.W. 2008. Imaging in Osteoarthritis. *Rheumatic Disease Clinics of North America* 34(3), pp. 645–687. doi: 10.1016/J.RDC.2008.04.006.

Gutírrez, L.F., Ozturk, C., McVeigh, E.R. and Lederman, R.J. 2008. A practical global distortion correction method for an image intensifier based x-ray fluoroscopy system. *Medical physics* 35(3), p. 997. Available at: /pmc/articles/PMC2679378/ [Accessed: 19 June 2023].

Hall, A.J., Stubbs, B., Mamas, M.A., Myint, P.K. and Smith, T.O. 2016. Association between osteoarthritis and cardiovascular disease: Systematic review and meta-analysis. *European Journal of Preventive Cardiology* 23(9), pp. 938–946. doi: 10.1177/2047487315610663.

Häller, T.V., Schenk, P., Jud, L., Hoch, A., Götschi, T. and Zingg, P.O. 2021. Consistency of 3D femoral torsion measurement from MRI compared to CT gold standard. *BMC Musculoskeletal Disorders* 22(1), pp. 1–7. Available at: https://link.springer.com/articles/10.1186/s12891-021-04633-7 [Accessed: 18 June 2023].

Hamai, S., Okazaki, K., Shimoto, T., Nakahara, H., Higaki, H. and Iwamoto, Y. 2015. Continuous Sagittal Radiological Evaluation of Stair-Climbing in Cruciate-Retaining and Posterior-Stabilized Total Knee Arthroplasties Using Image-Matching Techniques. *The Journal of Arthroplasty* 30(5), pp. 864–869. doi: 10.1016/J.ARTH.2014.12.027.

Hamilton, D., Henderson, G.R., Gaston, P., MacDonald, D., Howie, C. and Simpson, A.H.R.W. 2012a. Comparative outcomes of total hip and knee arthroplasty: a prospective cohort study. *Postgraduate Medical Journal* 88(1045), pp. 627–31. doi: 10.1136/postgradmedj-2011-130715.

Hamilton, D.F. et al. 2013. What determines patient satisfaction with surgery? A prospective cohort study of 4709 patients following total joint replacement. *BMJ Open* 3(4), pp. 1–7. Available at: http://bmjopen.bmj.com [Accessed: 12 May 2019].

Hamilton, D.F., Gaston, P. and Simpson, A.H.R.W. 2012b. Is patient reporting of physical function accurate following total knee replacement? *The Journal of Bone and Joint Surgery* 94-B(11), pp. 1506–1510. Available at: http://online.boneandjoint.org.uk/doi/10.1302/0301-620X.94B11.30081 [Accessed: 8 October 2019].

Hanavan, J.E.P. 1964. *A MATHEMATICAL MODEL OF THE HUMAN BODY*. Available at: https://apps.dtic.mil/sti/citations/AD0608463 [Accessed: 9 March 2023].

Handschin, C. and Spiegelman, B.M. 2008. The role of exercise and PGC1α in inflammation and chronic disease. *Nature* 454(7203), pp. 463–469. doi: 10.1038/nature07206.

Hanna, F.S. et al. 2009. Women have increased rates of cartilage loss and progression of cartilage defects at the knee than men: A gender study of adults without clinical knee osteoarthritis. *Menopause* 16(4), pp. 666–670. Available at: https://journals.lww.com/menopausejournal/Fulltext/2009/16040/Women_have_increased_rates_of_cartilage_loss_and.13.aspx [Accessed: 15 May 2023].

Harrington, M.E., Zavatsky, A.B., Lawson, S.E.M., Yuan, Z. and Theologis, T.N. 2007. Prediction of the hip joint centre in adults, children, and patients with cerebral palsy based on magnetic resonance imaging. *Journal of Biomechanics* 40(3), pp. 595–602. doi: https://doi.org/10.1016/j.jbiomech.2006.02.003.

Harris, K., Lim, C.R., Dawson, J., Fitzpatrick, R., Beard, D.J. and Price, A.J. 2017. The Oxford knee score and its subscales do not exhibit a ceiling or a floor effect in knee arthroplasty patients: an analysis of the National Health Service PROMs data set. *Knee Surgery, Sports Traumatology, Arthroscopy* 25(9), pp. 2736–2742. Available at: https://link.springer.com/article/10.1007/s00167-015-3788-0 [Accessed: 6 July 2023].

Harris, K.K., Dawson, J., Jones, L.D., Beard, D.J. and Price, A.J. 2013. Extending the use of PROMs in the NHS—using the Oxford Knee Score in patients undergoing non-operative management for knee osteoarthritis: a validation study. *BMJ Open* 3(8), p. e003365. Available at: https://bmjopen.bmj.com/content/3/8/e003365 [Accessed: 6 July 2023].

Hatfield, G.L., Stanish, W.D. and Hubley-Kozey, C.L. 2015. Three-Dimensional Biomechanical Gait Characteristics at Baseline Are Associated With Progression to Total Knee Arthroplasty. *Arthritis Care & Research* 67(7), pp. 1004–1014. doi: 10.1002/acr.22564.

Hawker, G.A., Mian, S., Kendzerska, T. and French, M. 2011. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). *Arthritis Care & Research* 63(S11), pp. S240–S252. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/acr.20543 [Accessed: 27 June 2023].

Henery, R.J. 1994. Methods for comparison. In: Michie, D., Spiegelhalter, D. J., and Taylor, C. C. eds. *Machine learning, neural and statistical classification*. Ellis Horwood Workshops Ltd., pp. 107–124.

Hensor, E.M.A., Dube, B., Kingsbury, S.R., Tennant, A. and Conaghan, P.G. 2015. Toward a Clinical Definition of Early Osteoarthritis: Onset of Patient-Reported Knee Pain Begins on Stairs. Data From the Osteoarthritis Initiative. *Arthritis Care & Research* 67(1), pp. 40–47. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/acr.22418 [Accessed: 27 June 2023].

Herssens, N., Verbecque, E., Hallemans, A., Vereeck, L., Van Rompaey, V. and Saeys, W. 2018. Do spatiotemporal parameters and gait variability differ across the lifespan of healthy adults? A systematic review. *Gait & Posture* 64, pp. 181–190. doi: 10.1016/J.GAITPOST.2018.06.012.

Hinman, R.S., Hunt, M.A., Creaby, M.W., Wrigley, T. v., McManus, F.J. and Bennell, K.L. 2010. Hip muscle weakness in individuals with medial knee osteoarthritis. *Arthritis Care & Research* 62(8), pp. 1190–1193. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/acr.20199 [Accessed: 23 November 2022].

Hofstede, S.N. et al. 2015. Mobile bearing vs fixed bearing prostheses for posterior cruciate retaining total knee arthroplasty for postoperative functional status in patients with osteoarthritis and rheumatoid arthritis. *Cochrane Database of Systematic Reviews* (2). doi: 10.1002/14651858.CD003130.pub3.

Houglum, P.A., Bertoti, D. and Brunnstrom, Signe. 2012. *Brunnstrom's clinical kinesiology.* Sixth. Philadelphia: F.A. Davis Company. Available at: https://books.google.co.uk/books?id=offuAwAAQBAJ&pg=PA456&dq=clinical+kinesiology+60%25+tibia&hl=en&sa=X&ved=0ahUKEwiKv_34k5TiAhVKTxUIHUexAP0Q6AEIKTAA#v=onepage&q=clinical kinesiology 60%25 tibia&f=false [Accessed: 11 May 2019].

Huber, E.O., Meichtry, A., de Bie, R.A. and Bastiaenen, C.H. 2016. Construct validity of change scores of the Chair Stand Test versus Timed Up and Go Test, KOOS questionnaire and the isometric muscle strength test in patients with severe knee osteoarthritis undergoing total knee replacement. *Manual Therapy* 21, pp. 262–267. doi: 10.1016/J.MATH.2015.09.012.

Hubley-Kozey, C.L. and Astephen Wilson, J. 2018. Effects of Knee Osteoarthritis and Joint Replacement Surgery on Gait. In: Müller, B. and Wolf, S. I. eds. *Handbook of Human Motion*. Berlin: Springer International Publishing, pp. 1521–1549. doi: 10.1007/978-3-319-14418-4/COVER.

Hunt, M.A., Birmingham, T.B., Bryant, D., Jones, I., Giffin, J.R., Jenkyn, T.R. and Vandervoort, A.A. 2008. Lateral trunk lean explains variation in dynamic knee joint load in patients with medial compartment knee osteoarthritis. *Osteoarthritis and Cartilage* 16(5), pp. 591–599. Available at: http://www.sciencedirect.com/science/article/pii/S1063458407003603.

Hunt, M.A., Birmingham, T.B., Giffin, J.R. and Jenkyn, T.R. 2006. Associations among knee adduction moment, frontal plane ground reaction force, and lever arm during walking in patients with knee osteoarthritis. *Journal of Biomechanics* 39(12), pp. 2213–2220. doi: 10.1016/J.JBIOMECH.2005.07.002.

Hunt, M.A., Simic, M., Hinman, R.S., Bennell, K.L. and Wrigley, T. V. 2011. Feasibility of a gait retraining strategy for reducing knee joint loading: Increased trunk lean guided by real-time biofeedback. *Journal of Biomechanics* 44(5), pp. 943–947. doi: 10.1016/J.JBIOMECH.2010.11.027.

Hunt, M.A., Wrigley, T. V., Hinman, R.S. and Bennell, K.L. 2010. Individuals with severe knee osteoarthritis (OA) exhibit altered proximal walking mechanics compared with individuals with less severe OA and those without knee pain. *Arthritis Care & Research* 62(10), pp. 1426–1432. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/acr.20248 [Accessed: 4 June 2023].

Hussain, S.M. et al. 2014. Incidence of Total Knee and Hip Replacement for Osteoarthritis in Relation to Circulating Sex Steroid Hormone Concentrations in Women. *Arthritis & Rheumatology* 66(8), pp. 2144–2151. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/art.38651 [Accessed: 15 May 2023].

Husted, J.A., Cook, R.J., Farewell, V.T. and Gladman, D.D. 2000. *Methods for assessing responsiveness: a critical review and recommendations*.

Iijima, H., Shimoura, K., Aoyama, T. and Takahashi, M. 2018. Biomechanical characteristics of stair ambulation in patients with knee OA: A systematic review with meta-analysis toward a better definition of clinical hallmarks. *Gait & Posture* 62, pp.

191–201. Available at: http://www.sciencedirect.com/science/article/pii/S0966636218301322.

Iijima, H., Shimoura, K., Ono, T., Aoyama, T. and Takahashi, M. 2019. Proximal gait adaptations in individuals with knee osteoarthritis: A systematic review and meta-analysis. *Journal of Biomechanics* 87, pp. 127–141. doi: 10.1016/j.jbiomech.2019.02.027.

Innocenti, B. 2022. Knee prosthesis: biomechanics and design. In: *Human Orthopaedic Biomechanics: Fundamentals, Devices and Applications*. Academic Press, pp. 377–407. doi: 10.1016/B978-0-12-824481-4.00015-9.

Ismailidis, P. et al. 2020. Kinematic changes in patients with severe knee osteoarthritis are a result of reduced walking speed rather than disease severity. *Gait & Posture* 79, pp. 256–261. doi: 10.1016/J.GAITPOST.2020.05.008.

Jacobs, C.A. and Christensen, C.P. 2009. Correlations between knee society function scores and functional force measures. *Clinical Orthopaedics and Related Research* 467(9), pp. 2414–2419. Available at: https://link.springer.com/article/10.1007/s11999-009-0811-0 [Accessed: 24 February 2021].

Jensen, C., Penny, J.T., Nielsen, D.B., Overgaard, S. and Holsgaard-Larsen, A. 2015a. Quantifying Gait Quality in Patients with Large-Head and Conventional Total Hip Arthroplasty-A Prospective Cohort Study. *Journal of Arthroplasty* 30(12), pp. 2343-2348.e1. doi: 10.1016/j.arth.2015.06.049.

Jensen, C., Rosenlund, S., Nielsen, D.B., Overgaard, S. and Holsgaard-Larsen, A. 2015b. The use of the Gait Deviation Index for the evaluation of participants following total hip arthroplasty: An explorative randomized trial. *Gait and Posture* 42(1), pp. 36–41. doi: 10.1016/j.gaitpost.2015.02.009.

Jiang, C., Liu, Z., Wang, Y., Bian, Y., Feng, B. and Weng, X. 2016. Posterior Cruciate Ligament Retention versus Posterior Stabilization for Total Knee Arthroplasty: A Meta-Analysis. *PLOS ONE* 11(1), p. e0147865. Available at: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0147865 [Accessed: 11 January 2024].

Jiang, Y., Sanchez-Santos, M.T., Judge, A.D., Murray, D.W. and Arden, N.K. 2017. Predictors of Patient-Reported Pain and Functional Outcomes Over 10 Years After Primary Total Knee Arthroplasty: A Prospective Cohort Study. *The Journal of Arthroplasty* 32(1), pp. 92-100.e2. doi: 10.1016/J.ARTH.2016.06.009.

Johnston, J.D., Masri, B.A. and Wilson, D.R. 2009. Computed tomography topographic mapping of subchondral density (CT-TOMASD) in osteoarthritic and normal knees: methodological development and preliminary findings. *Osteoarthritis and Cartilage* 17(10), pp. 1319–1326. doi: 10.1016/J.JOCA.2009.04.013.

Jones, L. 2004. *The development of a novel method for the classification of the osteoarthritic and normal knee function*. Cardiff University.

Jones, L., Beynon, M.J., Holt, C.A. and Roy, S. 2006. An application of the Dempster–Shafer theory of evidence to the classification of knee function and detection of improvement due to total knee replacement surgery. *Journal of Biomechanics* 39(13), pp. 2512–2520. doi: 10.1016/J.JBIOMECH.2005.07.024.

Jones, L., Holt, C.A. and Beynon, M.J. 2008. Reduction, classification and ranking of motion analysis data: an application to osteoarthritic and normal knee function data. *Computer Methods in Biomechanics and Biomedical Engineering* 11(1), pp. 31–40. Available at: https://www.tandfonline.com/doi/abs/10.1080/10255840701550956 [Accessed: 14 October 2021].

Joyce, C.R.B., Zutshi, D.W., Hrubes, V. and Mason, R.M. 1975. Comparison of fixed interval and visual analogue scales for rating chronic pain. *European Journal of Clinical Pharmacology* 8(6), pp. 415–420. Available at: https://link.springer.com/article/10.1007/BF00562315 [Accessed: 27 June 2023].

Julin, J., Jämsen, E., Puolakka, T., Konttinen, Y.T. and Moilanen, T. 2010. Younger age increases the risk of early prosthesis failure following primary total knee replacement for osteoarthritis. *New Pub: Medical Journals Sweden* 81(4), pp. 413–419. Available at: https://www.tandfonline.com/doi/abs/10.3109/17453674.2010.501747 [Accessed: 5 May 2023].

Jung, J.H., Bang, C.H., Song, G.G., Kim, C., Kim, J.H. and Choi, S.J. 2019. Knee osteoarthritis and menopausal hormone therapy in postmenopausal women: A

nationwide cross-sectional study. *Menopause* 26(6), pp. 598–602. Available at: https://journals.lww.com/menopausejournal/Fulltext/2019/06000/Knee_osteoarthritis_and_menopausal_hormone_therapy.6.aspx [Accessed: 15 May 2023].

Kage, T. et al. 2021. The Association between in Vivo Knee Kinematics and Patient-Reported Outcomes during Squatting in Bicruciate-Stabilized Total Knee Arthroplasty. *Journal of Knee Surgery*. Available at: http://www.thieme-connect.com/products/ejournals/html/10.1055/s-0041-1723763 [Accessed: 7 July 2022].

Kaiser, H.F. 1960. The application of electronic computers to factor analysis. *Educational and psychological measurement.* 20(1), pp. 141–151.

Kaufman, K., Miller, E., Kingsbury, T., Russell Esposito, E., Wolf, E., Wilken, J. and Wyatt, M. 2016a. Reliability of 3D gait data across multiple laboratories. *Gait & Posture* 49, pp. 375–381. doi: https://doi.org/10.1016/j.gaitpost.2016.07.075.

Kaufman, K., Miller, E., Kingsbury, T., Russell Esposito, E., Wolf, E., Wilken, J. and Wyatt, M. 2016b. Reliability of 3D gait data across multiple laboratories. *Gait & Posture* 49, pp. 375–381. doi: 10.1016/J.GAITPOST.2016.07.075.

Kefala, V., Cyr, A.J., Harris, M.D., Hume, D.R., Davidson, B.S., Kim, R.H. and Shelburne, K.B. 2017. Assessment of Knee Kinematics in Older Adults Using High-Speed Stereo Radiography. *Medicine and science in sports and exercise* 49(11), p. 2260. Available at: /pmc/articles/PMC8230732/ [Accessed: 7 July 2022].

Kellgren, J.H. and Lawrence, J.S. 1957. Radiological assessment of osteo-arthrosis. *Annals of the rheumatic diseases* 16(4), pp. 494–502. Available at: http://www.ncbi.nlm.nih.gov/pubmed/13498604 [Accessed: 13 March 2019].

Kennedy, D.M., Hanna, S.E., Stratford, P.W., Wessel, J. and Gollish, J.D. 2006. Preoperative Function and Gender Predict Pattern of Functional Recovery After Hip and Knee Arthroplasty. *The Journal of Arthroplasty* 21(4), pp. 559–566. doi: 10.1016/J.ARTH.2005.07.010.

Kennedy, D.M., Stratford, P.W., Wessel, J., Gollish, J.D. and Penney, D. 2005. Assessing stability and change of four performance measures: A longitudinal study evaluating outcome following total hip and knee arthroplasty. *BMC Musculoskeletal Disorders* 6(1), pp. 1–12. Available at:

https://link.springer.com/articles/10.1186/1471-2474-6-3 [Accessed: 13 August 2020].

Kerkhof, H.J.M. et al. 2014. Prediction model for knee osteoarthritis incidence, including clinical, genetic and biochemical risk factors. *Annals of the rheumatic diseases* 73(12), pp. 2116–21. doi: 10.1136/annrheumdis-2013-203620.

Khatib, Y., Madan, A., Naylor, J.M. and Harris, I.A. 2015. Do Psychological Factors Predict Poor Outcome in Patients Undergoing TKA? A Systematic Review. *Clinical Orthopaedics and Related Research* 473(8), pp. 2630–2638. Available at: http://link.springer.com/10.1007/s11999-015-4234-9 [Accessed: 7 January 2021].

Kiadaliri, A.A., Lohmander, L.S., Moradi-Lakeh, M., Petersson, I.F. and Englund, M. 2018. High and rising burden of hip and knee osteoarthritis in the Nordic region, 1990–2015. *Acta Orthopaedica* 89(2), pp. 177–183. doi: 10.1080/17453674.2017.1404791.

Kitagawa, A., Tsumura, N., Chin, T., Gamada, K., Banks, S.A. and Kurosaka, M. 2010. In Vivo Comparison of Knee Kinematics Before and After High-Flexion Posterior Cruciate-Retaining Total Knee Arthroplasty. *The Journal of Arthroplasty* 25(6), pp. 964–969. doi: 10.1016/J.ARTH.2009.07.008.

Klem, N.-R. et al. 2020. Satisfaction after total knee replacement for osteoarthritis is usually high, but what are we measuring? A systematic review. *Osteoarthritis and Cartilage Open* 2(1), p. 100032. doi: 10.1016/J.OCARTO.2020.100032.

Kluge, F. et al. 2018. Pre-operative sensor-based gait parameters predict functional outcome after total knee arthroplasty. *Gait and Posture* 66, pp. 194–200. doi: 10.1016/j.gaitpost.2018.08.026.

Knofczynski, G.T. and Mundfrom, D. 2008. Sample Sizes When Using Multiple Linear Regression for Prediction. *Educational and psychological measurement* 68(3), pp. 431–442.

Ko, Y., Lo, N.N., Yeo, S.J., Yang, K.Y., Yeo, W., Chong, H.C. and Thumboo, J. 2013. Comparison of the responsiveness of the SF-36, the Oxford Knee Score, and the Knee Society Clinical Rating System in patients undergoing total knee replacement. *Quality of Life Research* 22(9), pp. 2455–2459. Available at: https://link.springer.com/article/10.1007/s11136-013-0376-y [Accessed: 6 July 2023].

Kobsar, D., Charlton, J.M. and Hunt, M.A. 2019. Individuals with knee osteoarthritis present increased gait pattern deviations as measured by a knee-specific gait deviation index. *Gait and Posture* 72, pp. 82–88. doi: 10.1016/j.gaitpost.2019.05.020.

Kramers-de Quervain, I.A., Kämpfen, S., Munzinger, U. and Mannion, A.F. 2012. Prospective study of gait function before and 2 years after total knee arthroplasty. *The Knee* 19(5), pp. 622–627. doi: 10.1016/J.KNEE.2011.12.009.

Kroman, S.L., Roos, E.M., Bennell, K.L., Hinman, R.S. and Dobson, F. 2014. Measurement properties of performance-based outcome measures to assess physical function in young and middle-aged people known to be at high risk of hip and/or knee osteoarthritis: a systematic review. *Osteoarthritis and Cartilage* 22(1), pp. 26–39. Available at: http://www.sciencedirect.com/science/article/pii/S1063458413010042.

Kurihara, Y. et al. 2021. Relationships between early postoperative gait biomechanical factors and patient-reported outcome measures 6 months after total knee arthroplasty. *Knee* 28, pp. 354–361. doi: 10.1016/j.knee.2020.12.021.

Kuroyanagi, Y., Mu, S., Hamai, S., Robb, W.J. and Banks, S.A. 2012. In Vivo Knee Kinematics During Stair and Deep Flexion Activities in Patients With Bicruciate Substituting Total Knee Arthroplasty. *The Journal of Arthroplasty* 27(1), pp. 122–128. doi: 10.1016/J.ARTH.2011.03.005.

Kurtz, S., Ong, K., Lau, E., Mowat, F. and Halpern, M. 2007. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. *The Journal of Joint and Bone Surgery* 89(4), pp. 780–785.

Langley, B., Page, R. and Greig, M. 2019. The influence of different pelvic technical marker sets upon hip kinematics during gait. *Gait & Posture* 71, pp. 74–78. doi: https://doi.org/10.1016/j.gaitpost.2019.04.012.

Leardini, A., Chiari, L., Croce, U. Della and Cappozzo, A. 2005. Human movement analysis using stereophotogrammetry: Part 3. Soft tissue artifact assessment and compensation. *Gait & Posture* 21(2), pp. 212–225. Available at: http://www.sciencedirect.com/science/article/pii/S0966636204000773.

Lee, S.J. et al. 2020. Preoperative physical factors that predict stair-climbing ability at one month after total knee arthroplasty. *Journal of Rehabilitation Medicine* 52(5),

pp. 1–8. Available at: https://medicaljournalssweden.se/jrm/article/view/3799 [Accessed: 1 May 2023].

Levinger, P., Menz, H.B., Morrow, A.D., Feller, J.A., Bartlett, J.R. and Bergman, N.R. 2013. Lower limb biomechanics in individuals with knee osteoarthritis before and after total knee arthroplasty surgery. *Journal of Arthroplasty* 28(6), pp. 994–999. doi: 10.1016/j.arth.2012.10.018.

Li, J.S., Hosseini, A., Cancre, L., Ryan, N., Rubash, H.E. and Li, G. 2013. Kinematic characteristics of the tibiofemoral joint during a step-up activity. *Gait & Posture* 38(4), pp. 712–716. doi: 10.1016/J.GAITPOST.2013.03.004.

Liebensteiner, M.C., Herten, A., Gstoettner, M., Thaler, M., Krismer, M. and Bach, C.M. 2008. Correlation between objective gait parameters and subjective score measurements before and after total knee arthroplasty. *The Knee* 15(6), pp. 461–6. doi: 10.1016/j.knee.2008.07.001.

Liebs, T.R., Herzberg, W., Roth-Kroeger, A.M., Rüther, W. and Hassenpflug, J. 2011. Women recover faster than men after standard knee arthroplasty. *Clinical Orthopaedics and Related Research* 469(10), pp. 2855–2865. Available at: https://link.springer.com/article/10.1007/s11999-011-1921-z [Accessed: 6 April 2023].

Lin, Y.C., Davey, R.C. and Cochrane, T. 2001. Tests for physical function of the elderly with knee and hip osteoarthritis. *Scandinavian Journal of Medicine & Science in Sports* 11(5), pp. 280–286. Available at: https://onlinelibrary.wiley.com/doi/full/10.1034/j.1600-0838.2001.110505.x [Accessed: 14 May 2023].

Linley, H.S., Sled, E.A., Culham, E.G. and Deluzio, K.J. 2010. A biomechanical analysis of trunk and pelvis motion during gait in subjects with knee osteoarthritis compared to control subjects. *Clinical Biomechanics* 25(10), pp. 1003–1010. Available at: http://www.sciencedirect.com/science/article/pii/S0268003310002056.

Loening, A.M. et al. 2000. Injurious Mechanical Compression of Bovine Articular Cartilage Induces Chondrocyte Apoptosis. *Archives of Biochemistry and Biophysics* 381(2), pp. 205–212. doi: 10.1006/ABBI.2000.1988.

Longo, U.G., Ciuffreda, M., Mannering, N., D'Andrea, V., Cimmino, M. and Denaro, V. 2018. Patellar Resurfacing in Total Knee Arthroplasty: Systematic Review and Meta-Analysis. *The Journal of Arthroplasty* 33(2), pp. 620–632. doi: 10.1016/J.ARTH.2017.08.041.

Losina, E. et al. 2013. Lifetime Risk and Age at Diagnosis of Symptomatic Knee Osteoarthritis in the US. *Arthritis Care & Research* 65(5), pp. 703–711. doi: 10.1002/acr.21898.

Luyten, F.P., Denti, M., Filardo, G., Kon, E. and Engebretsen, L. 2012. Definition and classification of early osteoarthritis of the knee. *Knee Surgery, Sports Traumatology, Arthroscopy* 20(3), pp. 401–406. doi: 10.1007/s00167-011-1743-2.

Lynch, J.T. et al. 2019. Statistical shape modelling reveals large and distinct subchondral bony differences in osteoarthritic knees. *Journal of Biomechanics* 93, pp. 177–184. doi: 10.1016/J.JBIOMECH.2019.07.003.

Lynch, J.T. et al. 2020. Shape is only a weak predictor of deep knee flexion kinematics in healthy and osteoarthritic knees. *Journal of Orthopaedic Research* 38(10), pp. 2250–2261. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/jor.24622 [Accessed: 3 September 2021].

Lynch, J.T., Perriman, D.M., Scarvell, J.M., Pickering, M.R., Galvin, C.R., Neeman, T. and Smith, P.N. 2021. The influence of total knee arthroplasty design on kneeling kinematics: A prospective randomized clinical trial. *Bone and Joint Journal* 103(1), pp. 105–112. Available at: https://online.boneandjoint.org.uk/doi/abs/10.1302/0301-620X.103B1.BJJ-2020-0958.R1 [Accessed: 7 July 2022].

Maempel, J.F., Clement, N.D., Brenkel, I.J. and Walmsley, P.J. 2016. Range of movement correlates with the Oxford knee score after total knee replacement: A prediction model and validation. *Knee* 23(3), pp. 511–516. doi: 10.1016/j.knee.2016.01.009.

Malahias, M.A., Gu, A., Addona, J., Nocon, A.A., Carli, A. v. and Sculco, P.K. 2019. Different clinical outcomes on the second side after staged total knee replacement. A systematic review. *The Knee* 26(3), pp. 530–536. doi: 10.1016/J.KNEE.2019.04.011.

Mandzuk, L.L., McMillan, D.E. and Bohm, E.R. 2015. A longitudinal study of quality of life and functional status in total hip and total knee replacement patients. *International Journal of Orthopaedic and Trauma Nursing* 19(2), pp. 102–113. doi: 10.1016/j.ijotn.2014.07.001.

Mayhew, D., Mendonca, V. and Murthy, B.V.S. 2019. A review of ASA physical status – historical perspectives and modern developments. *Anaesthesia* 74(3), pp. 373–379. Available at: https://onlinelibrary.wiley.com/doi/full/10.1111/anae.14569 [Accessed: 6 January 2023].

Maynard, V., Bakheit, A.M.O., Oldham, J. and Freeman, J. 2003. Intra-rater and inter-rater reliability of gait measurements with CODA mpx30 motion analysis system. *Gait & Posture* 17(1), pp. 59–67. doi: https://doi.org/10.1016/S0966-6362(02)00051-6.

McAlindon, T.E. et al. 2014. OARSI guidelines for the non-surgical management of knee osteoarthritis. *Osteoarthritis and Cartilage* 22(3), pp. 363–388. doi: 10.1016/J.JOCA.2014.01.003.

McClelland, J.A., Feller, J.A., Menz, H.B. and Webster, K.E. 2017. Patients with total knee arthroplasty do not use all of their available range of knee flexion during functional activities. *Clinical Biomechanics* 43, pp. 74–78. doi: 10.1016/J.CLINBIOMECH.2017.01.022.

McClelland, J.A., Feller, J.A. and Webster, K.E. 2018. Sex Differences in Gait After Total Knee Arthroplasty. *The Journal of Arthroplasty* 33(3), pp. 897–902. Available at: https://www.sciencedirect.com/science/article/pii/S0883540317308690 [Accessed: 15 May 2019].

McClelland, J.A., Grant, C., Webster, K.E. and Feller, J.A. 2007a. Alternative pelvic markers for Plug-In-Gait. In: *Gait & Posture 26 (Suppl. 1), S81*.

McClelland, J.A., Webster, K.E. and Feller, J.A. 2007b. Gait analysis of patients following total knee replacement: A systematic review. *The Knee* 14(4), pp. 253–263. doi: 10.1016/J.KNEE.2007.04.003.

McClelland, J.A., Wittwer, J.E., Webster, K.E. and Feller, J.A. 2014. Long-term assessment of gait biomechanics in patients with total knee arthroplasty. *Osteoarthritis and Cartilage* 22, p. S90. Available at:

http://www.oarsijournal.com/article/S1063458414002155/fulltext [Accessed: 4 October 2021].

McGibbon, C.A. and Krebs, D.E. 2002. Compensatory gait mechanics in patients with unilateral knee arthritis. *The Journal of Rheumatology* 29(11).

McGinley, J.L., Baker, R., Wolfe, R. and Morris, M.E. 2009. The reliability of three-dimensional kinematic gait measurements: A systematic review. *Gait & Posture* 29(3), pp. 360–369. doi: 10.1016/J.GAITPOST.2008.09.003.

McGinnis, K., Snyder-Mackler, L., Flowers, P. and Zeni, J. 2013. Dynamic joint stiffness and co-contraction in subjects after total knee arthroplasty. *Clinical Biomechanics* 28(2), pp. 205–210. doi: 10.1016/J.CLINBIOMECH.2012.11.008.

McKean, K.A., Landry, S.C., Hubley-Kozey, C.L., Dunbar, M.J., Stanish, W.D. and Deluzio, K.J. 2007. Gender differences exist in osteoarthritic gait. *Clinical Biomechanics* 22(4), pp. 400–409. Available at: https://www.sciencedirect.com/science/article/pii/S0268003306002221 [Accessed: 28 August 2019].

Mehta, S.P., Morelli, N., Prevatte, C., White, D. and Oliashirazi, A. 2019. Validation of Physical Performance Tests in Individuals with Advanced Knee Osteoarthritis. *Hospital for Special Surgery Journal* 15, pp. 261–268. doi: 10.1007/s11420-019-09702-1.

Metcalfe, A.J. 2014. *Knee osteoarthritis is a bilateral disease*. Cardiff University.

Metcalfe, A.J., Stewart, C.J., Postans, N.J., Biggs, P.R., Whatling, G.M., Holt, C.A. and Roberts, A.P. 2017. Abnormal loading and functional deficits are present in both limbs before and after unilateral knee arthroplasty. *Gait and Posture* 55, pp. 109–115. doi: 10.1016/j.gaitpost.2017.04.008.

Michael, J.W.-P., Schlüter-Brust, K.U. and Eysel, P. 2010. The epidemiology, etiology, diagnosis, and treatment of osteoarthritis of the knee. *Deutsches Arzteblatt international* 107(9), pp. 152–62. doi: 10.3238/arztebl.2010.0152.

Miller, D.E. and Kunce, J.T. 1973. Prediction and Statistical Overkill Revisited. *Measurement and Evaluation in Guidance* 6(3), pp. 157–163. doi: 10.1080/00256307.1973.12022590.

Mills, K., Hunt, M.A. and Ferber, R. 2013. Biomechanical Deviations During Level Walking Associated With Knee Osteoarthritis: A Systematic Review and Meta-Analysis. *Arthritis Care & Research* 65(10), pp. 1643–1665. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/acr.22015 [Accessed: 28 April 2023].

Mills, K., Wymenga, A.B., Bénard, M.R., Kaptein, B.L., Defoort, K.C., van Hellemondt, G.G. and Heesterbeek, P.J.C. 2023. Fluoroscopic and radiostereometric analysis of a bicruciate-retaining versus a posterior cruciate-retaining total knee arthroplasty: a randomized controlled trial. *The bone & joint journal* 105-B(1), pp. 35–46. doi: 10.1302/0301-620X.105B1.BJJ-2022-0465.R2/LETTERTOEDITOR.

Milner, C.E. 2009. Is gait normal after total knee arthroplasty? Systematic review of the literature. *Journal of Orthopaedic Science 2009 14:1* 14(1), pp. 114–120. Available at: https://link.springer.com/article/10.1007/s00776-008-1285-8 [Accessed: 29 November 2022].

Mizner, R.L., Petterson, S.C., Clements, K.E., Zeni, J.A., Irrgang, J.J. and Snyder-Mackler, L. 2011. Measuring Functional Improvement After Total Knee Arthroplasty Requires Both Performance-Based and Patient-Report Assessments A Longitudinal Analysis of Outcomes. *Journal of Arthroplasty* 26, pp. 728–737. Available at: https://pdf.sciencedirectassets.com/272519/1-s2.0-S0883540311X00049/1-s2.0-S0883540310003499/main.pdf?x-amz-security-token=AgoJb3JpZ2luX2VjEK%2F%2F%2F%2F%2F%2F%2F%2F%2F%2FwEaCXVzLWVhc3QtMSJHMEUCIFZXpPXN4nI4Ji%2FBq2s23DeCJLKA3OR%2BH2II5DozhNbUAiEAg9Uc3k [Accessed: 12 May 2019].

Mizner, R.L., Petterson, S.C., Stevens, J.E., Axe, M.J. and Snyder-Mackler, L. 2005. Preoperative Quadriceps Strength Predicts Functional Ability One Year After Total Knee Arthroplasty. *The Journal of Rheumatology* 32(8), pp. 1533–1539. Available at: www.jrheum.org [Accessed: 24 November 2022].

Moonot, P., Shang, M., Railton, G.T., Field, R.E. and Banks, S.A. 2010. In vivo weight-bearing kinematics with medial rotation knee arthroplasty. *The Knee* 17(1), pp. 33–37. doi: 10.1016/J.KNEE.2009.06.009.

Mostafaee, N., Rashidi, F., Negahban, H. and Ebrahimzadeh, M.H. 2022. Responsiveness and minimal important changes of the OARSI core set of performance-based measures in patients with knee osteoarthritis following

physiotherapy intervention. *Physiotherapy Theory and Practice*, pp. 1–12. Available at: https://www.tandfonline.com/doi/abs/10.1080/09593985.2022.2143253.

Muijs, D. 2011. *Doing Quantitative Research in Education with SPSS*. 2nd ed. London: SAGE publications Ltd.

Mündermann, A., Dyrby, C.O. and Andriacchi, T.P. 2005. Secondary gait changes in patients with medial compartment knee osteoarthritis: Increased load at the ankle, knee, and hip during walking. *Arthritis & Rheumatism* 52(9), pp. 2835–2844. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/art.21262 [Accessed: 16 November 2022].

Murakami, K. et al. 2018. Knee kinematics in bi-cruciate stabilized total knee arthroplasty during squatting and stair-climbing activities. *Journal of Orthopaedics* 15(2), pp. 650–654. doi: 10.1016/J.JOR.2018.05.003.

Murphy, L. et al. 2008. Lifetime risk of symptomatic knee osteoarthritis. *Arthritis & Rheumatism* 59(9), pp. 1207–1213. doi: 10.1002/art.24021.

Myers, A.M., Holliday, P.J., Harvey, K.A. and Hutchinson, K.S. 1993. Functional Performance Measures: Are They Superior to Self-Assessments? *Journal of Gerontology* 48(5), pp. M196–M206. doi: 10.1093/geronj/48.5.M196.

Naili, J.E., Broström, E.W., Clausen, B. and Holsgaard-Larsen, A. 2019a. Measures of knee and gait function and radiographic severity of knee osteoarthritis – A cross-sectional study. *Gait and Posture* 74, pp. 20–26. doi: 10.1016/j.gaitpost.2019.08.003.

Naili, J.E., Esbjörnsson, A.C., Iversen, M.D., Schwartz, M.H., Hedström, M., Häger, C.K. and Broström, E.W. 2017a. The impact of symptomatic knee osteoarthritis on overall gait pattern deviations and its association with performance-based measures and patient-reported outcomes. *Knee* 24(3), pp. 536–546. Available at: https://www.sciencedirect.com/science/article/pii/S0968016017300418 [Accessed: 13 May 2019].

Naili, J.E., Hedström, M. and Broström, E.W. 2019b. Changes of and interrelationships between performance-based function and gait and patient-reported function 1 year after total hip arthroplasty. *Journal of orthopaedics and traumatology : official journal of the Italian Society of Orthopaedics and Traumatology* 20(1), p. 14.

Available at: https://jorthoptraumatol.springeropen.com/articles/10.1186/s10195-019-0521-7 [Accessed: 2 November 2022].

Naili, J.E., Iversen, M.D., Esbjörnsson, A.C., Hedström, M., Schwartz, M.H., Häger, C.K. and Broström, E.W. 2017b. Deficits in functional performance and gait one year after total knee arthroplasty despite improved self-reported function. *Knee Surg Sports Traumatol Arthrosc* 25(11), pp. 3378–3386. Available at: https://link-springer-com.abc.cardiff.ac.uk/article/10.1007/s00167-016-4234-7 [Accessed: 7 July 2020].

Naili, J.E., Lindgren, V., Iversen, M.D., Hedström, M. and Broström, E.W. 2017c. Improved knee biomechanics among patients reporting a good outcome in knee-related quality of life one year after total knee arthroplasty. *BMC Musculoskeletal Disorders* 1(18), pp. 1–11. doi: 10.1186/s12891-017-1479-3.

National Clinical Guideline Centre. 2014. *Osteoarthritis*. National Institute for Health and Care Excellence (UK).

National Joint Registry. 2019. *16th Annual Report 2019 - National Joint Registry for England, Wales, Northern Ireland and the Isle of Man*. Hemel Hempstead.

Nayebi, H. 2020. *Advanced Statistics for Testing Assumed Casual Relationships*. Cham: Springer International Publishing. Available at: http://link.springer.com/10.1007/978-3-030-54754-7 [Accessed: 10 January 2023].

Naylor, J.M., Hayen, A., Davidson, E., Hackett, D., Harris, I.A., Kamalasena, G. and Mittal, R. 2014. Minimal detectable change for mobility and patient-reported tools in people with osteoarthritis awaiting arthroplasty. *BMC Musculoskeletal Disorders* 15(1), pp. 1–9. Available at: https://bmcmusculoskeletdisord.biomedcentral.com/articles/10.1186/1471-2474-15-235 [Accessed: 14 May 2023].

Nemes, S. et al. 2015. Historical view and future demand for knee arthroplasty in Sweden. *Acta Orthopaedica* 86(4), pp. 426–431. Available at: https://www.tandfonline.com/doi/abs/10.3109/17453674.2015.1034608 [Accessed: 11 January 2024].

Neogi, T. and Zhang, Y. 2013. Epidemiology of Osteoarthritis. *Rheumatic Disease Clinics* 39(1), pp. 1–19. Available at:

http://www.rheumatic.theclinics.com/article/S0889857X12001160/fulltext [Accessed: 15 May 2023].

NHS. 2019. *What is the body mass index (BMI)?* Available at: https://www.nhs.uk/common-health-questions/lifestyle/what-is-the-body-mass-index-bmi/ [Accessed: 13 May 2019].

NHS Digital. 2017. *A Guide to PROMS Methodology (updated March 2017)*.

NHS Digital. 2018a. *Finalised Patient Reported Outcome Measures (PROMs) in England*.

NHS Digital. 2018b. *PROMs Special Topic, Satisfaction with and Success of surgery- 2010-2011*. Available at: https://webarchive.nationalarchives.gov.uk/20180307190418/http://digital.nhs.uk/catalogue/PUB10037 [Accessed: 12 May 2019].

NHS Digital. 2019. *Patient Reported Outcome Measures (PROMs) - NHS Digital*. Available at: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/patient-reported-outcome-measures-proms [Accessed: 12 May 2019].

NICE. 2014. *Osteoarthritis: care and management. Clinical guideline.* Available at: https://www.nice.org.uk/guidance/cg177 [Accessed: 13 May 2019].

NICE. 2020a. *Joint replacement (primary): hip, knee and shoulder: Evidence review for inpatient hip and knee postoperative rehabilitation. NICE guideline*.

NICE. 2020b. *Joint replacement (primary): hip, knee and shoulder [L] Evidence review for patella resurfacing*.

NICE. 2022. *Osteoarthritis in over 16s: diagnosis and management. NICE guideline [NG226]*. Available at: https://www.nice.org.uk/guidance/ng226 [Accessed: 15 April 2023].

NICE. 2023. *Glossary*. Available at: https://www.nice.org.uk/glossary?letter=q [Accessed: 5 November 2022].

Niinimäki, T.T., Eskelinen, A., Mann, B.S., Junnila, M., Ohtonen, P. and Leppilahti, J. 2012. Survivorship of high tibial osteotomy in the treatment of osteoarthritis of the knee: Finnish registry-based study of 3195 knees. *Journal of Bone and Joint Surgery*

- *Series B* 94 B(11), pp. 1517–1521. doi: 10.1302/0301-620X.94B11.29601/LETTERTOEDITOR.

Nishio, Y., Onodera, T., Kasahara, Y., Takahashi, D., Iwasaki, N. and Majima, T. 2014. Intraoperative Medial Pivot Affects Deep Knee Flexion Angle and Patient-Reported Outcomes After Total Knee Arthroplasty. *The Journal of Arthroplasty* 29(4), pp. 702–706. doi: 10.1016/J.ARTH.2013.06.035.

NJR. 2018a. *Bearing type for unicondylar implant used in primary knee replacement*. Available at: http://www.njrreports.org.uk/knees-all-procedures-activity/K05v1NJR?reportid=A2FAF203-E834-47B8-A372-BFBA45C635CF&defaults=DC__Reporting_Period__Date_Range=%22MAX%22,JYS__Filter__Calendar_Year__From__To=%22MIN-MAX%22,H__Filter__Joint=%22Knee%22 [Accessed: 13 May 2019].

NJR. 2018b. *Types of primary knee replacements undertaken*. Available at: http://www.njrreports.org.uk/knees-all-procedures-activity/K03v1NJR?reportid=FFCEA144-54BC-486D-81A6-C6A58DDCA079&defaults=DC__Reporting_Period__Date_Range=%22MAX%22,H__JYS__Filter__Calendar_Year__From__To=%22MIN-MAX%22,R__Filter__Country=%22All%22,H__Fil [Accessed: 13 May 2019].

Nüesch, C., Laffer, D., Netzer, C., Pagenstert, G. and Mündermann, A. 2016. Effect of gait retraining for reducing ambulatory knee load on trunk biomechanics and trunk muscle activity. *Gait and Posture* 47, pp. 24–30. doi: 10.1016/J.GAITPOST.2016.03.019.

OARSI. 2023. *Physical Performance Measures | Osteoarthritis Research Society International (OARSI)*. Available at: https://oarsi.org/research-oarsi-success/physical-performance-measures [Accessed: 28 October 2018].

Øiestad, B.E., Engebretsen, L., Storheim, K. and Risberg, M.A. 2009. Winner of the 2008 Systematic Review Competition: Knee Osteoarthritis after Anterior Cruciate Ligament Injury. *The American Journal of Sports Medicine* 37(7), pp. 1434–1443. doi: 10.1177/0363546509338827.

Okamoto, N., Nakamura, E., Nishioka, H., Karasugi, T., Okada, T. and Mizuta, H. 2014. In Vivo Kinematic Comparison Between Mobile-Bearing and Fixed-Bearing

Total Knee Arthroplasty During Step-Up Activity. *The Journal of Arthroplasty* 29(12), pp. 2393–2396. Available at: http://www.sciencedirect.com/science/article/pii/S0883540314001144 [Accessed: 7 July 2022].

O'Neill, T., Jinks, C. and Ong, B.N. 2007. Decision-making regarding total knee replacement surgery: A qualitative meta-synthesis. *BMC Health Services Research* 7(1), pp. 1–9. Available at: https://link.springer.com/articles/10.1186/1472-6963-7-52 [Accessed: 4 March 2021].

Onodera, C.M.K., Coelho-Júnior, H.J., Sampaio, R.A.C., Santos Duarte Lana, J.F., Teixeira, L.F.M., Uchida, M.C. and Bizzacchi, J.M.A. 2020. The importance of objectively measuring functional tests in complement to self-report assessments in patients with knee osteoarthritis. *Gait & Posture* 82, pp. 33–37. doi: 10.1016/J.GAITPOST.2020.08.121.

Van Onsem, S., Verstraete, M., Van Eenoo, W., Van Der Straeten, C. and Victor, J. 2020. Are TKA Kinematics During Closed Kinetic Chain Exercises Associated with Patient-reported Outcomes? A Preliminary Analysis. *Clinical Orthopaedics and Related Research* 478(2), p. 255. Available at: /pmc/articles/PMC7438127/ [Accessed: 19 June 2023].

Ornetti, P., Maillefert, J.F., Laroche, D., Morisset, C., Dougados, M. and Gossec, L. 2010. Gait analysis as a quantifiable outcome measure in hip or knee osteoarthritis: A systematic review. *Joint Bone Spine* 77(5), pp. 421–425. doi: 10.1016/j.jbspin.2009.12.009.

Osteoarthritis Research Society International. 2016. *OARSI White Paper-Osteoarthritis as a Serious Disease*. Available at: https://www.oarsi.org/education/oarsi-resources/oarsi-white-paper-oa-serious-disease [Accessed: 3 April 2019].

Ouellet, D. and Moffet, H. 2002. Locomotor deficits before and two months after knee arthroplasty. *Arthritis & Rheumatism* 47(5), pp. 484–493. Available at: http://doi.wiley.com/10.1002/art.10652 [Accessed: 20 July 2020].

Paschos, N.K. 2017. Anterior cruciate ligament reconstruction and knee osteoarthritis. *World journal of orthopedics* 8(3), pp. 212–217. doi: 10.5312/wjo.v8.i3.212.

Paterson, K.L. et al. 2018. Effects of sex and obesity on gait biomechanics before and six months after total knee arthroplasty: A longitudinal cohort study. *Gait and Posture* 61, pp. 263–268. doi: 10.1016/j.gaitpost.2018.01.014.

Paterson, K.L. et al. 2020. The influence of sex and pre-operative obesity on biomechanics two years after total knee arthroplasty: A longitudinal cohort study. *Gait and Posture* 76, pp. 74–84. doi: 10.1016/j.gaitpost.2019.10.031.

Peter, W.F., Poolman, R.W., Scholtes, V.A.B., de Vet, H.C.W. and Terwee, C.B. 2019. Responsiveness and interpretability of the Animated Activity Questionnaire for assessing activity limitations of patients with hip or knee osteoarthritis. *Musculoskeletal Care* 17(4), pp. 327–334. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/msc.1418 [Accessed: 1 November 2022].

Pol, F., Baharlouei, H., Taheri, A., Menz, H.B. and Forghany, S. 2021a. Foot and ankle biomechanics during walking in older adults: A systematic review and meta-analysis of observational studies. *Gait & Posture* 89, pp. 14–24. doi: 10.1016/J.GAITPOST.2021.06.018.

Pol, F., Baharlouei, H., Taheri, A., Menz, H.B. and Forghany, S. 2021b. Foot and ankle biomechanics during walking in older adults: A systematic review and meta-analysis of observational studies. *Gait & Posture* 89, pp. 14–24. doi: 10.1016/J.GAITPOST.2021.06.018.

Portney, L.G. and Watkins, M.P. 2014. *Foundations of Clinical Research: Application to Practice*. Third Edit. Harlow: Pearson.

Preece, S.J., Algarni, A.S. and Jones, R.K. 2019. Trunk flexion during walking in people with knee osteoarthritis. *Gait & Posture* 72, pp. 202–205. doi: 10.1016/J.GAITPOST.2019.06.012.

Preece, S.J. and Alghamdi, W. 2021. Increased trunk flexion may underlie elevated knee flexor activity in people with knee osteoarthritis. *Knee* 33, pp. 216–225. doi: 10.1016/j.knee.2021.10.006.

Price, A.J. et al. 2018. Knee replacement. *The Lancet* 392(10158), pp. 1672–1682. doi: 10.1016/S0140-6736(18)32344-4.

Price, A.J. et al. 2020. The use of patient-reported outcome measures to guide referral for hip and knee arthroplasty: Part 1: The development of an evidence-based model linking preoperative score to the probability of gaining benefit from surgery. *Bone and Joint Journal* 102-B(7), pp. 941–949. doi: 10.1302/0301-620X.102B7.BJJ-2019-0102.R2/LETTERTOEDITOR.

Prieto-Alhambra, D., Judge, A., Javaid, K., Cooper, C., Diez-Perez, A. and Arden, N.K. 2013. Incidence and risk factors for clinically diagnosed knee, hip and hand osteoarthritis: influences of age, gender and osteoarthritis affecting other joints. *Annals of the Rheumatic Diseases* 73(9), pp. 1659–1664. Available at: http://ard.bmj.com/ [Accessed: 1 December 2022].

Prieto-Alhambra, D., Judge, A., Javaid, M.K., Cooper, C., Diez-Perez, A. and Arden, N.K. 2014. Incidence and risk factors for clinically diagnosed knee, hip and hand osteoarthritis: influences of age, gender and osteoarthritis affecting other joints. *Annals of the rheumatic diseases* 73(9), pp. 1659–64. doi: 10.1136/annrheumdis-2013-203355.

Primeau, C.A. et al. 2021. Total knee replacement after high tibial osteotomy: Time-to-event analysis and predictors. *CMAJ* 193(5), pp. E158–E166. Available at: /pmc/articles/PMC7954572/ [Accessed: 1 May 2023].

Pua, Y.H. et al. 2019. Predicting individual knee range of motion, knee pain, and walking limitation outcomes following total knee arthroplasty. *Acta Orthopaedica* 90(2), pp. 179–186. Available at: /pmc/articles/PMC6461070/?report=abstract [Accessed: 8 January 2021].

Pua, Y.H., Ong, P.H., Chong, H.C., Yeo, W., Tan, C.I.C. and Lo, N.N. 2013. Associations of self-report physical function with knee strength and knee range-of-motion in total knee arthroplasty possible nonlinear and threshold effects. *Journal of Arthroplasty* 28(9), pp. 1521–1527. doi: 10.1016/j.arth.2012.10.020.

Raja, K. and Dewan, N. 2011. Efficacy of Knee Braces and Foot Orthoses in Conservative Management of Knee Osteoarthritis. *American Journal of Physical*

Medicine & Rehabilitation 90(3), pp. 247–262. doi: 10.1097/PHM.0b013e318206386b.

Rehman, Y., Koster, L.A., Röhrl, S.M. and Aamodt, A. 2022. Comparison of the in-vivo kinematics of three different knee prosthesis designs during a step-up movement. *Clinical Biomechanics* 100, p. 105824. doi: 10.1016/J.CLINBIOMECH.2022.105824.

Resende, R.A., Fonseca, S.T., Silva, P.L., Magalhães, C.M.B. and Kirkwood, R.N. 2012a. Power at hip, knee and ankle joints are compromised in women with mild and moderate knee osteoarthritis. *Clinical Biomechanics* 27(10), pp. 1038–1044. doi: 10.1016/J.CLINBIOMECH.2012.08.001.

Resende, R.A., Fonseca, S.T., Silva, P.L., Magalhães, C.M.B. and Kirkwood, R.N. 2012b. Power at hip, knee and ankle joints are compromised in women with mild and moderate knee osteoarthritis. *Clinical Biomechanics* 27(10), pp. 1038–1044. doi: 10.1016/J.CLINBIOMECH.2012.08.001.

Riddle, D.L., Macfarlane, G.J., Hamilton, D.F., Beasley, M. and Dumenci, L. 2022. Cross-validation of good versus poor self-reported outcome trajectory types following knee arthroplasty. *Osteoarthritis and Cartilage* 30(1), pp. 61–68. doi: 10.1016/J.JOCA.2021.09.004.

Ro, D.H., Han, H.S., Lee, D.Y., Kim, S.H., Kwak, Y.H. and Lee, M.C. 2018. Slow gait speed after bilateral total knee arthroplasty is associated with suboptimal improvement of knee biomechanics. *Knee Surgery, Sports Traumatology, Arthroscopy* 26(6), pp. 1671–1680. Available at: https://link.springer.com/article/10.1007/s00167-017-4682-8 [Accessed: 7 November 2022].

Ro, D.H., Kang, T., Han, D. hwan, Lee, D.Y., Han, H.S. and Lee, M.C. 2020. Quantitative evaluation of gait features after total knee arthroplasty: Comparison with age and sex-matched controls. *Gait & Posture* 75, pp. 78–84. doi: 10.1016/J.GAITPOST.2019.09.026.

Ro, D.H., Lee, J., Lee, J., Park, J.Y., Han, H.S. and Lee, M.C. 2019. Effects of Knee Osteoarthritis on Hip and Ankle Gait Mechanics. *Advances in Orthopedics* 2019. doi: 10.1155/2019/9757369.

Robertson, G.E., Caldwell, G.E., Hamill, J., Kamen, G. and Whittlesey, S.N. 2013. *Research Methods in Biomechanics*. 2nd ed. Champaign: Human Kinetics.

Roos, E.M., Herzog, W., Block, J.A. and Bennell, K.L. 2011. Muscle weakness, afferent sensory dysfunction and exercise in knee osteoarthritis. *Nature Reviews Rheumatology* 7(1), pp. 57–63. doi: 10.1038/nrrheum.2010.195.

Roos, E.M., Roos, H.P., Lohmander, L.S., Ekdahl, C. and Beynnon, B.D. 1998. Knee Injury and Osteoarthritis Outcome Score (KOOS)—Development of a Self-Administered Outcome Measure. *Journal of Orthopaedic & Sports Physical Therapy* 28(2), pp. 88–96. Available at: http://www.jospt.org/doi/10.2519/jospt.1998.28.2.88 [Accessed: 12 May 2019].

Rosenlund, S., Broeng, L., Overgaard, S., Jensen, C. and Holsgaard-Larsen, A. 2016a. The efficacy of modified direct lateral versus posterior approach on gait function and hip muscle strength after primary total hip arthroplasty at 12 months follow-up. An explorative randomised controlled trial. *Clinical Biomechanics* 39, pp. 91–99. doi: 10.1016/j.clinbiomech.2016.09.011.

Rosenlund, S., Holsgaard-Larsen, A., Overgaard, S. and Jensen, C. 2016b. The Gait Deviation Index is associated with hip muscle strength and patient-reported outcome in patients with severe hip osteoarthritis - A cross-sectional study. *PLoS ONE* 11(4), p. e0153177. Available at: www.bevica.dk/bevica-fonden/formalansogning [Accessed: 11 May 2021].

Rowe, E., Beauchamp, M.K. and Astephen Wilson, J. 2021. Age and sex differences in normative gait patterns. *Gait and Posture* 88, pp. 109–115. doi: 10.1016/J.GAITPOST.2021.05.014.

Rowe, P.J., Myles, C.M., Walker, C. and Nutton, R. 2000. Knee joint kinematics in gait and other functional activities measured using flexible electrogoniometry: how much knee motion is sufficient for normal daily life? *Gait & Posture* 12(2), pp. 143–155. doi: 10.1016/S0966-6362(00)00060-6.

Rozumalski, A. and Schwartz, M.H. 2011. The GDI-Kinetic: A new index for quantifying kinetic deviations from normal gait. *Gait & Posture* 33(4), pp. 730–732. doi: 10.1016/J.GAITPOST.2011.02.014.

Saari, T., Carlsson, L., Karlsson, J. and Kärrholm, J. 2005a. Knee kinematics in medial arthrosis. Dynamic radiostereometry during active extension and weight-bearing. *Journal of Biomechanics* 38(2), pp. 285–292. doi: 10.1016/J.JBIOMECH.2004.02.009.

Saari, T., Tranberg, R., Zügner, R., Uvehammer, J. and Kärrholm, J. 2005b. Changed gait pattern in patients with total knee arthroplasty but minimal influence of tibial insert design: Gait analysis during level walking in 39 TKR patients and 18 healthy controls. *Acta Orthopaedica* 76(2), pp. 253–260. Available at: https://www.tandfonline.com/action/journalInformation?journalCode=iort20 [Accessed: 12 May 2021].

Sabah, S.A., Alvand, A., Beard, D.J. and Price, A.J. 2022. Minimal important changes and differences were estimated for Oxford hip and knee scores following primary and revision arthroplasty. *Journal of Clinical Epidemiology* 143, pp. 159–168. Available at: http://www.jclinepi.com/article/S0895435621004212/fulltext [Accessed: 15 June 2022].

Safranek, R.J., Gottschlich, S. and Kak, A.C. 1990. Evidence Accumulation Using Binary Frames of Discernment for Verification Vision. *IEEE Transactions on Robotics and Automation* 6(4), pp. 405–417. doi: 10.1109/70.59366.

Sagawa, Y., Armand, S., Lubbeke, A., Hoffmeyer, P., Fritschy, D., Suva, D. and Turcot, K. 2013. Associations between gait and clinical parameters in patients with severe knee osteoarthritis: A multiple correspondence analysis. *Clinical Biomechanics* 28(3), pp. 299–305. doi: 10.1016/J.CLINBIOMECH.2013.01.008.

Sanchez-Santos, M.T. et al. 2018. Development and validation of a clinical prediction model for patient-reported pain and function after primary total knee replacement surgery. *Scientific Reports* 8(1), p. 3381. Available at: /pmc/articles/PMC5821875/?report=abstract [Accessed: 6 January 2021].

Scarvell, J.M., Galvin, C.R., Perriman, D.M., Lynch, J.T. and van Deursen, R.W.M. 2018. Kinematics of knees with osteoarthritis show reduced lateral femoral roll-back and maintain an adducted position. A systematic review of research using medical imaging. *Journal of Biomechanics* 75, pp. 108–122. doi: 10.1016/J.JBIOMECH.2018.05.007.

Scarvell, J.M., Pickering, M.R. and Smith, P.N. 2010. New registration algorithm for determining 3D knee kinematics using CT and single-plane fluoroscopy with improved out-of-plane translation accuracy. *Journal of Orthopaedic Research* 28(3), pp. 334–340. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/jor.21003 [Accessed: 7 December 2022].

Scarvell, J.M., Smith, P.N., Refshauge, K.M. and Galloway, H.R. 2007. Magnetic Resonance Imaging Analysis of Kinematics in Osteoarthritic Knees. *The Journal of Arthroplasty* 22(3), pp. 383–393. doi: 10.1016/J.ARTH.2006.06.006.

Schache, M.B., McClelland, J.A. and Webster, K.E. 2019. Incorporating hip abductor strengthening exercises into a rehabilitation program did not improve outcomes in people following total knee arthroplasty: a randomised trial. *Journal of Physiotherapy* 65(3), pp. 136–143. doi: 10.1016/J.JPHYS.2019.05.008.

Schmid, S., Studer, D., Hasler, C.C., Romkes, J., Taylor, W.R., Lorenzetti, S. and Brunner, R. 2016. Quantifying spinal gait kinematics using an enhanced optical motion capture approach in adolescent idiopathic scoliosis. *Gait & Posture* 44, pp. 231–237. doi: 10.1016/J.GAITPOST.2015.12.036.

Schwartz, M.H. and Rozumalski, A. 2008. The gait deviation index: A new comprehensive index of gait pathology. *Gait and Posture* 28(3), pp. 351–357. doi: 10.1016/j.gaitpost.2008.05.001.

Scott, C.E.H., Howie, C.R., Macdonald, D. and Biant, L.C. 2010. Predicting dissatisfaction following total knee replacement. A prospective study of 1217 patients. *J Bone Joint Surg [Br]* 92(9), pp. 92–1253. doi: 10.1302/0301-620X.92B9.

Sedgwick, P. 2014. Multiple hypothesis testing and Bonferroni's correction. *BMJ* 349. Available at: https://www.bmj.com/content/349/bmj.g6284 [Accessed: 4 July 2023].

Senden, R., Grimm, B., Meijer, K., Savelberg, H. and Heyligers, I.C. 2011. The importance to including objective functional outcomes in the clinical follow up of total knee arthroplasty patients. *The Knee* 18(5), pp. 306–11. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20817472 [Accessed: 13 May 2019].

Serna-Berna, R., Lizaur-Utrilla, A., Vizcaya-Moreno, M.F., Miralles Muñoz, F.A., Gonzalez-Navarro, B. and Lopez-Prats, F.A. 2018. Cruciate-Retaining vs Posterior-Stabilized Primary Total Arthroplasty. Clinical Outcome Comparison With a Minimum

Follow-Up of 10 Years. *The Journal of Arthroplasty* 33(8), pp. 2491–2495. doi: 10.1016/J.ARTH.2018.02.094.

Shafer, G. 1976. *A mathematical theory of evidence*. Princeton university press.

Shakoor, N. and Block, J.A. 2006. Walking barefoot decreases loading on the lower extremity joints in knee osteoarthritis. *Arthritis & Rheumatism* 54(9), pp. 2923–2927. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/art.22123 [Accessed: 24 November 2022].

Sharma, L., Song, J., Felson, D.T., Cahue, S., Shamiyeh, E. and Dunlop, D.D. 2001. The Role of Knee Alignment in Disease Progression and Functional Decline in Knee Osteoarthritis. *JAMA* 286(2), p. 188. doi: 10.1001/jama.286.2.188.

Sharma, S. et al. 2023. Reliability, Validity, Responsiveness, and Minimum Important Change of the Stair Climb Test in Adults With Hip and Knee Osteoarthritis. *Arthritis Care & Research* 75(5), pp. 1147–1157. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/acr.24821 [Accessed: 14 May 2023].

Shim, J. and Hamilton, D.F. 2019. Comparative responsiveness of the PROMIS-10 global health and EQ-5D questionnaires in patients undergoing total knee arthroplasty. *Bone and Joint Journal* 101 B(7), pp. 832–837. Available at: https://online.boneandjoint.org.uk/doi/abs/10.1302/0301-620X.101B7.BJJ-2018-1543.R1 [Accessed: 1 November 2022].

Silder, A., Heiderscheit, B. and Thelen, D.G. 2008. Active and passive contributions to joint kinetics during walking in older adults. *Journal of Biomechanics* 41(7), pp. 1520–1527. doi: 10.1016/J.JBIOMECH.2008.02.016.

Simic, M., Hunt, M.A., Bennell, K.L., Hinman, R.S. and Wrigley, T. V. 2012. Trunk lean gait modification and knee joint load in people with medial knee osteoarthritis: The effect of varying trunk lean angles. *Arthritis Care & Research* 64(10), pp. 1545–1553. Available at: https://doi.org/10.1002/acr.21724.

Sims, E.L., Carland, J.M., Keefe, F.J., Kraus, V.B., Guilak, F. and Schmitt, D. 2009. Sex Differences in Biomechanics Associated with Knee Osteoarthritis. *Journal of Women & Aging* 21(3), pp. 159–170. Available at: https://www.tandfonline.com/action/journalInformation?journalCode=wjwa20 [Accessed: 12 July 2019].

Sinclair, J., Taylor, P.J. and Hobbs, S.J. 2013. Digital Filtering of Three-Dimensional Lower Extremity Kinematics: an Assessment. *Journal of Human Kinetics* 39, pp. 25–36.

Singh, J.A., Dowsey, M.M., Dohm, M., Goodman, S.M., Leong, A.L., Voshaar, M.M.J.H.S. and Choong, P.F. 2017. Achieving Consensus on Total Joint Replacement Trial Outcome Reporting Using the OMERACT Filter: Endorsement of the Final Core Domain Set for Total Hip and Total Knee Replacement Trials for Endstage Arthritis. *The Journal of Rheumatology* 44(11), pp. 1723–1726. Available at: https://www.jrheum.org/content/44/11/1723 [Accessed: 20 April 2023].

Singh, J.A., Yu, S., Chen, L. and Cleveland, J.D. 2019. Rates of Total Joint Replacement in the United States: Future Projections to 2020-2040 Using the National Inpatient Sample. *The Journal of Rheumatology* 46(9), pp. 1134–1140. Available at: https://www.scinapse.io/papers/2938070949 [Accessed: 16 April 2023].

Smith, T.O. et al. 2019. Uptake of the OMERACT-OARSI Hip and Knee Osteoarthritis Core Outcome Set: Review of Randomized Controlled Trials from 1997 to 2017. *The Journal of Rheumatology* 46(8), pp. 976–980. Available at: https://www-jrheum-org.abc.cardiff.ac.uk/content/46/8/976 [Accessed: 10 June 2023].

Standifird, T.W., Saxton, A.M., Coe, D.P., Cates, H.E., Reinbolt, J.A. and Zhang, S. 2016. Influence of Total Knee Arthroplasty on Gait Mechanics of the Replaced and Non-Replaced Limb During Stair Negotiation. *Journal of Arthroplasty* 31(1), pp. 278–283. doi: 10.1016/j.arth.2015.06.052.

Stanhope, S.J., Kepple, T.M., McGuire, D.A. and Roman, N.L. 1990. Kinematic-based technique for event time determination during gait. *Medical & Biological Engineering & Computing* 28(4), pp. 355–360. Available at: https://link.springer.com/article/10.1007/BF02446154 [Accessed: 3 July 2023].

Stevens-Lapsley, J.E., Petterson, S.C., Mizner, R.L. and Snyder-Mackler, L. 2010. Impact of Body Mass Index on Functional Performance After Total Knee Arthroplasty. *The Journal of Arthroplasty* 25(7), pp. 1104–1109. doi: 10.1016/J.ARTH.2009.08.009.

Stevens-Lapsley, J.E., Schenkman, M.L. and Dayton, M.R. 2011a. Comparison of Self-Reported Knee Injury and Osteoarthritis Outcome Score to Performance Measures in Patients After Total Knee Arthroplasty. *PM&R* 3(6), pp. 541–549. Available at: www.pmrjournal.org [Accessed: 12 May 2019].

Stevens-Lapsley, J.E., Schenkman, M.L. and Dayton, M.R. 2011b. Comparison of Self-Reported Knee Injury and Osteoarthritis Outcome Score to Performance Measures in Patients After Total Knee Arthroplasty. *PM&R* 3(6), pp. 541–549. doi: 10.1016/J.PMRJ.2011.03.002.

Stickles, B., Phillips, L., Brox, W.T., Owens, B. and Lanzer, W.L. 2001. Defining the Relationship between Obesity and Total Joint Arthroplasty. *Obesity Research* 9(3), pp. 219–223. Available at: https://onlinelibrary.wiley.com/doi/full/10.1038/oby.2001.24 [Accessed: 2 March 2023].

Stratford, P.W., Kennedy, D.M., Maly, M.R. and MacIntyre, N.J. 2010. Quantifying Self-Report Measures' Overestimation of Mobility Scores Postarthroplasty. *Physical Therapy* 90(9), pp. 1288–1296. Available at: http://academic.oup.com/ptj/article/90/9/1288/2738048 [Accessed: 12 May 2019].

Stürmer, T., Günther, K.-P. and Brenner, H. 2000. Obesity, overweight and patterns of osteoarthritis: the Ulm Osteoarthritis Study. *Journal of Clinical Epidemiology* 53(3), pp. 307–313. doi: 10.1016/S0895-4356(99)00162-6.

Suri, P., Morgenroth, D.C. and Hunter, D.J. 2012. Epidemiology of Osteoarthritis and Associated Comorbidities. *Physical Medicine and Rehabilitation* 4, pp. S10–S19. doi: 10.1016/j.pmrj.2012.01.007.

Terwee, C.B. et al. 2007. Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology* 60(1), pp. 34–42. doi: 10.1016/J.JCLINEPI.2006.03.012.

Terwee, C.B., Mokkink, L.B., Steultjens, M.P.M. and Dekker, J. 2006. Performance-based methods for measuring the physical function of patients with osteoarthritis of the hip or knee: a systematic review of measurement properties. *Rheumatology* 45(7), pp. 890–902. Available at:

http://academic.oup.com/rheumatology/article/45/7/890/1788526/Performancebased-methods-for-measuring-the [Accessed: 12 May 2019].

Thorp, L.E., Sumner, D.R., Wimmer, M.A. and Block, J.A. 2007. Relationship between pain and medial knee joint loading in mild radiographic knee osteoarthritis. *Arthritis Care & Research* 57(7), pp. 1254–1260. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/art.22991 [Accessed: 28 May 2023].

Tibesku, C.O., Daniilidis, K., Vieth, V., Skwara, A., Heindel, W. and Fuchs-Winkelmann, S. 2011. Sagittal plane kinematics of fixed- and mobile-bearing total knee replacements. *Knee Surgery, Sports Traumatology, Arthroscopy* 19(9), pp. 1488–1495. Available at: https://link.springer.com/article/10.1007/s00167-011-1462-8 [Accessed: 7 July 2022].

Toda, H., Maruyama, T., Fujita, K., Yamauchi, Y. and Tada, M. 2021. Self-Reported Walking Difficulty Associated with Stiff-Knee Gait in Japanese Patients with Knee Osteoarthritis: A Preliminary Cross-Sectional Study. *Healthcare 2021, Vol. 9, Page 1308* 9(10), p. 1308. Available at: https://www.mdpi.com/2227-9032/9/10/1308/htm [Accessed: 16 January 2024].

Tolk, J.J., Janssen, R.P.A., Prinsen, C.A.C., Latijnhouwers, D.A.J.M., van der Steen, M.C., Bierma-Zeinstra, S.M.A. and Reijman, M. 2017. The OARSI core set of performance-based measures for knee osteoarthritis is reliable but not valid and responsive. *Knee Surg Sports Traumatol Arthrosc* 27(9), pp. 2898–2909. Available at: http://link.springer.com/10.1007/s00167-017-4789-y [Accessed: 12 May 2019].

Tsai, T.Y., Lu, T.W., Chen, C.M., Kuo, M.Y. and Hsu, H.C. 2010. A volumetric model-based 2D to 3D registration method for measuring kinematics of natural knees with single-plane fluoroscopy. *Medical Physics* 37(3), pp. 1273–1284. Available at: https://onlinelibrary.wiley.com/doi/full/10.1118/1.3301596 [Accessed: 26 June 2023].

Turcot, K., Armand, S., Lübbeke, A., Fritschy, D., Hoffmeyer, P. and Suvà, D. 2013. Does knee alignment influence gait in patients with severe knee osteoarthritis? *Clinical Biomechanics* 28(1), pp. 34–39. doi: 10.1016/J.CLINBIOMECH.2012.09.004.

Twiggs, J.G. et al. 2019. Clinical and Statistical Validation of a Probabilistic Prediction Tool of Total Knee Arthroplasty Outcome. *The Journal of Arthroplasty* 34(11), pp. 2624–2631. doi: 10.1016/J.ARTH.2019.06.007.

UK Health Security Agency. 2011. *Ionising radiation: dose comparisons*. Available at: https://www.gov.uk/government/publications/ionising-radiation-dose-comparisons/ionising-radiation-dose-comparisons#:~:text=The%202.7%20mSv%20dose%20that,radiation%20in%20homes%20and%20workplaces. [Accessed: 18 June 2022].

Vaienti, E., Scita, G., Ceccarelli, F. and Pogliacomi, F. 2017. Understanding the human knee and its relationship to total knee replacement. *Acta Bio Medica : Atenei Parmensis* 88(Suppl 2), p. 6. Available at: /pmc/articles/PMC6178997/ [Accessed: 11 January 2024].

Vajapey, S.P., Morris, J., Spitzer, A.I., Glassman, A.H., Greco, N.J. and Li, M. 2020. Outcome reporting patterns in total knee arthroplasty: A systematic review. *Journal of Clinical Orthopaedics and Trauma* 11, pp. S464–S471. doi: 10.1016/J.JCOT.2020.05.014.

Varadarajan, K.M., Harry, R.E., Johnson, T. and Li, G. 2009. Can in vitro systems capture the characteristic differences between the flexion–extension kinematics of the healthy and TKA knee? *Medical Engineering & Physics* 31(8), pp. 899–906. doi: 10.1016/J.MEDENGPHY.2009.06.005.

De Vecchis, M., Holt, C.A., Holsgaard-Larsen, A., Whatling, G.M., Broström, E.W., Wilson, C. and Naili, J.E. 2022. RELATIONSHIP BETWEEN GAIT CLASSIFICATION METHODS IN TWO COHORTS OF PATIENTS WITH SEVERE KNEE OSTEOARTHRITIS WHO RECEIVED A TOTAL KNEE REPLACEMENT. In: *OARSI World Congress*.

Verra, W.C., Van Den Boom, L.G.H., Jacobs, W.C.H., Schoones, J.W., Wymenga, A.B. and Nelissen, R.G.H.H. 2015. Similar outcome after retention or sacrifice of the posterior cruciate ligament in total knee arthroplasty: A systematic review and meta-analysis. *Acta Orthopaedica* 86(2), pp. 195–201. Available at: https://www.tandfonline.com/doi/abs/10.3109/17453674.2014.973329 [Accessed: 13 April 2023].

Versus Arthritis. 2019. *The State of Musculoskeletal Health 2019*. Available at: https://www.versusarthritis.org/media/14594/state-of-musculoskeletal-health-2019.pdf [Accessed: 28 August 2019].

Vignon, É. et al. 2006. Osteoarthritis of the knee and hip and activity: a systematic international review and synthesis (OASIS). *Joint Bone Spine* 73(4), pp. 442–455. doi: 10.1016/J.JBSPIN.2006.03.001.

Vos, T. et al. 2020. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. doi: 10.1016/S0140-6736(20)30925-9.

De Vroey, H., Staes, F., Weygers, I., Bekteshi, S., Vereecke, E., Hallez, H. and Claeys, K. 2018. P 161 – Correlation between patient reported outcome measures and objective clinical findings in the knee arthroplasty population. *Gait & Posture* 65, pp. 507–508. doi: 10.1016/J.GAITPOST.2018.07.082.

Wang, L., Lin, L., Feng, Y., Fernandes, T.L., Asnis, P., Hosseini, A. and Li, G. 2015. Anterior cruciate ligament reconstruction and cartilage contact forces—A 3D computational simulation. *Clinical Biomechanics* 30(10), pp. 1175–1180. doi: 10.1016/J.CLINBIOMECH.2015.08.007.

Wang, W. et al. 2023. Effects of soft tissue artifacts on the calculated kinematics of the knee during walking and running. *Journal of Biomechanics* 150, p. 111474. doi: 10.1016/J.JBIOMECH.2023.111474.

Warth, L.C., Ishmael, M.K., Deckard, E.R., Ziemba-Davis, M. and Meneghini, R.M. 2017. Do Medial Pivot Kinematics Correlate With Patient-Reported Outcomes After Total Knee Arthroplasty? *The Journal of Arthroplasty* 32(8), pp. 2411–2416. doi: 10.1016/J.ARTH.2017.03.019.

Watling, D.J. 2013. *Development of Novel Methodologies to Quantify, Analyse and Classify In-Vivo Knee Function Affected by Aging, Osteoarthritis and Total Knee Replacement*. PhD Thesis, Cardiff University.

W-Dahl, A., Robertsson, O. and Lohmander, L.S. 2012. High tibial osteotomy in Sweden, 1998–2007. *New Pub: Medical Journals Sweden* 83(3), pp. 244–248. Available at: https://www.tandfonline.com/doi/abs/10.3109/17453674.2012.688725 [Accessed: 7 May 2023].

Whatling, G.M. 2009. *A contribution to the clinical validation of a generic method for the classification of osteoarthritic and non-pathological knee function, PhD Thesis, Cardiff University*. PhD Thesis, Cardiff University.

Whatling, G.M., Biggs, P.R., Elson, D.W., Metcalfe, A., Wilson, C. and Holt, C. 2020. High tibial osteotomy results in improved frontal plane knee moments, gait patterns and patient-reported outcomes. *Knee Surgery, Sports Traumatology, Arthroscopy* 28(9), pp. 2872–2882. Available at: https://link.springer.com/article/10.1007/s00167-019-05644-7 [Accessed: 7 November 2022].

Whatling, G.M., Biggs, P.R., Wilson, C. and Holt, C.A. 2022. Assessing functional recovery following total knee replacement surgery using objective classification of level gait data and patient-reported outcome measures. *Clinical Biomechanics* 95, p. 105625. doi: 10.1016/J.CLINBIOMECH.2022.105625.

WHO. 2010. *A healthy lifestyle - WHO recommendations*. Available at: https://www.who.int/europe/news-room/fact-sheets/item/a-healthy-lifestyle---who-recommendations [Accessed: 1 May 2022].

WHO. 2019. *Chronic rheumatic conditions*. Available at: https://www.who.int/chp/topics/rheumatic/en/ [Accessed: 18 April 2019].

WHO. 2023. *International Classification of Functioning Disability and Health - Activities and Participation: Mobility*. Available at: https://icd.who.int/dev11/l-icf/en#/http%3A%2F%2Fid.who.int%2Ficd%2Fentity%2F2048203604 [Accessed: 4 May 2022].

Wiik, A.V., Aqil, A., Brevadt, M., Jones, G. and Cobb, J. 2017. Abnormal ground reaction forces lead to a general decline in gait speed in knee osteoarthritis patients. *World Journal of Orthopedics* 8(4), p. 322. Available at: /pmc/articles/PMC5396017/ [Accessed: 16 November 2022].

Williams, D. et al. 2020. The relationship between alignment, function and loading in total knee replacement: In-vivo analysis of a unique patient population. *Journal of Biomechanics* 112, p. 110042. doi: 10.1016/J.JBIOMECH.2020.110042.

Williams, D.E. 2018. *Development and Validation of a Biplane Fluoroscopy System to Quantify In-Vivo Knee Kinematics*. Cardiff: Cardiff University.

Williams, D.P., Price, A.J., Beard, D.J., Hadfield, S.G., Arden, N.K., Murray, D.W. and Field, R.E. 2013. The effects of age on patient-reported outcome measures in total knee replacements. *J Bone Joint Surg* 95 B(1), pp. 38–44. Available at: https://online.boneandjoint.org.uk/doi/abs/10.1302/0301-620X.95B1.28061 [Accessed: 21 December 2022].

Wluka, A.E., Lombard, C.B. and Cicuttini, F.M. 2013. Tackling obesity in knee osteoarthritis. *Nature Reviews Rheumatology* 9(4), pp. 225–235. doi: 10.1038/nrrheum.2012.224.

Worsley, P.R. 2011. *Assessment of short-term knee arthroplasty function using clinical measures, motion analysis, and musculoskeletal modelling*. University of Southampton.

Worsley, P.R., Whatling, G., Barrett, D., Holt, C., Stokes, M. and Taylor, M. 2015. Assessing changes in subjective and objective function from pre- to post-knee arthroplasty using the Cardiff Dempster–Shafer theory classifier. *Computer Methods in Biomechanics and Biomedical Engineering* 19(4), pp. 418–427. Available at: https://www.tandfonline.com/doi/abs/10.1080/10255842.2015.1034115 [Accessed: 14 October 2021].

Wu, G. et al. 2002. 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(4), pp. 543–548. Available at: https://www.sciencedirect.com/science/article/pii/S0021929001002226 [Accessed: 10 October 2019].

Wu, G. et al. 2005. ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion—Part II: shoulder, elbow, wrist and hand. *Journal of Biomechanics* 38(5), pp. 981–992. doi: https://doi.org/10.1016/j.jbiomech.2004.05.042.

Wu, G. and Cavanagh, P.R. 1995. ISB recommendations for standardization in the reporting of kinematic data. *Journal of Biomechanics* 28(10), pp. 1257–1262. Available at: http://isbweb.org/images/documents/standards/Wu and Cavanagh J Biomech 28 (1995) 1258-1261.pdf [Accessed: 9 March 2023].

Wylde, V., Blom, A.W., Whitehouse, S.L., Taylor, A.H., Pattison, G.T. and Bannister, G.C. 2009. Patient-Reported Outcomes After Total Hip and Knee Arthroplasty: Comparison of Midterm Results. *The Journal of Arthroplasty* 24(2), pp. 210–216. doi: 10.1016/J.ARTH.2007.12.001.

Xie, F., Ye, H., Zhang, Y., Liu, X., Lei, T. and Li, S.C. 2011. Extension from inpatients to outpatients: validity and reliability of the Oxford Knee Score in measuring health outcomes in patients with knee osteoarthritis. *International journal of rheumatic diseases* 14(2), pp. 206–210. Available at: https://pubmed.ncbi.nlm.nih.gov/21518321/ [Accessed: 6 July 2023].

Xu, S. et al. 2018. The influence of obesity on functional outcome and quality of life after total knee arthroplasty. *The Bone & Joint Journal* 100B(5), pp. 579–583. Available at: https://online.boneandjoint.org.uk/doi/abs/10.1302/0301-620X.100B5.BJJ-2017-1263.R1 [Accessed: 6 January 2023].

Yoo, J.D., Papannagari, R., Park, S.E., DeFrate, L.E., Gill, T.J. and Li, G. 2005. The Effect of Anterior Cruciate Ligament Reconstruction on Knee Joint Kinematics under Simulated Muscle Loads. *The American Journal of Sports Medicine* 33(2), pp. 240–246. doi: 10.1177/0363546504267806.

Yoshida, Y., Mizner, R.L., Ramsey, D.K. and Snyder-Mackler, L. 2008. Examining outcomes from total knee arthroplasty and the relationship between quadriceps strength and knee function over time. *Clinical Biomechanics* 23(3), pp. 320–328. doi: 10.1016/j.clinbiomech.2007.10.008.

Yoshida, Y., Zeni, J. and Snyder-Mackler, L. 2012. Do patients achieve normal gait patterns 3 years after total knee arthroplasty? *Journal of Orthopaedic and Sports Physical Therapy* 42(12), pp. 1039–1049. doi: 10.2519/jospt.2012.3763.

Yu, D., Peat, G., Bedson, J. and Jordan, K.P. 2015. Annual consultation incidence of osteoarthritis estimated from population-based health care data in England. *Rheumatology* 54(11), pp. 2051–2060. doi: 10.1093/rheumatology/kev231.

Yue, B., Varadarajan, K.M., Moynihan, A.L., Liu, F., Rubash, H.E. and Li, G. 2011. Kinematics of medial osteoarthritic knees before and after posterior cruciate ligament retaining total knee arthroplasty. *Journal of Orthopaedic Research* 29(1), pp. 40–46.

Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/jor.21203 [Accessed: 7 November 2022].

Yuksel, E., Kalkan, S., Cekmece, S., Unver, B. and Karatosun, V. 2017. Assessing Minimal Detectable Changes and Test-Retest Reliability of the Timed Up and Go Test and the 2-Minute Walk Test in Patients With Total Knee Arthroplasty. *The Journal of Arthroplasty* 32(2), pp. 426–430. doi: 10.1016/J.ARTH.2016.07.031.

Zeni, J.A., Flowers, P., Bade, M., Cheuy, V., Stevens-Lapsley, J. and Snyder-Mackler, L. 2019. Stiff knee gait may increase risk of second total knee arthroplasty. *Journal of Orthopaedic Research* 37(2), pp. 397–402. Available at: https://pubmed.ncbi.nlm.nih.gov/30387528/ [Accessed: 5 March 2021].

Zeni, J.A. and Snyder-Mackler, L. 2010a. Early Postoperative Measures Predict 1- and 2-Year Outcomes After Unilateral Total Knee Arthroplasty: Importance of Contralateral Limb Strength. *Physical Therapy* 90(1), pp. 43–54. Available at: https://academic.oup.com/ptj/article/90/1/43/2737635 [Accessed: 1 May 2023].

Zeni, J.A. and Snyder-Mackler, L. 2010b. Preoperative predictors of persistent impairments during stair ascent and descent after total knee arthroplasty. *Journal of Bone and Joint Surgery - Series A* 92(5), pp. 1130–1136. Available at: /pmc/articles/PMC2859903/?report=abstract [Accessed: 8 February 2021].

Zingde, S.M. and Slamin, J. 2017. Biomechanics of the knee joint, as they relate to arthroplasty. *Orthopaedics and Trauma* 31(1), pp. 1–7. doi: 10.1016/J.MPORTH.2016.10.001.

Appendix A - Knee Osteoarthritis and Injury Outcome Score (KOOS)

KOOS KNEE SURVEY

Today's date: _____/_____/_____ Date of birth: _____/_____/_____

Name: _____

INSTRUCTIONS: This survey asks for your view about your knee. This information will help us keep track of how you feel about your knee and how well you are able to perform your usual activities.
Answer every question by ticking the appropriate box, only one box for each question. If you are unsure about how to answer a question, please give the best answer you can.

Symptoms
These questions should be answered thinking of your knee symptoms during the **last week**.

S1. Do you have swelling in your knee?

Never	Rarely	Sometimes	Often	Always
☐	☐	☐	☐	☐

S2. Do you feel grinding, hear clicking or any other type of noise when your knee moves?

Never	Rarely	Sometimes	Often	Always
☐	☐	☐	☐	☐

S3. Does your knee catch or hang up when moving?

Never	Rarely	Sometimes	Often	Always
☐	☐	☐	☐	☐

S4. Can you straighten your knee fully?

Always	Often	Sometimes	Rarely	Never
☐	☐	☐	☐	☐

S5. Can you bend your knee fully?

Always	Often	Sometimes	Rarely	Never
☐	☐	☐	☐	☐

Stiffness
The following questions concern the amount of joint stiffness you have experienced during the **last week** in your knee. Stiffness is a sensation of restriction or slowness in the ease with which you move your knee joint.

S6. How severe is your knee joint stiffness after first wakening in the morning?

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

S7. How severe is your knee stiffness after sitting, lying or resting **later in the day**?

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Pain

P1. How often do you experience knee pain?

Never	Monthly	Weekly	Daily	Always
☐	☐	☐	☐	☐

What amount of knee pain have you experienced the **last week** during the following activities?

P2. Twisting/pivoting on your knee

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

P3. Straightening knee fully

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

P4. Bending knee fully

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

P5. Walking on flat surface

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

P6. Going up or down stairs

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

P7. At night while in bed

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

P8. Sitting or lying

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

P9. Standing upright

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Function, daily living

The following questions concern your physical function. By this we mean your ability to move around and to look after yourself. For each of the following activities please indicate the degree of difficulty you have experienced in the **last week** due to your knee.

A1. Descending stairs

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

A2. Ascending stairs

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

For each of the following activities please indicate the degree of difficulty you have experienced in the **last week** due to your knee.

A3. Rising from sitting

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

A4. Standing

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

A5. Bending to floor/pick up an object

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

A6. Walking on flat surface

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

A7. Getting in/out of car

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

A8. Going shopping

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

A9. Putting on socks/stockings

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

A10. Rising from bed

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

A11. Taking off socks/stockings

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

A12. Lying in bed (turning over, maintaining knee position)

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

A13. Getting in/out of bath

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

A14. Sitting

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

A15. Getting on/off toilet

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

For each of the following activities please indicate the degree of difficulty you have experienced in the **last week** due to your knee.

A16. Heavy domestic duties (moving heavy boxes, scrubbing floors, etc)

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

A17. Light domestic duties (cooking, dusting, etc)

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Function, sports and recreational activities

The following questions concern your physical function when being active on a higher level. The questions should be answered thinking of what degree of difficulty you have experienced during the **last week** due to your knee.

SP1. Squatting

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

SP2. Running

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

SP3. Jumping

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

SP4. Twisting/pivoting on your injured knee

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

SP5. Kneeling

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Quality of Life

Q1. How often are you aware of your knee problem?

Never	Monthly	Weekly	Daily	Constantly
☐	☐	☐	☐	☐

Q2. Have you modified your life style to avoid potentially damaging activities to your knee?

Not at all	Mildly	Moderately	Severely	Totally
☐	☐	☐	☐	☐

Q3. How much are you troubled with lack of confidence in your knee?

Not at all	Mildly	Moderately	Severely	Extremely
☐	☐	☐	☐	☐

Q4. In general, how much difficulty do you have with your knee?

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Thank you very much for completing all the questions in this questionnaire.

Appendix B – Oxford Knee Score (OKS)

The Oxford Knee Score

During the past four weeks (**please, complete both L & R**):

1. How would you describe the pain you usually have from your knee?

L R

☐ ☐ None
☐ ☐ Very mild
☐ ☐ Mild
☐ ☐ Moderate
☐ ☐ Severe

2. Have you had any trouble with washing and drying yourself (all over) because of your knee?

L R

☐ ☐ 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 a car or using public transport because of your knee? (whichever you tend to use)

L R

☐ ☐ No trouble at all
☐ ☐ Very little trouble
☐ ☐ Moderate trouble
☐ ☐ Extreme difficulty
☐ ☐ Impossible to do

4. For how long have you been able to walk before the pain from your knee becomes severe? (with or without a stick)

L R

☐ ☐ No Pain/ > 30min
☐ ☐ 16 to 30 min
☐ ☐ 5 to 15 min
☐ ☐ Around the house only
☐ ☐ Not at all – severe on walking

5. After a meal (sat at table), how painful has it been for you to stand up from a chair because of your knee?

L R

☐ ☐ Not at all painful

☐ ☐ Slightly painful

☐ ☐ Moderately painful

☐ ☐ Very painful

☐ ☐ Unbearable

6. Have you been limping when walking, because of your knee?

L R

☐ ☐ 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?

L R

☐ ☐ Yes, easily

☐ ☐ With a little difficulty

☐ ☐ With moderate difficulty

☐ ☐ With extreme difficulty

☐ ☐ No, impossible

8. Have you been troubled by pain from your knee in bed at night?

L R

☐ ☐ No nights

☐ ☐ Only 1 or 2 nights

☐ ☐ Some nights

☐ ☐ Most nights

☐ ☐ Every night

9. How much has pain from your knee interfered with your usual work? (including housework)

L R

- ☐ ☐ Not at all
- ☐ ☐ A little bit
- ☐ ☐ Moderately
- ☐ ☐ Greatly
- ☐ ☐ Totally

10. Have you felt that your knee might suddenly "give way" or let you down?

L R

- ☐ ☐ Rarely/never
- ☐ ☐ Sometimes or just at first
- ☐ ☐ Often, not just at first
- ☐ ☐ Most of the time
- ☐ ☐ All of the time

11. Could you do the household shopping on your own?

L R

- ☐ ☐ Yes, easily
- ☐ ☐ With little difficulty
- ☐ ☐ With moderate difficulty
- ☐ ☐ With extreme difficulty
- ☐ ☐ No, impossible

12. Could you walk down a flight of stairs?

L R

- ☐ ☐ Yes, easily
- ☐ ☐ With little difficulty
- ☐ ☐ With moderate difficulty
- ☐ ☐ With extreme difficulty
- ☐ ☐ No, impossible

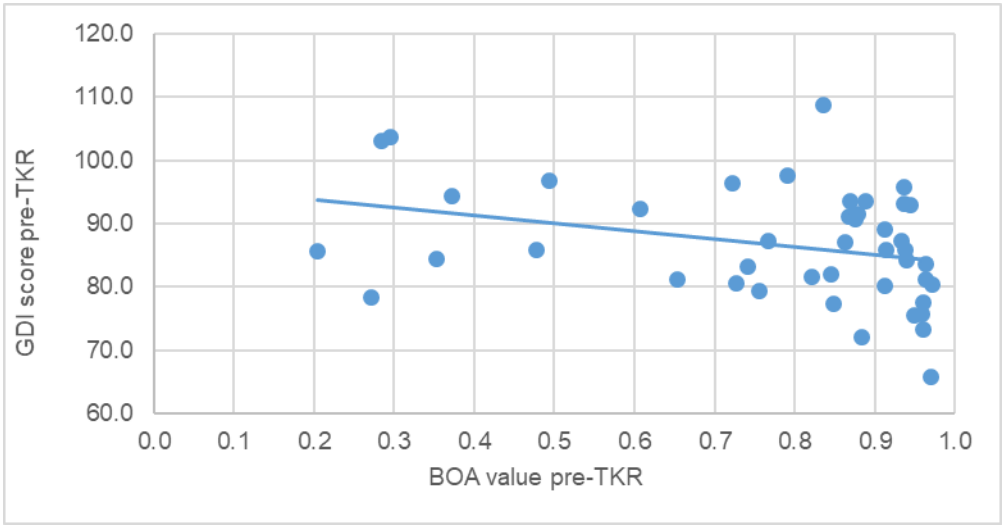
Thank you for completing these questions!

Appendix C – Scatter plots showing the relationship between the BOA vs. GDI and BOA vs. GDI-kinetic pre and one year post-TKR

Classifier and GDI relationship pre-TKR

The scatter plot in Figure 48b showed two outliers in the Karolinska cohort, due to 2 patients having a very low BOA pre-TKR (minimally compromised gait) compared to the rest of the group.

a



b

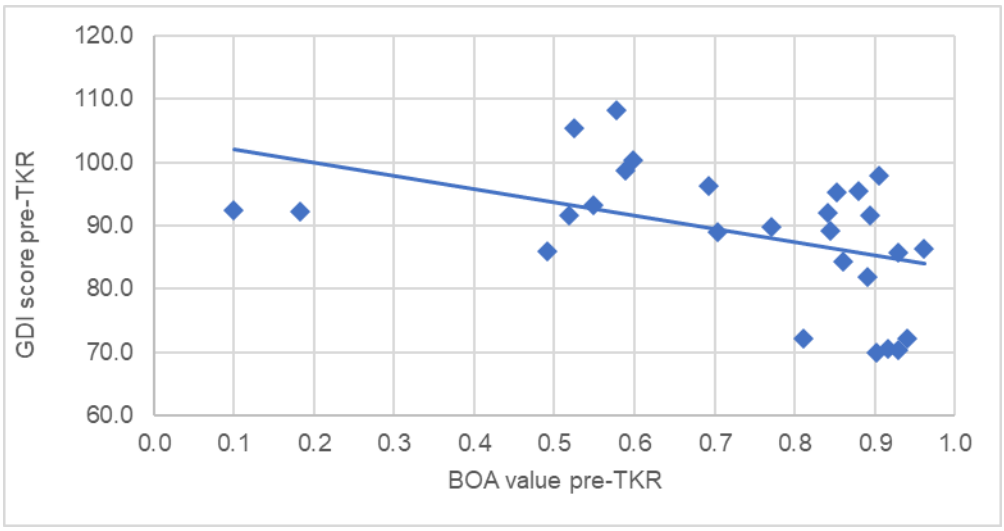
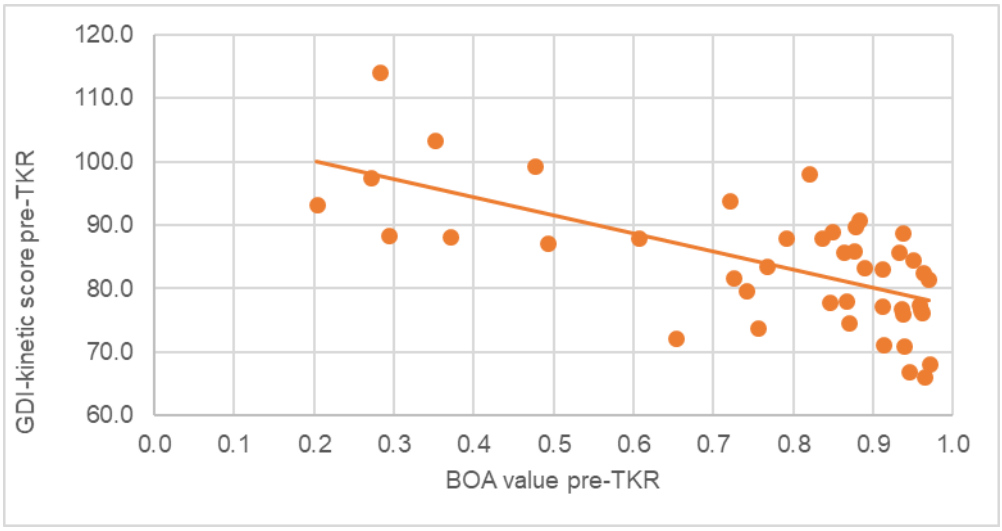


Figure 48 Scatterplots on the relationship between the BOA (= Belief of Osteoarthritis) values and the GDI (= Gait Deviation Index) scores pre-TKR (= Total Knee Replacement). **a**) Cardiff patients (light blue dots); **b**) Karolinska patients (blue diamonds).

Classifier and GDI-kinetic relationship pre-TKR

Figure 49a revealed no outliers in the Cardiff cohort. Conversely, two outliers were identified in the Karolinska group, Figure 49b, due to 2 patients having a very low BOA pre-TKR (minimally compromised gait).

a



b

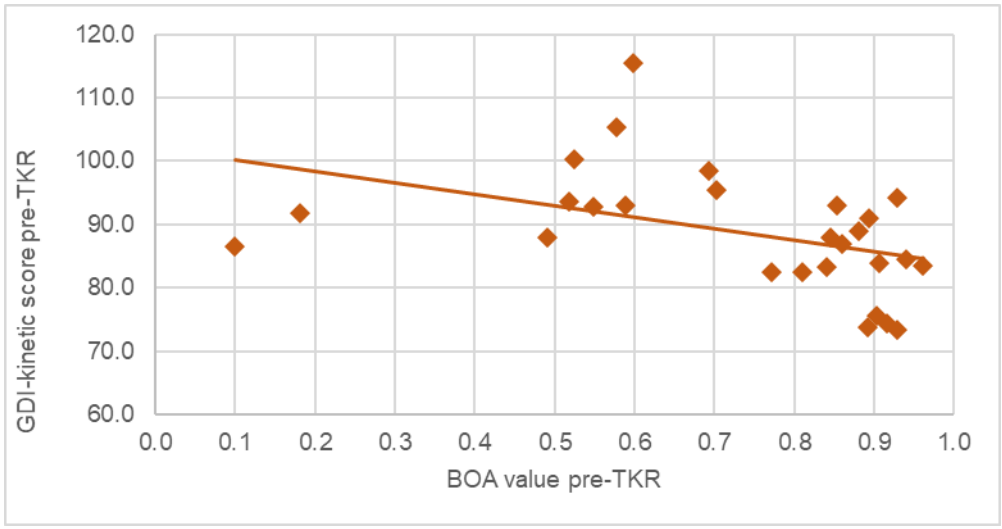
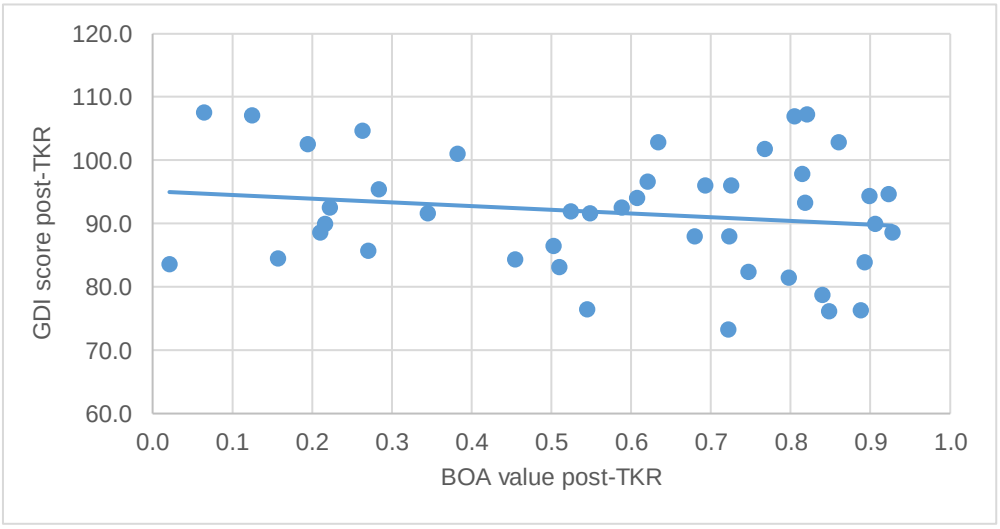


Figure 49 Scatterplots on the relationship between the BOA (= Belief of Osteoarthritis) values and the GDI-kinetic (= Gait Deviation Index) scores pre-TKR (= Total Knee Replacement). **a**) Cardiff patients (light orange dots); **b**) Karolinska patients (dark orange diamonds).

Classifier and GDI relationship post-TKR

a



b

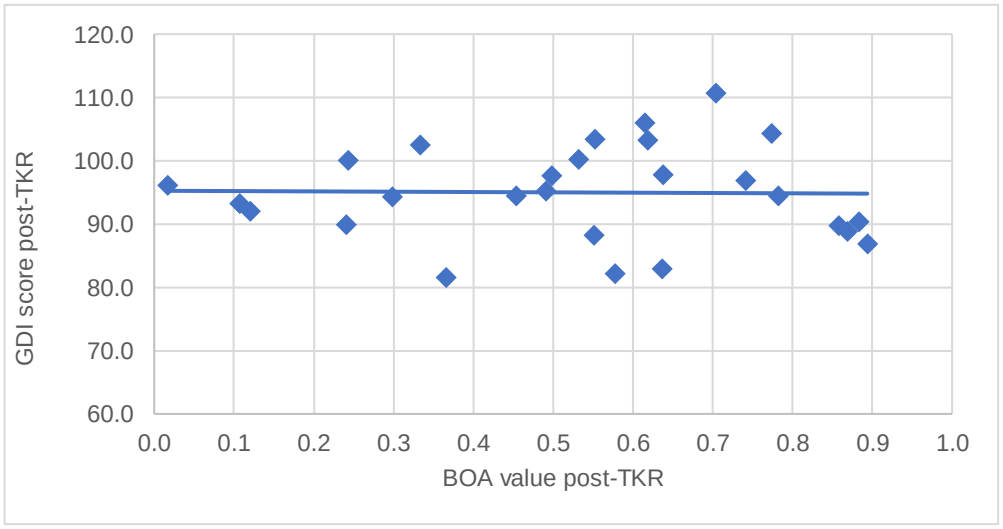
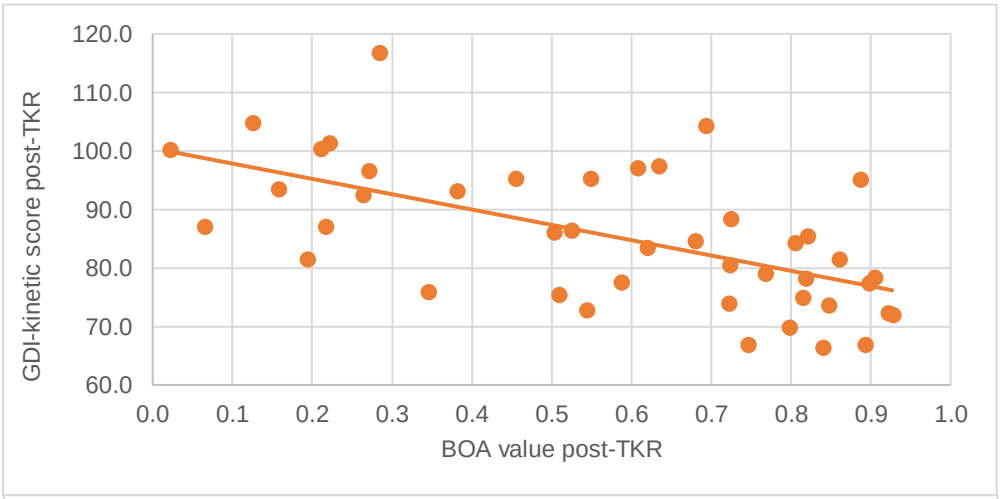


Figure 50 Scatterplots on the relationship between the BOA (= Belief of Osteoarthritis) values and the GDI (= Gait Deviation Index) scores post-TKR (= Total Knee Replacement). **a**) Cardiff patients (light blue dots); **b**) Karolinska patients (blue diamonds).

Classifier and GDI-kinetic relationship post-TKR

a



b

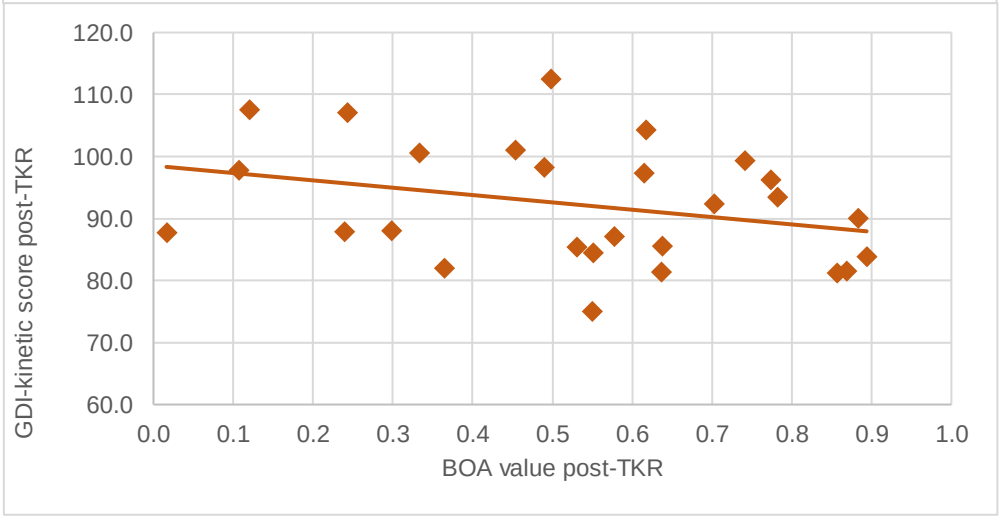
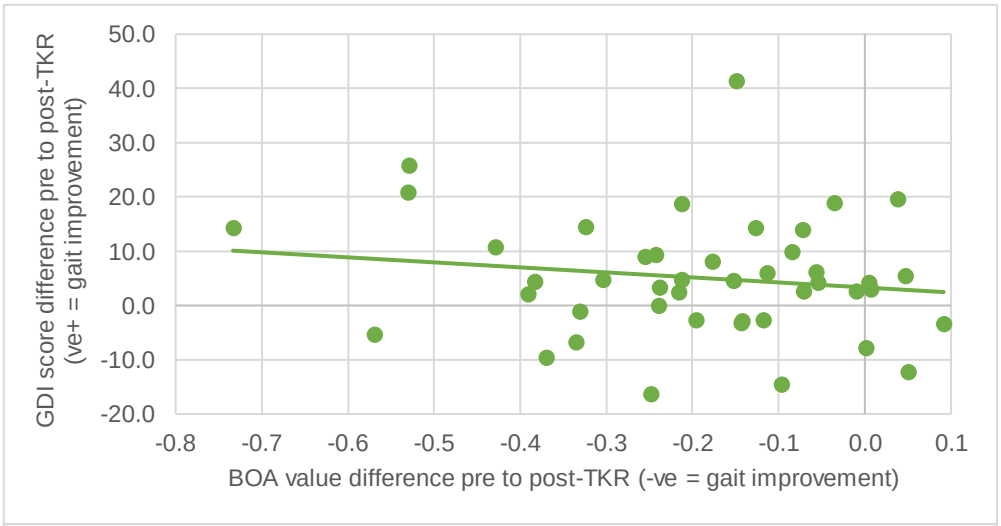


Figure 51 Scatterplots on the relationship between the BOA (= Belief of Osteoarthritis) values and the GDI-kinetic (= Gait Deviation Index Kinetic) scores post-TKR (= Total Knee Replacement). **a**) Cardiff patients (light orange dots); **b**) Karolinska patients (dark orange diamonds).

Appendix D – Scatter plots showing the relationship between the BOA change vs. GDI change pre to one year post-TKR

BOA change and GDI change relationship

a



b

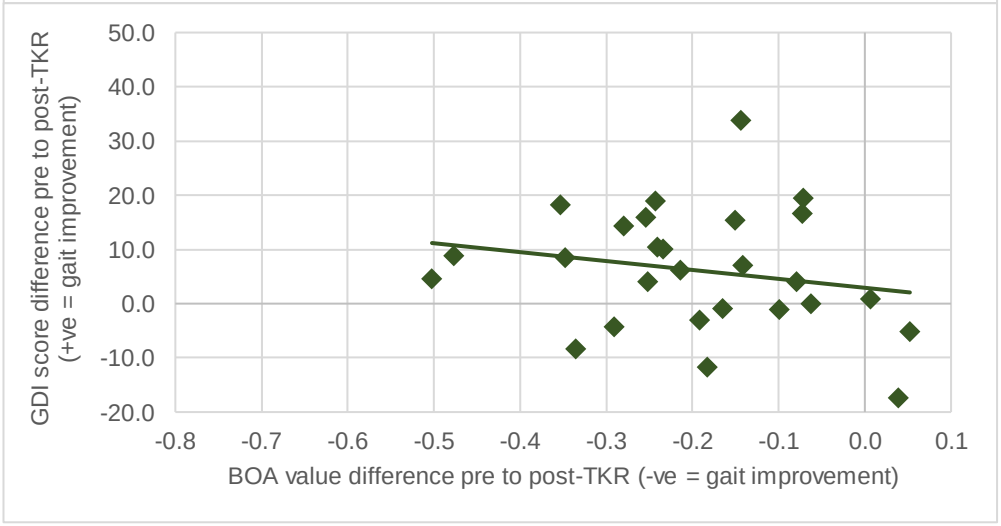


Figure 52 Scatterplots on the relationship between the BOA (= Belief of Osteoarthritis) and the GDI (= Gait Deviation Index) change pre to post-TKR (= Total Knee Replacement). **a**) Cardiff patients (light green dots); **b**) Karolinska patients (dark green diamonds).

There were some outliers in the data, the following cases had an improvement in gait function that was greater than the rest of the cohort:

- Cardiff patient 00090 changed BOA by -0.73 points (beyond 3 SD from the group mean change: -0.20 ± 0.18) (Figure 52a)

- Cardiff patient 02596, change in GDI by 41.6 points (almost 3 SD from the group mean change: 5.14 ± 10.9) but a decrease in the BOA only by 0.15 points (Figure 52 a)
- Karolinska patient K20 who had an improvement in GDI by 33.8 points (almost 3 SD from the group mean change: 6.1 ± 11.12) and a limited decrease in the BOA of 0.14 (Figure 52b)

Classifier change and GDI-kinetic change relationship

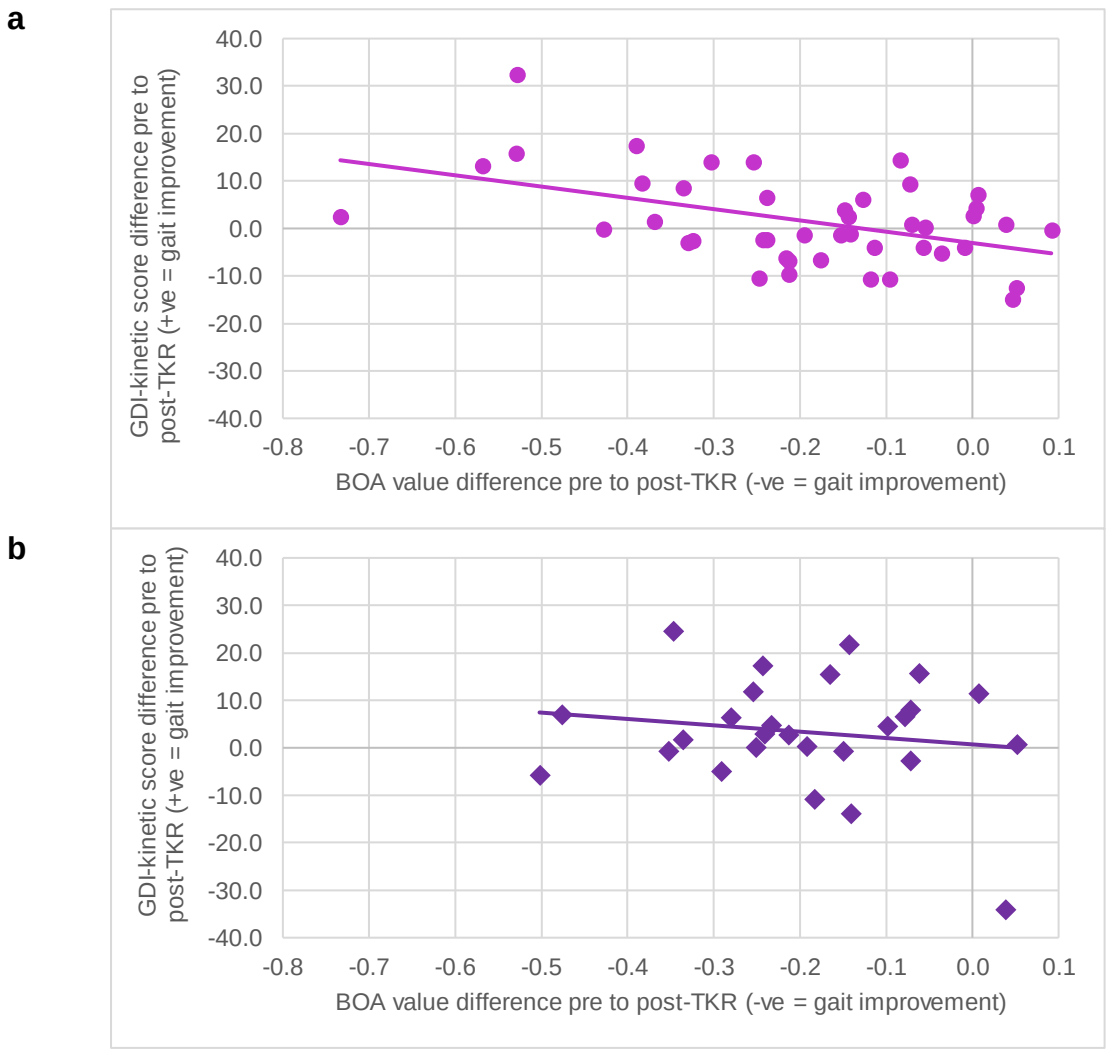


Figure 53 Scatterplots on the relationship between the BOA (= Belief of Osteoarthritis) and the GDI-kinetic (= Gait Deviation Index) change pre to post-TKR (= Total Knee Replacement). a) Cardiff patients (light purple dots); b) Karolinska patients (dark purple diamonds).

There was one outlier in the Cardiff patients (Figure 53a), patient 02518, with a change in GDI-kinetic by 32.6 points (almost 3 SD from the group mean: 1.62 ± 9.24). and one outlier in the Karolinska patients (K23) (Figure 53b), whose GDI kinetic decreased (=gait deterioration) by 34.0 points (over 3 SD from the group mean change: 3.3 ± 11.8).

Appendix E – Case Report Form (surgical data)

ARTHRITIS RESEARCH UK | Biomechanics & Bioengineering Centre | CARDIFF UNIVERSITY / PRIFYSGOL CAERDYDD

CASE REPORT FORM

Total knee replacement surgery Visit – Knee Project

Centre Project Name: _____

Patient Initials ☐☐☐ Centre Database ID: ☐☐☐☐☐

Date of Visit: __ __ / __ __ __ / __ __ __ __ Location: _____

Ensure relevant informed consent is in place.

Please, tick/cross the following boxes when each of the inclusion and exclusion criteria have been checked.

INCLUSION CRITERIA	**Yes**	**No**
1. Within the age range from 18-80.	☐	☐
EXCLUSION CRITERIA (if any of the following answers is yes, the participant will be excluded)	**Yes**	**No**
2. Inability to provide written informed consent.	☐	☐
3. Presence of any previous injury to the joint under investigation that the treating clinician or researcher deems unsuitable.	☐	☐
4. Other pathologies e.g. neurological (Parkinson's disease, Alzheimer's disease, MS, previous cerebrovascular accident), visual or vestibular conditions which might affect the way they move.	☐	☐
5. Inability to walk 10 meters without the use of a mobility aid.	☐	☐

CASE REPORT FORM

Total knee replacement surgery Visit – Knee Project

PRE-SURGERY INFORMATION		
1.	Compartments Affected by OA	☐ Medial ☐ Lateral ☐ Patella
2.	Patient ASA Grade	☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Charlson Comorbidity Index		
1.	Age	☐ <50 years (0) ☐ 50-59 years (+1) ☐ 60-69 years (+2) ☐ 70-79 years (+3) ☐ ≥80 years (+4)
2.	Myocardial Infarction	☐ No (0) ☐ Yes (+1)
3.	Congestive Heart Failure	☐ No (0) ☐ Yes (+1)
4.	Peripheral Vascular disease	☐ No (0) ☐ Yes (+1)
5.	Cerebrovascular disease	☐ No (0) ☐ Yes (+1)
6.	Dementia	☐ No (0) ☐ Yes (+1)
7.	Chronic Obstructive Pulmonary disease	☐ No (0) ☐ Yes (+1)
8.	Connective Tissue disease	☐ No (0) ☐ Yes (+1)
9.	Peptic Ulcer disease	☐ No (0) ☐ Yes (+1)
10.	Diabetes Mellitus	☐ None or diet-controlled (0) ☐ Uncomplicated (+1) ☐ End-organ damage (+2)
11.	Moderate to severe Chronic Kidney disease	☐ No (0) ☐ Yes (+2)
12.	Hemiplegia	☐ No (0) ☐ Yes (+2)

CASE REPORT FORM

Total knee replacement surgery Visit – Knee Project

13.	Leukemia	☐ No (0) ☐ Yes (+2)
14.	Malignant Lymphoma	☐ No (0) ☐ Yes (+2)
15.	Solid Tumor	☐ No (0) ☐ Localised (+2) ☐ Metastatic (+6)
16.	Liver Disease	☐ None (0) ☐ Mild (+1) ☐ Moderate to severe (+3)
17.	AIDS	☐ No (0) ☐ Yes (+6)

POST-SURGERY INFORMATION		
5.	Minimally invasive technique?	☐ Yes ☐ No
6.	Intra-operative events	☐ Fracture ☐ Patella Tendon Avulsion ☐ Ligament Injury ☐ Other (specify:_____) ☐ No
7.	Patella resurfacing?	☐ Yes ☐ No
8.	Implant design	☐ Posterior-Stabilised ☐ Posterior Cruciate Ligament Retaining ☐ Bi-cruciate Ligaments Retaining
9.	Tibial Component	☐ Fixed-Bearing ☐ Mobile-Bearing
10.	Post-surgery complications?	☐ Infection ☐ Other (specify:_____) ☐ No

CASE REPORT FORM

Total knee replacement surgery Visit – Knee Project

I have reviewed this CRF and confirm that, to the best of my knowledge, it accurately reflects the study information obtained for this participant. All entries were made either by me or by a person under my supervision who is named on the local Delegation and Signature Log.

Name of Researcher: _____

Signature: _____

Date: 16 February 2023

This form must be sent to the Centre Research Manager within 14 calendar days of the visit being carried out – if a consent form was completed at the visit then then include the consent form in the same envelope. A copy may also be kept with the researcher for the project records.

Appendix F - Preliminary regression model assumptions (n = 31 patients)

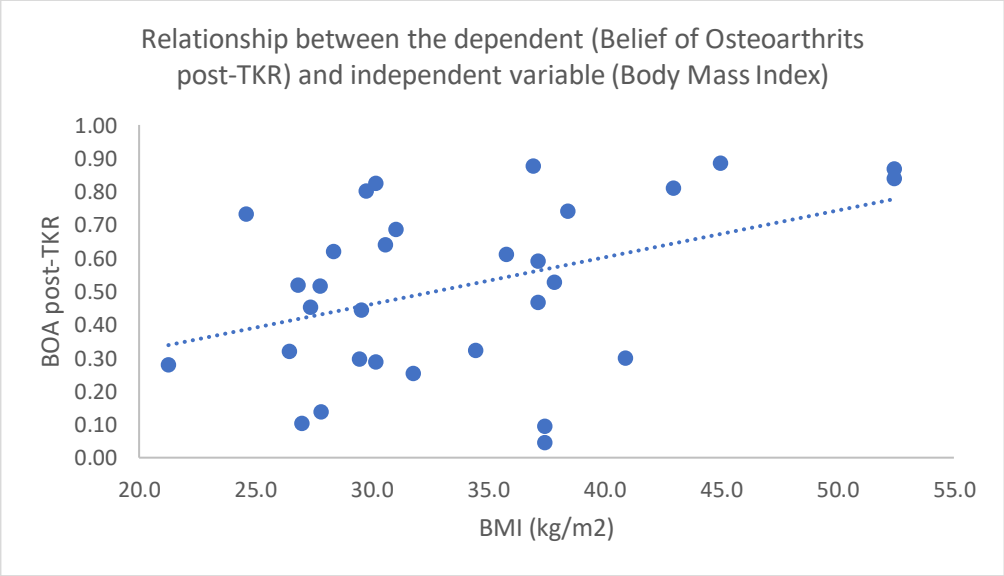
Multivariable regression model (outcome: BOA post-TKR; predictors: BMI, sex, previous major knee surgeries on the affected side, ASA 1 and ASA 2, BOA pre-TKR)

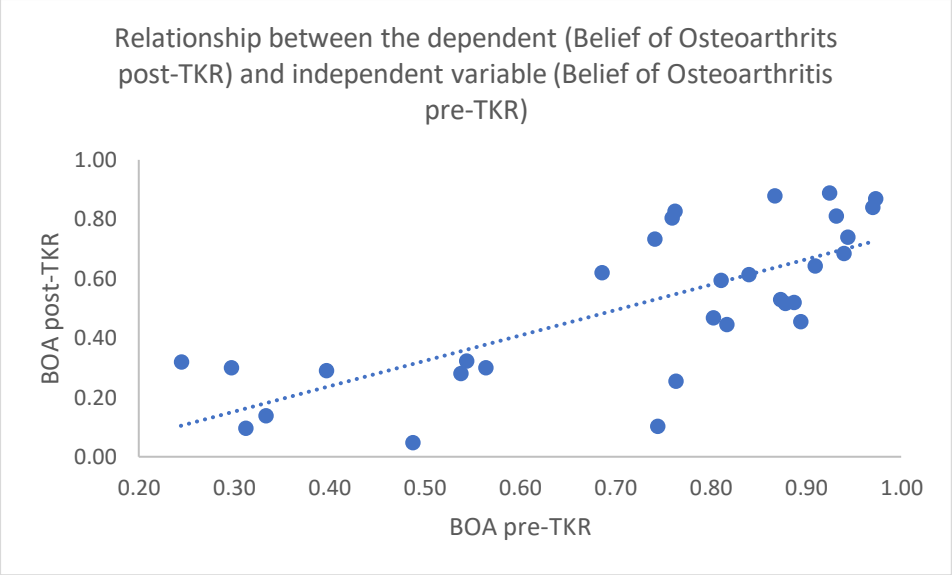
The regression equation was:

$$Predicted\ BOA\ postTKR = \beta_0 + (\beta_1\ x\ BMI) + (\beta_2\ x\ Sex) + (\beta_3\ x\ Previous\ major\ surgeries\ on\ affected\ knee) + (\beta_4\ x\ ASA1) + (\beta_5\ x\ ASA2) + (\beta_6\ x\ BOA\ pre\ TKR)$$

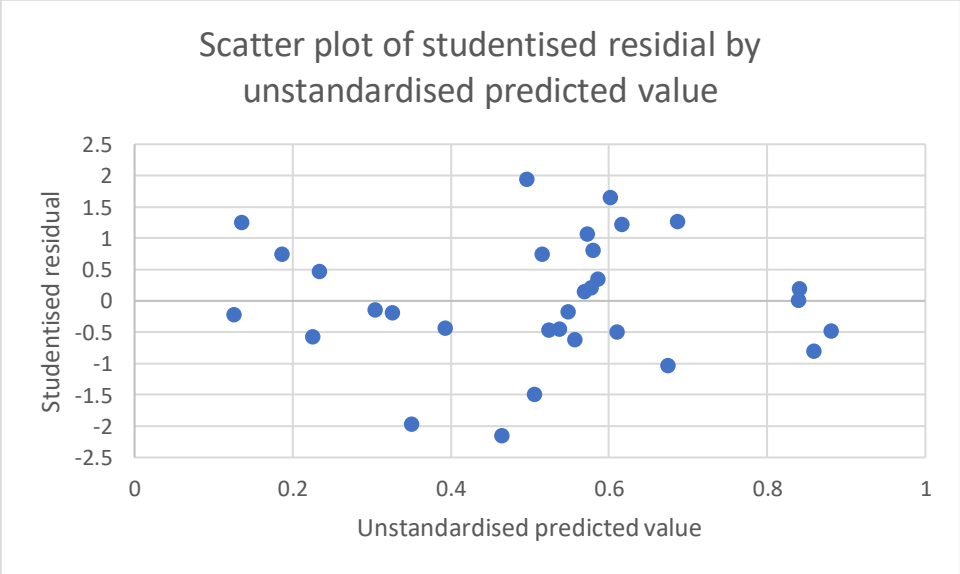
All assumptions of the multivariable regression model were met:

- There was independence of residuals (Durbin-Watson statistics = 2.236)
- There was a linear relationship between all the continuous variables predictors (BMI and BOA pre-TKR) and the outcome (BOA post-TKR)





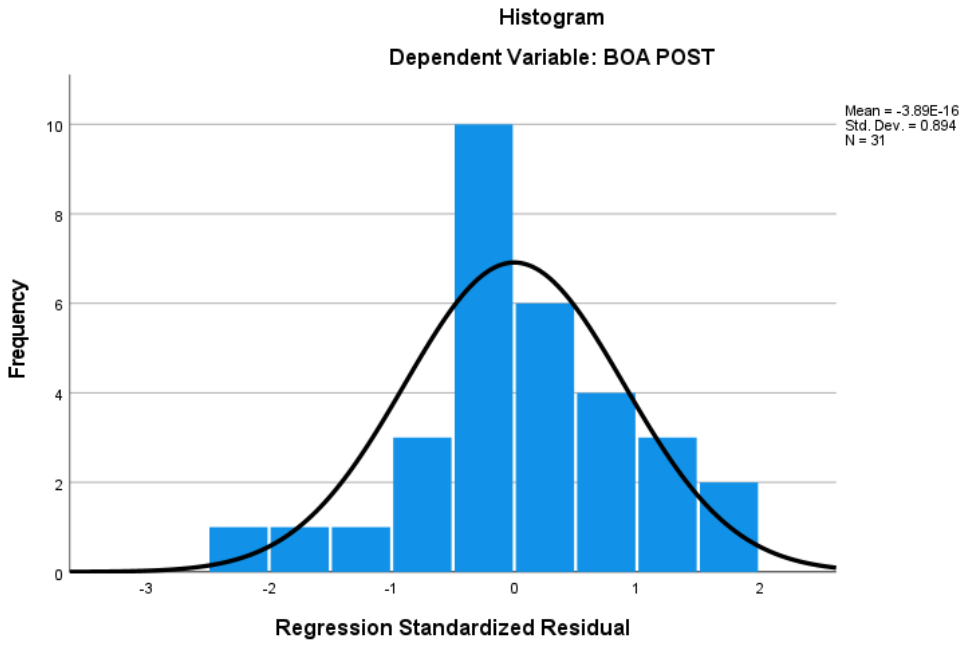
- The plot between the studentised residuals and the unstandardised predicted values showed approximately homoscedasticity

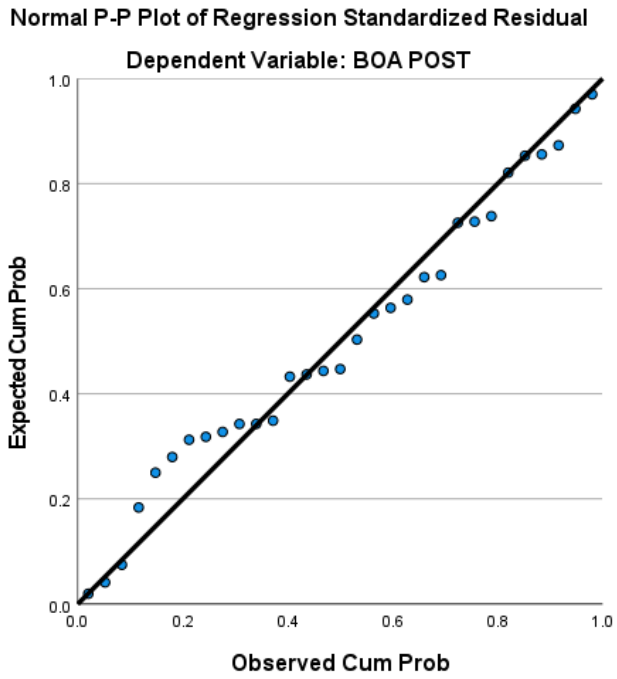


- There were no multicollinearity issues (the variable with the smallest tolerance, ASA 2, had a value of 0.437)

		BOA post-TKR	BMI	Sex	Previous major surgeries (affected knee)	ASA 1	ASA 2	BOA pre-TKR
Pearson Correlation	BOA post-TKR	1.000	0.409	0.358	-0.368	-0.375	0.020	0.738
	BMI		1.000	-0.063	0.107	-0.039	-0.347	0.372
	Sex			1.000	-0.144	-0.215	-0.012	0.255
	Previous major surgeries (affected knee)				1.000	0.116	-0.188	-0.447
	ASA 1					1.000	-0.616	-0.469
	ASA 2						1.000	0.212
	BOA pre-TKR							1.000

- There were no outliers, but there were three patients with higher leverage points that, however, were not influential in the model
- The errors of the prediction were approximately normally distributed as shown by the histogram of the regression standardised residual and the relative P-P plot




Normal P-P Plot of Regression Standardized Residual
Dependent Variable: BOA POST
Expected Cum Prob
Observed Cum Prob
0.0
0.2
0.4
0.6
0.8
1.0

Appendix G - Final regression model assumptions (n = 43 patients)

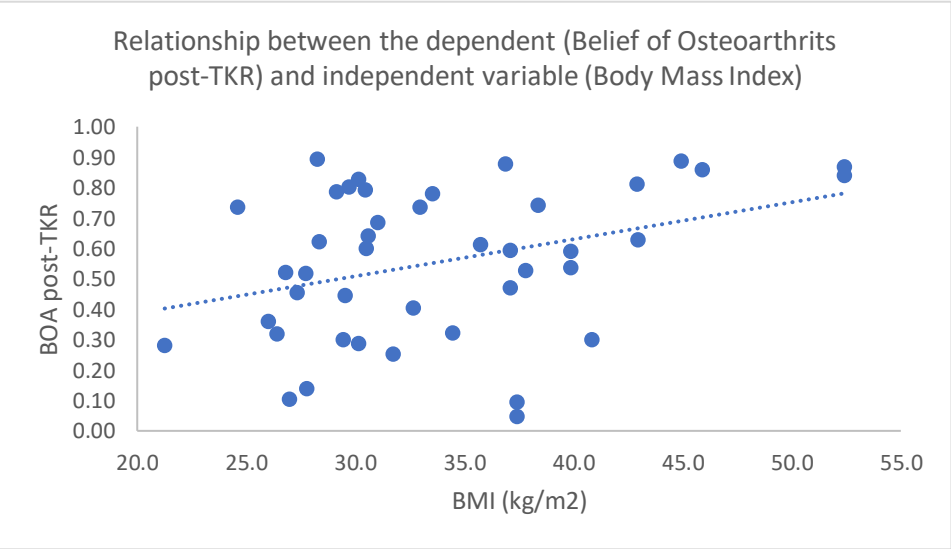
Multivariable regression model (outcome: BOA post-TKR; predictors: BMI, sex, BOA pre-TKR)

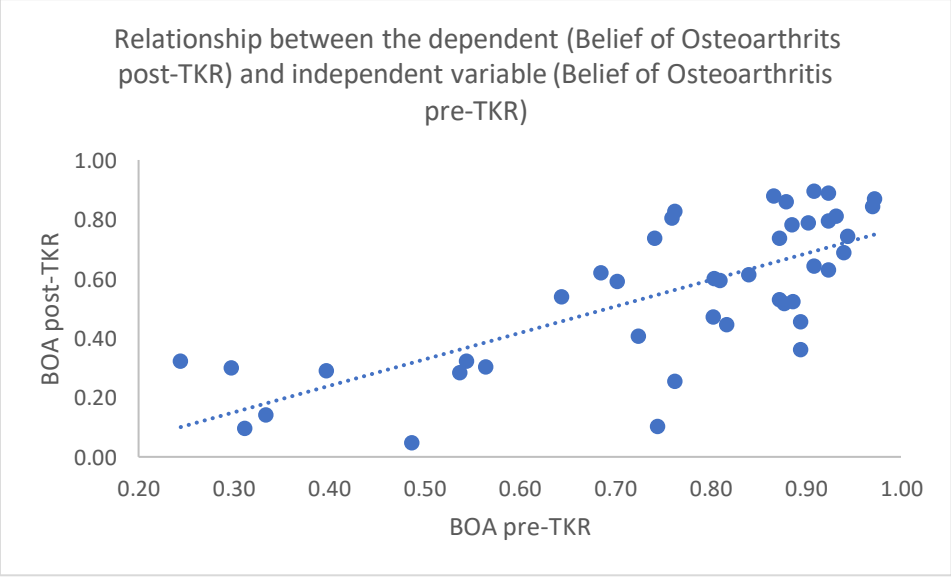
The regression equation was:

$$Predicted\ BOA\ postTKR = \beta_0 + (\beta_1\ x\ BMI) + (\beta_2\ x\ Sex) + (\beta_3\ x\ BOA\ pre\ TKR)$$

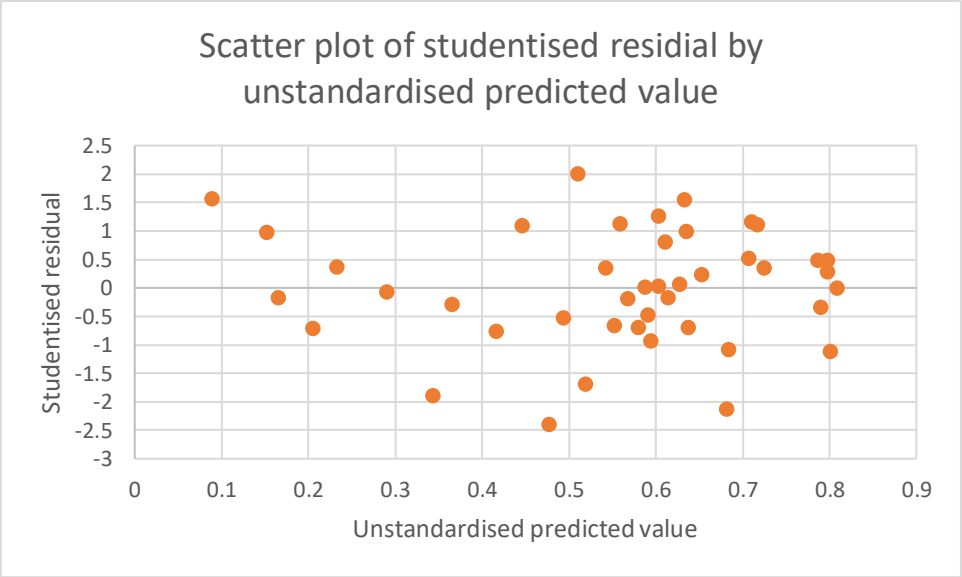
All assumptions of the multivariable regression model were met:

- There was independence of residuals (Durbin-Watson statistics = 1.604)
- There was an approximate linear relationship between the continuous variables predictors (BMI and BOA pre-TKR) and the outcome (BOA post-TKR)





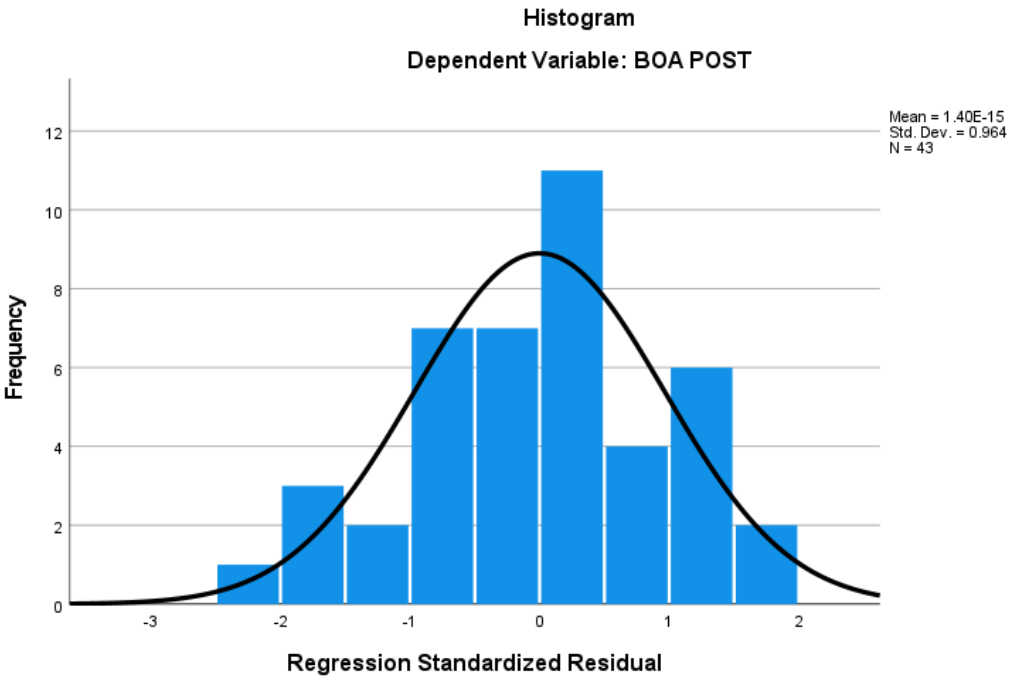
- The plot between the studentised residuals and the unstandardised predicted values showed approximately homoscedasticity

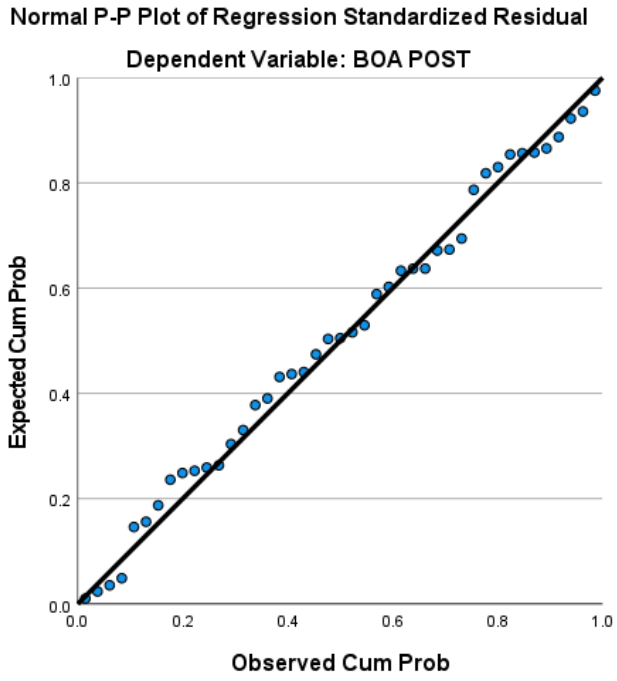


- There were no multicollinearity issues (the variable with the smallest tolerance, BOA pre-TKR, had a value of 0.840)

		BOA post-TKR	BMI	Sex	BOA pre-TKR
Pearson Correlation	BOA post-TKR	1.000	0.356	0.379	0.734
	BMI		1.000	0.059	0.285
	Sex			1.000	0.297
	BOA PRE				1.000

- There were no outliers
- The errors of the prediction were normally distributed, as shown by the histogram of the regression standardised residuals and the P-P plot of the regression standardised residual


Histogram
Dependent Variable: BOA POST
Mean = 1.40E-15
Std. Dev. = 0.964
N = 43
Frequency
Regression Standardized Residual


Normal P-P Plot of Regression Standardized Residual
Dependent Variable: BOA POST
Expected Cum Prob
Observed Cum Prob

Appendix H - MATLAB code for data filtering and normalisation

```
% This code was written by Marina De Vecchis (Cardiff University) in autumn
% 2021 to extract the fluoroscopy data (step-up activity) collected in
% Australia. The code picks the data for each participant from an excel
% file, filters the data, selects the data from the start of the movement
% (beginning of knee extension) to the end of step-up (knee straigh),
% without including the parts at the start where
% the participant's acceleration is close to zero. The code then normalises
% the data from 0 to 100% of the task.

% What you need: an excel file with the participants' data. The
% phases (or stages) of the task should have already be determined via an
"R"
% code previously developed (see Catherine's Galvin PhD thesis appendix for
reference).
% The phases have been determined using the angular acceleration from the
% knee flexion-extension angles.

% Follow the instructions that came with the code

%% Path where all the data is stored:

clear all
clearvars %Remove items from workspace, freeing up system memory

%Specify default file path (saves time)
cd 'C:';
Defaultpath = 'C:\';

[FileName,PathName] = uigetfile({'*.xls;*.xlsx','Excel Files
(*.xls,*.xlsx)'},'Select the data file',Defaultpath);

%% To extract initial info from the file
%Reads the excel file, sheet 1. Separates out (and creates a numeric matrix
for each)
%the numbers, text, and then raw has both numbers and text.
[num,txt,raw] = xlsread(FileName, 1);

%Looks at the dimensions of the raw data (and creates a numeric matrix with
%the values for "height" and "width")
[height, width] = size(raw);

%Takes the content of raw 1 (Which are the variables)
Labels=raw(1, :);

%To sort the data in participant number ascending order, based on column 1
in "num"
numsorted = sortrows(num, 1, 'ascend');

%Takes the content of columnn 1 from "numsorted" (which is the subjects ID)
ParticipantID=numsorted(:, 1);

%% To return the number of participants in the document opened
%rmmissing removes the "NaN" values that the command "unique", unique
treats
%NaN values as distinct.
ParticipantIDwithoutNaN=rmmissing(ParticipantID);
```

```
%unique=returns the same data as the ones specified in the round brackets
%but with no repetitions and in a sorted order
Participants = unique(ParticipantIDwithoutNaN);

%numel=returns the number of elements in the object in the brackets
n = numel (Participants);

%f = msgbox (message) creates a pop-up message
%num2str converts a numeric array into a character array that represents
%the numbers. In this case it reports the number of participants in the
matrix "n"
f = msgbox(['The code found ' num2str(n) ' participants.']);

%% To define in which column the data for a new participant starts and
finishes
% from the Matlab help page: Find locations of repeated values?

A = numsorted(:,1); % take the first column from "numsorted" which contains
the participants' IDs
A = A'; % flip the matrix so the data are in a raw rather than in a column
% (this is to simplify the next step)

buffer = [true, diff(A)~=0];
% with a 1 when a new number is present in A and 0 when the following
% number is the same as the preceding one

groupstart = find(buffer); % returns the column number in "buffer" where
there is a "1"
% find=returns a vector containing the linear indices of each nonzero
element in the array specified in the brackets.

groupId = cumsum(buffer); % returns a matrix containing the cumulative
% sums for each column of buffer

groupSize = accumarray(groupId.', ones(size(groupId))).'; % returns a
series of numbers
% each of which indicate how many columns have the same number

start = groupstart(groupSize > 2); % returns the column number where the ID
of
% a different patient starts. This is the important bit that can be used
% further

% to create a new element that contains the number of the columns where the
data for one participant finish
finish = start(1, 2:end); %takes all the values from the varible "start"
from column 2 to the end of raw 1
finish = finish-1; %subtracts 1 to each value in the new variable "finish"
finish (end + 1) = (height-1); %adds a column with a number that
corresponds to the height of the
% original raw data -1. This is because the data we are using in the
% following loop are retrieved from the array "NumDataFlipped" which has a
% width that is one column smaller than the height of the raw data (this is
% because we removed the column names).

%% Get the data for each single participant and filter+smooth+normalise it
% Loop preparation
NumDataFlipped = numsorted'; %to transpose to make it easier for the for
loop to pick the values needed
allFlexNorm = [];
allExtIntNorm = [];
allAbdAddNorm = [];
```

```
allMLNorm = [];
allAPNorm = [];
allCompDistNorm = [];

%%the following is for the butterworth filter calculations
  freqCutoff = 4 %cutoff frequency
   freqSample = 30 %frequency at which the data was recorded (30Hz)
   [b,a]=butter(4,freqCutoff/(freqSample/2))

for i = 1:n %n is the number of participants and it will always be
equivalent to the number of columns in the array "start"
    B1 = start(1, i);
    B2 = finish(1, i);
    EachParticipant = NumDataFlipped(:, B1:B2); %extracts the data for each
participant

    %% NORMALISE

%data preparation
flexExt1 = EachParticipant(4, :);
intExtRot1 = EachParticipant(5, :);
abdAdd1 = EachParticipant(6, :);
MLshift1 = EachParticipant(7, :);
APshift1 = EachParticipant(8, :);
compDist1 = EachParticipant(9, :);

% To find where the task starts and ends exluding the beginning and the end
% of the task where there is no knee movement in the sagittal plane
    phases = EachParticipant (2, :); % extracts raw 2 where there is the
phase data
    buffer2 = [true, diff(phases)~=0];
    % with a 1 when a new number is present in A and 0 when the following
    % number is the same as the preceding one
    phaseFind = find(buffer2); % returns the column number where there is a
"1"
    phaseId = cumsum(buffer2); % returns a matrix containing the cumulative
    % sums for each column of buffer2

    % find where is the first column that has a "2"
    Cond1 = phaseId == 2;
    x = cumsum(Cond1,2) == 1 & Cond1;
    out = sum(x.*phaseId,2);
    out(~any(Cond1,2)) = nan;
    idxFlexStart = x*(1:size(phaseId,2))';

    % find where is the first column that has a "5" (= the end of knee
    % flexion at the end of the step-down)
    % change the value to "3" if you just want to keep the step-up data up
    % to the end of step-up
    Cond1 = phaseId == 3;
    x2 = cumsum(Cond1,2) == 1 & Cond1;
    out2 = sum(x2.*phaseId,2);
    out2(~any(Cond1,2)) = nan;
    idxExtEnd = x2*(1:size(phaseId,2))';
    idxExtEnd = idxExtEnd - 1; % column number where the flexion ends

%%================================
%% Extracts the data from the beginning of the step up to the end of the
% step up
    flexExt = flexExt1 (idxFlexStart:idxExtEnd);
    intExtRot = intExtRot1 (idxFlexStart:idxExtEnd);
```

```
    abdAdd = abdAdd1 (idxFlexStart:idxExtEnd);
    MLshift = MLshift1 (idxFlexStart:idxExtEnd);
    APshift = APshift1 (idxFlexStart:idxExtEnd);
    compDist = compDist1 (idxFlexStart:idxExtEnd);
%=======================================
  n2 = numel (flexExt(1, :)); %number of frames in the data cropped

% This produces and error message and stops the execution of the code
% in case the code was not able to calculate the start and end of the task
% correctly for a participant due to the lack of one or more of the five
% phases
    if n2 == 0
        disp (['The code found that at least one participant does not have
all the five phases for the step-up task, please check and correct the data
in the Excel file'])
    break
    end

% Normalises the data to 101 datapoints (0-100% of the step-up task) using
% cubic spline data interpolation
    x = linspace(0, 100, n2); %y = linspace(x1,x2,n) generates n points.
The spacing between the points is (x2-x1)/(n-1)
    y = 0:100;
    FlexExtNorm = spline(x, flexExt, y);
    IntExtRotNorm = spline(x, intExtRot, y);
    AbdAddNorm = spline(x, abdAdd, y);
    MLShiftNorm = spline(x, MLshift, y);
    APShiftNorm = spline(x, APshift, y);
    CompDistNorm = spline(x, compDist, y);


%% filtering the data with a 4th order lowpass Butterworth filter with a
cutoff frequency of 4Hz
%  and then making it so the filter does not make all data to start from a
%  "zero" value, in other words, making it so the filter has zero lag
%  (=zero-phase filtering)
   flexExtButterA = filter(b,a,FlexExtNorm); %this would make all the data
to start from a value of zero which is not realistic
   flexExtButter = filtfilt(b,a,FlexExtNorm); %this is the zero-phase
filtering
   intExtRotButterA = filter(b,a,IntExtRotNorm);
   intExtRotButter = filtfilt(b,a,IntExtRotNorm);
   abdAddButterA = filter(b,a,AbdAddNorm);
   abdAddButter = filtfilt(b,a,AbdAddNorm);
   MLShiftButterA = filter(b,a,MLShiftNorm);
   MLShiftButter = filtfilt(b,a,MLShiftNorm);
   APShiftButterA = filter(b,a,APShiftNorm);
   APShiftButter = filtfilt(b,a,APShiftNorm);
   compDistButterA = filter(b,a,CompDistNorm);
   compDistButter = filtfilt(b,a,CompDistNorm);

% Saves the data from all loops in new matrices
    allFlexNorm = [allFlexNorm; flexExtButter];
    allExtIntNorm = [allExtIntNorm; intExtRotButter];
    allAbdAddNorm = [allAbdAddNorm; abdAddButter];
    allMLNorm = [allMLNorm; MLShiftButter];
    allAPNorm = [allAPNorm; APShiftButter];
    allCompDistNorm = [allCompDistNorm; compDistButter];

%=============================================
    allFlexRaw{i} = (flexExt);
```

```
%     allExtIntRaw = [intExtRot];
%     allAbdAddRaw = [abdAdd];
%     allMLRaw = [MLshift];
%     allAPRaw = [APshift];
%     allCompDistRaw = [compDist];
end


%% SAVE RESULTS IN AN EXCEL FILE
% creates an array with the %of activity cycle in the first raw (the first
% column name is empty as this is where the participant IDs will go later
on
% in the code)
G = [nan, 0:100];

% creates a new array that merges the participant's IDs (1st column) and
the
% calculated biomechanical outputs
KneeFlexExt = [Participants allFlexNorm];
KneeAbdAdd = [Participants allAbdAddNorm];
KneeIntExtRot = [Participants allExtIntNorm];
KneeAPShift = [Participants allAPNorm];
KneeMLShift = [Participants allMLNorm];
KneeCompression = [Participants allCompDistNorm];

% creates a new array that merges the participant's IDs (1st column) and
the
% calculated biomechanical outputs with the %gait cycle created before
KneeFlexExt = [G; KneeFlexExt];
KneeAbdAdd = [G; KneeAbdAdd];
KneeIntExtRot = [G; KneeIntExtRot];
KneeAPShift = [G; KneeAPShift];
KneeMLShift = [G; KneeMLShift];
KneeCompression = [G; KneeCompression];

% creates a new folder where to save the results, changes the path to the
new folder
currDate = strrep(datestr(datetime), ':', '_'); %gets the current date and
time
Datafolder = [PathName 'Results_', currDate]; %creates an element with the
text address
% of the current folder to which \Results is added at the end + the current
% date and time
mkdir (Datafolder); %Creates the results folder
oldFolder = cd(Datafolder); %changes the current folder to the one where
the Results folder was created


%save results in separate excel files
sheetname = 'FlexionExtension';
writematrix (KneeFlexExt, 'FlexExt.xlsx', 'Sheet', sheetname)

sheetname2 = 'AddAbd';
writematrix (KneeAbdAdd, 'AddAbd.xlsx', 'Sheet', sheetname2)

sheetname3 = 'IntExtRot';
writematrix (KneeIntExtRot, 'IntExtRot.xlsx', 'Sheet', sheetname3)

sheetname4 = 'APShift';
writematrix (KneeAPShift, 'APShift.xlsx', 'Sheet', sheetname4)

sheetname5 = 'MLShift';
```

```
writematrix (KneeMLShift, 'MLShift.xlsx', 'Sheet', sheetname5)

sheetname6 = 'CompDistr';
writematrix (KneeCompression, 'CompDist.xlsx', 'Sheet', sheetname6)

%% DISPLAY PLOTS OF THE RESULTS
% Flexion-extension
figure
plot(y, allFlexNorm)
line(xlim(), [0,0], 'LineWidth', 2, 'Color', 'k'); % to have a line on zero
grid on; %to have a grid
grid minor; %to show the minor grid lines
yticks(-20:5:100)
title('Knee flexion (+)')
xlabel('% Step-up')
ylabel('Degrees')

% Adduction/abduction
figure
plot(y, allAbdAddNorm)
line(xlim(), [0,0], 'LineWidth', 2, 'Color', 'k'); % to have a line on zero
grid on; %to have a grid
grid minor; %to show the minor grid lines
title('Knee adduction - varus (+)')
xlabel('% Step-up')
ylabel('Degrees')

% Internal/external rotation
figure
plot(y, allExtIntNorm)
line(xlim(), [0,0], 'LineWidth', 2, 'Color', 'k'); % to have a line on zero
grid on; %to have a grid
grid minor; %to show the minor grid lines
title('Knee internal rotation (+) - Femur relative to tibia')
xlabel('% Step-up')
ylabel('Degrees')

% AP translation
figure
plot(y, allAPNorm)
line(xlim(), [0,0], 'LineWidth', 2, 'Color', 'k'); % to have a line on zero
grid on; %to have a grid
grid minor; %to show the minor grid lines
title('Femur anterior translation (+)')
xlabel('% Step-up')
ylabel('mm')

% ML shift
figure
plot(y, allMLNorm)
line(xlim(), [0,0], 'LineWidth', 2, 'Color', 'k'); % to have a line on zero
grid on; %to have a grid
grid minor; %to show the minor grid lines
title('Knee medial shift (+) - Femur relative to tibia')
xlabel('% Step-up')
ylabel('mm')

%Compression/distraction
figure
plot(y, allCompDistNorm)
line(xlim(), [0,0], 'LineWidth', 2, 'Color', 'k'); % to have a line on zero
```

```
grid on; %to have a grid
grid minor; %to show the minor grid lines
title('Knee distraction (+)')
xlabel('% Step-up')
ylabel('mm')
```