

CARDIFF UNIVERSITY

**Classification of
Biomechanical Changes in
Gait Following Total Knee
Replacement: An Objective,
Multi-feature Analysis**

Paul Robert Biggs

PhD Thesis

Biomedical Engineering Research Group

School of Engineering

December 2016

Abstract

Incidence of osteoarthritis (OA) is steadily increasing amongst the developed world, with the knee being the most commonly affected joint. Knee OA is a complex, progressive and multifactorial disease which can result in severe disability, pain, and reduced quality of life. Numerous biomechanical changes have been associated with OA disease progression within both the affected and unaffected joints. Total knee replacement (TKR) is a common surgical intervention which aims to replace the degenerated articular surfaces. As longevity of the prostheses have improved, TKR surgery is being recommended to an increasingly younger population. There is, however, a growing body of evidence to suggest a proportion of patients exhibit several functional limitations following surgery. Measuring functional changes is challenging, and numerous studies suggest patient-reported changes in physical function aren't reflective of objectively measured changes. This study builds upon techniques to objectively assess biomechanical function during level gait using three-dimensional stereophotogrammetry, with an aim to quantify biomechanical changes that occur as a result of late-stage OA, and measure and summarise functional changes following TKR surgery.

Firstly, the appropriateness of principal component analysis (PCA) and the Cardiff Dempster-Shafer Theory (DST) classifier to reduce and summarise level gait biomechanics is investigated within a cohort of 85 OA and 38 non-pathological (NP) subjects. The validity of previously adopted rules for retaining principal components (PCs) is assessed; namely the application of Kaiser's rule, and a factor loading threshold of ± 0.71 . Through the reconstruction of biomechanical waveforms using individual PCs, it is demonstrated that this rule discards biomechanical features which can accurately distinguish between OA and NP gait biomechanics. The currently accepted definitions of two control parameters of the DST classifier, which define the shape of the sigmoid activation function, are shown to introduce a bias under certain conditions. New definitions are proposed and tested, which result in an increase in classification accuracy. The robustness of the leave-one-out (LOO) cross-validation algorithm to assess the performance of the classification is investigated, and findings suggest little benefit of retaining larger cohorts within the cross-validation set. Training bodies of different sizes are investigated, and their ability to classify the remaining data is evaluated. Results indicated that a training body of ten subjects in each group resulted in high classification accuracy ($92\% \pm 2.5\%$), and improvements in accuracy then began to steadily plateau.

The techniques developed thus far are then adopted to classify the hip, knee and ankle biomechanics of 41 OA and 31 NP subjects, to describe the biomechanical characteristics of late-stage OA. There were numerous methodological changes within this section of the study, and it was proved necessary to recalculate new PCs using this cohort. These new PCs were contextualised and used to classify OA biomechanics, resulting in a LOO classification accuracy of 98.6%. Anecdotally, the single misclassified subject had late-stage OA, but reported only mild functional impairments. The biomechanical features which consistently distinguished OA gait are ranked and discussed.

The trained DST classifier was used to quantify the biomechanical function of 22 subjects pre and 12-months post-TKR surgery. In contrast to previous findings using the DST technique, biomechanical improvements varied, with no clear group of improvers. Five subjects were classified as NP post-operatively, seven were classified as "non-dominant OA", and ten as "dominant OA". Objectively measured function was significantly correlated with two out of nine patient-reported outcome measures both before surgery, and in all nine post-operatively. This might explain discrepancies in the literature between patient-reported and objectively measured changes. A retrospective analysis explored pre-operative predictors highlighted knee and ankle coronal plane angulation at heel strike, ankle range of motion, and timing of peak knee flexion as potential predictors of post-operative function.

Acknowledgements

I would firstly like to acknowledge how recklessly I embarked on this PhD journey, and how incredibly fortunate I am to have chosen a studentship supervised by Dr Gemma Whatling and Prof Cathy Holt. I have no doubt in my mind that without their encouragement and support, in both technical and emotional matters, I would never have had the opportunity to write my PhD acknowledgements. I would also like to thank the Arthritis Research UK Biomechanics and Bioengineering Centre (ARUKBBC) at Cardiff University for funding my research, and for providing a strong interdisciplinary research network.

My appreciation goes to my fellow researchers, particularly David Williams, Nidal Khatib, Dr Sarah Forrest, Dr Philippa Jones, Aseel Ghazwan, and Hassanain Ali Lafta, whom became both close companions and collaborators. Late-night data processing to meet looming conference abstract deadlines was made unnervingly enjoyable by our office challenges.

I am indebted to all the participants of this study, and those who helped with the long motion analysis sessions. A special thank you goes to Health and Care Research Wales, who provided support during the data collection process, both in patient recruitment and within the motion laboratory.

A very special thank you goes to Georgie Le'Fjord. My life was made richer by your presence, and my vigour mightier by your absence. Your passion for science, perpetual intrigue, and blissful wonderment of the universe overpowered my cynicism and pessimism and reignited my passion for research.

Lastly, a huge thank you for my incredible partner Tasmin. Writing this accursed document was rapidly leeching the joy from life, and you have been a limitless IV infusion of fun and silliness which has kept me sane, happy, and content. When I started this process I envisaged this write-up period to be one of the hardest challenges of my life,

but with your support I was somehow able to stay relatively calm as I was continually thwarted by Murphy's Law, and Microsoft's word processor.

Contents

Abstract	i
Acknowledgements	ii
Contents	iv
Abbreviations.....	x
Chapter 1 - Introduction	1
1.1 Scope.....	1
1.2 Aims and Objectives.....	2
1.3 Motivation	3
Chapter 2 - Literature Review	6
2.1 Osteoarthritis and Mechanical Loading	6
2.1.1 <i>Envelope of Function</i>	9
2.1.2 <i>Joint Mechanics and Biological Changes</i>	9
2.1.3 <i>Conservative Management of Knee OA.</i>	12
2.2 Total Knee Replacement.....	16
2.2.1 <i>Choice of TKR Design Within the UK</i>	16
2.2.2 <i>TKR Outcomes</i>	17
2.2.3 <i>Outcome Measures</i>	19
2.2.4 <i>Rehabilitative Factors</i>	21
2.3 Human Motion Analysis.....	22
2.4 Data Reduction.....	26
2.4.1 <i>Computing Principal Components</i>	27
2.4.2 <i>Further Techniques</i>	29
2.5 Classification / Data Summation	30
2.5.1 <i>Gait Indexes</i>	31
2.5.2 <i>Artificial Intelligence</i>	34

2.5.3	<i>Supervised Training</i>	35
2.5.4	<i>Dempster-Shafer Theory</i>	38
2.5.5	<i>Conversion to a Body of Evidence</i>	41
2.5.6	<i>Dempster's Combination of Evidence</i>	42
2.5.7	<i>Comparisons with Neural Networks</i>	43
Chapter 3 - Objective Assessment of Knee Function During Gait		46
3.1	<i>Introduction</i>	46
3.2	<i>Data Collection</i>	49
3.2.1	<i>Non-pathological Subject Recruitment</i>	49
3.2.2	<i>Osteoarthritis and TKR Patient Recruitment</i>	49
3.2.3	<i>Gait Assessment</i>	50
3.3	<i>Optimising the Pointer Method Pipeline</i>	53
3.3.1	<i>Design Criteria for the Development of the Pointer Method Pipeline</i>	56
3.3.2	<i>Patient Spreadsheet and Data Organisation</i>	57
3.3.3	<i>Knee Kinematics</i>	58
3.3.4	<i>Calculating Knee Kinetics</i>	59
3.3.5	<i>Estimating the Moments About the Knee</i>	62
3.3.6	<i>Data Verification and Saving</i>	65
3.4	<i>Principal Component Analysis</i>	68
3.4.1	<i>Standardisation</i>	68
3.4.2	<i>Correlation Matrix</i>	69
3.4.3	<i>Eigendecomposition</i>	69
3.4.4	<i>Transforming Data Points</i>	71
3.4.5	<i>Calculating Factor Loadings</i>	74
3.4.6	<i>Expanding from 2D to N-Dimensions</i>	75
3.4.7	<i>Optimising the Calculation of Principal Components</i>	75
3.4.8	<i>Retention of Principal Components</i>	76
3.4.9	<i>Reconstructing Data Using PCs</i>	79
3.5	<i>The DST Classifier</i>	82

3.5.1	<i>Defining K</i>	82
3.5.2	<i>Defining Theta</i>	87
3.5.3	<i>Defining the Uncertainty Boundaries</i> :.....	90
3.5.4	<i>Evaluation of Classification Error</i>	94
3.6	Results and Discussions	96
3.6.1	<i>Classifying Using the Same Variables and Principal Components as Jones (2004)</i> 96	
3.6.2	<i>Updated Principal Components</i>	99
3.6.3	<i>Classification Using Updated Principal Component Definitions</i>	110
3.6.4	<i>Leave-P-Out Classification</i>	111
3.6.5	<i>Increasing the Classification Cohort Size</i>	112
3.6.6	<i>Adding ML Forces and Moments to the Classification</i>	114
3.7	Conclusions	125
3.7.1	<i>Exploring the Validity of the Classifier Control Variables</i>	125
3.7.2	<i>Exploring the Sample Size Required to Classify Osteoarthritic Subjects Accurately</i>	126
3.7.3	<i>Assess the Reliability of the LOO Cross-Validation Technique as an Estimate of Classification Accuracy.</i>	127
3.7.4	<i>Does the Inclusion of Mediolateral GRF Force and Knee Joint Moments Have a Significant Impact the Ability to Classify Osteoarthritic Subjects?</i>	128
3.8	Clinical Summary.....	129
3.8.1	<i>Key methodological developments:</i>	129
3.8.2	<i>Key clinical findings</i>	130
Chapter 4 - Classification of Osteoarthritic Hip, Knee and Ankle Gait Biomechanics		132
4.1	Introduction	132
4.2	Methodology	136
4.2.1	<i>Marker Placement</i>	136

4.2.2	<i>Defining the Pelvis</i>	136
4.2.3	<i>Hip Joint Centre Definition</i>	138
4.2.4	<i>Hip, Knee and Ankle Axis Definitions</i>	140
4.2.5	<i>Upsampling to the Analogue Capture Frequency</i>	141
4.2.6	<i>Filtering Data.....</i>	144
4.2.7	<i>Comparison of Previously Defined PCs</i>	147
4.2.8	<i>Initial PCA Selection.....</i>	147
4.2.9	<i>Further PC Retention Using Classification Ranking.....</i>	148
4.3	<i>Results and Discussion</i>	151
4.3.1	<i>Subject Demographics.....</i>	151
4.3.2	<i>Assessing the Appropriateness of Previously Defined PC in Representing Variance Between Subjects Collected with the Updated Methodology.....</i>	152
4.3.3	<i>Principal Component Analysis and Retention</i>	154
4.3.4	<i>Classification Using Top 18 Ranked Variables</i>	155
4.3.5	<i>About the Misclassified Subject.....</i>	171
4.3.6	<i>NP Subject with Lowest Belief in NP Function</i>	172
4.3.7	<i>Assessing the Validity of a Combined Healthy and Elderly Cohort in Classifying OA Subjects</i>	173
4.4	<i>Conclusions</i>	175
4.5	<i>Clinical Summary.....</i>	177
Chapter 5 - Quantifying Functional Changes Following Total Knee Replacement Surgery	179	
5.1	<i>Introduction</i>	179
5.2	<i>Preliminary Work</i>	185
5.3	<i>Methods</i>	189
5.3.1	<i>Participants.....</i>	189
5.3.2	<i>Data Analysis, Processing and Classification</i>	189
5.3.3	<i>Patient-reported Outcome Measures</i>	190

5.3.4	<i>Temporal-spatial Parameters</i>	190
5.3.5	<i>Objective Improvement in Function</i>	192
5.4	<i>Results and Discussion</i>	194
5.4.1	<i>Does Functional Recovery Return Following TKR?</i>	194
5.4.2	<i>Defining Functional Improvement.....</i>	199
5.4.3	<i>Functional Improvement of Each Limb.....</i>	200
5.4.4	<i>Greatest Functional Improvement</i>	204
5.4.5	<i>Do Pre, Post-, and the Relative Change in Subjective Outcome Measures Correlate With Changes in Biomechanical Gait Classification?</i>	207
5.4.6	<i>Predicting Post-Operative Improvement.....</i>	211
5.5	<i>Conclusions</i>	217
5.6	<i>Clinical Summary.....</i>	219
Chapter 6 - Discussions	221	
6.1	<i>Objective 1: Assess the validity and robustness of Jones' application of PCA dimensionality reduction and DST classification in characterising OA gait.</i>	221
6.1.1	<i>Dimensionality Reduction</i>	221
6.1.2	<i>Challenges in PC Reconstruction:</i>	225
6.1.3	<i>DST Classification Control Parameters</i>	226
6.1.4	<i>Robustness of Classification.....</i>	227
6.2	<i>Objective 2: Determine the biomechanical changes in the ankle, knee and hip and due to late-stage osteoarthritis using the methods developed in Objective 1 ..</i>	228
6.2.1	<i>Ground Reaction Forces.....</i>	228
6.2.2	<i>Knee Kinematics</i>	229
6.2.3	<i>Knee Kinetics</i>	230
6.2.4	<i>Hip Kinematics</i>	231
6.2.5	<i>Hip Kinetics</i>	231
6.2.6	<i>Ankle Kinematics.....</i>	231
6.2.7	<i>Ankle Kinetics.....</i>	232

6.3 Objective 3: Objectively measure biomechanical changes following TKR surgery, and elucidate the relationship between pre and post-operative gait biomechanics, and patient-reported outcome.	233
6.3.1 Comparison With PROMs.....	234
6.4 Contributions to Knowledge	235
Chapter 7 - Limitations.....	238
7.1 Variability	238
7.2 Patient Cohort.....	241
7.2.1 Heterogeneity.....	241
7.2.2 Sample Bias	242
7.3 Hardware Changes	242
7.4 Inter-operator Errors.....	243
7.5 Sensitivity and Specificity	243
Chapter 8 Recommendations for future work	244
8.1 PCA Using Multiple Waveforms in One State Space.....	244
8.2 Non-linear PCA.....	245
8.3 Subgrouping Using PCA	246
Chapter 9 References	248

Abbreviations

ACS – Anatomical Coordinate System

ADLs – Activities of Daily Living

AI – Artificial Intelligence

ANN – Artificial Neural Networks

ASIS – Anterior Superior Iliac Spine

BMI – Body Mass Index

BOE – Body of Evidence

CBOE – Combined Body of Evidence

COM – Centre of Mass

COP – Centre of Pressure

DST – Dempster-Shafer Theory

EKAM - External Knee Adduction Moment

EMG – Electromyographic

FGI – Functional Gait Improvement

GCS – Global Coordinate System

GGI – Gillete Gait Index

GRF – Ground Reaction Force

HJC – Hip Joint Centre

HMA – Human Motion Analysis

HS – Heel Strike

IR – Infrared Light

ISB – International Society of Biomechanics

JCS – Joint Coordinate System

KJC – Knee Joint Centre

KL – Kellgren-Lawrence Scale

KOS – Knee Outcome Survey

LCS – Local Coordinate System

LOO – Leave-One-Out

MCS – Marker Clusters

MCS – Marker Coordinate System

ML – Mediolateral

MOCAP - Motion Capture using Opto-Electronic Stereophotogrammetry

NJR – National Joint Registry

NP – Non-Pathological

OA – Osteoarthritis

OKS – Oxford Knee Score

PC – Principal Component

PCA – Principal Component Analysis

PCL - Posterior Cruciate Ligament

PD – Pelvic Depth

PROM - Patient-Reported Outcome Measures

PSIS - Posterior Superior Iliac Spine

QTM – Qualysis Tracker Manager

ROM – Range of Motion

SOP – Standard Operating Procedure

STA – Soft Tissue Artefact

STD – Standard Deviation

SVD – Singular Value Decomposition

TKR – Total Knee Replacement

UKR - Unicondular Knee Replacement

VAS – Visual Analogue Scale

WOMAC - The Western Ontario and McMaster Universities Osteoarthritis Index

Chapter 1 - Introduction

1.1 Scope

This PhD thesis focuses on the objective quantification of changes in lower limb biomechanics during level gait resulting from severe osteoarthritis (OA) of the knee, and subsequent total knee replacement (TKR).

This “objective quantification” could be considered a three-step technique:

1. Human motion analysis – To quantify lower limb biomechanics during level gait.
2. Principal component analysis – To define discrete metrics from temporal biomechanical information.
3. Data classification using Dempster-Shafer Theory – To objectively summarise and weight biomechanical changes associated with OA, and hence quantify recovery following surgery.

The use of Human Motion Analysis (HMA) to calculate lower limb biomechanics during level gait has been adopted by numerous studies to characterise biomechanical changes during OA disease progression, and changes following surgical intervention. Biomechanical changes following TKR surgery have been subject to numerous studies within this research group (Jones *et al.*, 2006, Whatling, 2009, Watling, 2014, Metcalfe, 2014), in collaboration with this research group (Worsley, 2011), and also within a number of other studies well summarised in the systematic review of McClelland *et al.* (2007).

The application of HMA results in a great wealth of temporal data, and objectively describing biomechanical changes following TKR surgery is challenging. Since its application to human gait biomechanics was described by Deluzio *et al.* (1997), PCA has been adopted as a dimensional reduction technique within this research group (Jones and Holt, 2008, Whatling *et al.*, 2008, Watling, 2014), and in the wider biomechanics community (Sadeghi *et al.*, 2002, Chester and Wrigley, 2008, Kirkwood *et al.*, 2011).

While it has been proven as a useful tool in objectively describing biomechanical features, the clinical interpretation of results can be challenging (Brandon *et al.*, 2013).

The Dempster Shafer Theory (DST) classifier has been proven as an alternative to traditional statistics in summarising osteoarthritic changes to gait biomechanics, and progression following TKR surgery (Jones and Holt, 2008). The classification technique has been shown to accurately distinguish late-stage OA subjects from non-pathological (NP) controls (Jones *et al.*, 2008), to out-perform various machine-learning techniques (Jones *et al.*, 2008, Parisi *et al.*, 2015), and to characterise changes in function following TKR surgery (Worsley, 2011). The underlying classification framework is not one widely used within the biomechanics community, and little research has further developed this specific technique since its application to gait biomechanics within Jones (2004). The classification control parameters; θ , k , A and B have a significant impact on classification results, but have undergone little investigation. Furthermore, the robustness of classification has mainly been tested using a leave-one-out (LOO) cross-validation technique, which has come under scrutiny.

1.2 Aims and Objectives

The primary aim of this research was to further develop the application of PCA and the DST classification technique in order to quantify biomechanical changes in gait following severe osteoarthritis and subsequent TKR surgery. This is achieved through the following objectives:

Objective 1 (Chapter 3): *Assess the validity and robustness of Jones' application of PCA dimensionality reduction and DST classification in characterising OA gait.*

Within this chapter, the techniques used to extract, contextualise and select biomechanical features using PCA, will be investigated for their appropriateness. The introduction of the overarching study design, inclusion criteria, and data collection techniques will be introduced. The choice of classification control parameters, choice of

input variables, and robustness of the output classification will then be investigated, and recommendations produced for future studies.

Objective 2 (Chapter 4): *Determine the biomechanical changes in the ankle, knee and hip and due to late-stage osteoarthritis using the methods developed in Objective 1.*

The methods of input variable selection and the choice of NP controls will be investigated. The most robust biomechanical features of OA gait will be determined and ranked in order of their ability to classify between NP and OA subjects.

Objective 3 (Chapter 5): *Objectively measure biomechanical changes following TKR surgery, and elucidate the relationship between pre and post-operative gait biomechanics, and patient-reported outcome.*

The techniques developed within the previous two objectives, including the biomechanical classification of osteoarthritic hip, knee and ankle biomechanics, will be applied to quantify biomechanical changes following TKR surgery. The relationship between gait biomechanics and patient reported outcome measures (PROMs) will be investigated to elucidate the association between subjective patient-reported function and objectively measured joint kinetics and kinematics. Changes in the body of evidence of individual input variables shall be investigated, alongside the summative combined body of evidence, to investigate which specific biomechanical features changed following surgery in both the operative and non-operative limb. The relationship between biomechanical and clinical factors pre and post-TKR surgery will also be investigated to identify clinically feasible predictors of biomechanical outcome following surgery.

1.3 Motivation

The use of human motion analysis had been adopted to measure biomechanical differences or changes within a number of different context: defining “non-pathological movement”, measuring how this changes due to certain conditions or pathologies, how biomechanics change through surgical or conservative interventions, injury risk and several more. The application of human motion analysis allows us to concurrently

measure a great wealth of temporal information. In many contexts, much more is measured than can be concisely analysed or reported. This is particularly true when investigating biomechanical differences between two different cohorts.

This PhD studentship was supported by Arthritis Research UK, and the work was carried out within the Arthritis Research UK Biomechanics and Bioengineering Centre. The interdisciplinary research centre involves close collaborations with surgeons, engineers, biomedical scientists and physiotherapists to investigate NP joint biomechanics, and determine how this is influenced by arthritis, and hence inform clinical intervention. This research often involves the combined collection of biomechanical, biological and clinical factors and outcome measures before and after surgical intervention (see Figure 1.1). Transparent summative measures of changes in biomechanical function following surgical intervention have proven particularly useful

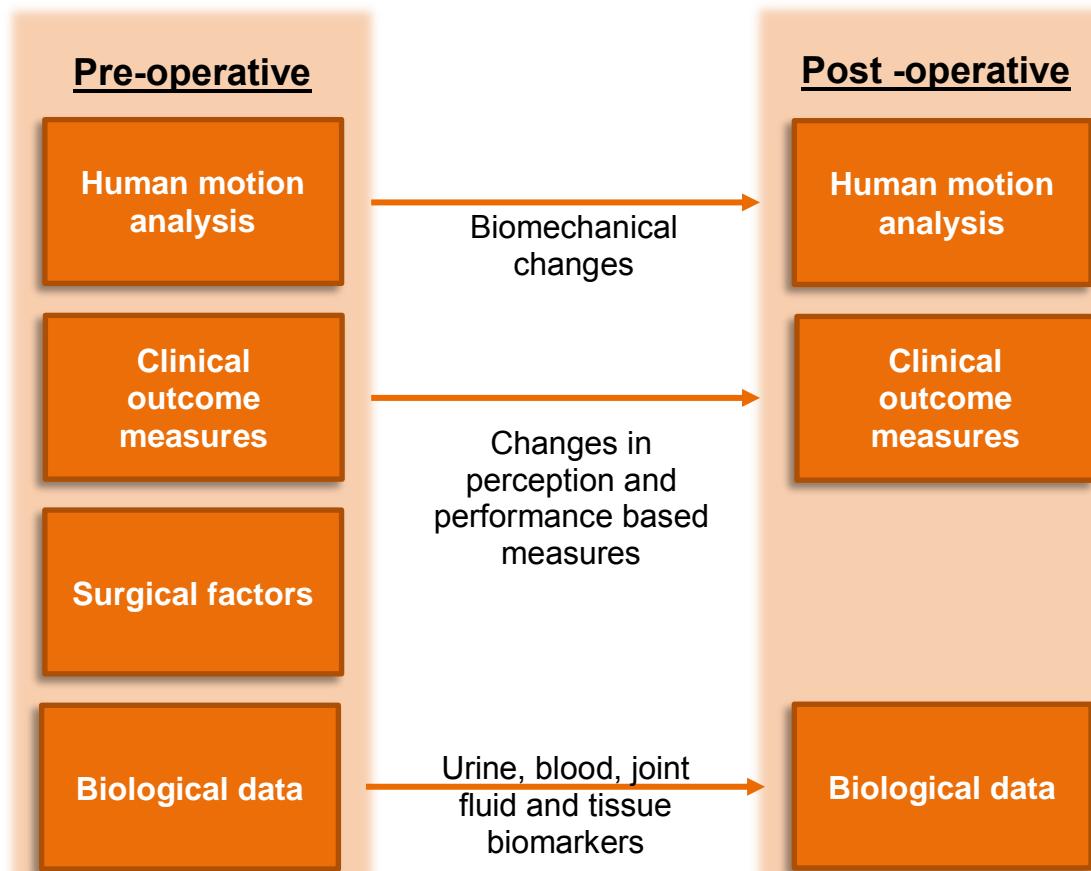


Figure 1.1 Illustrative example of the range of measures taken and examined within the Arthritis Research UK Biomechanics and Bioengineering Centre before and after surgery.

within the context, such as to reveal early signs of a link between biomechanical and biological changes following high-tibial osteotomy surgery (Holt *et al.*, 2016).

While the combined use of PCA and DST classification has shown useful in discriminating OA biomechanics, further work is required to explore the validity of currently adopted techniques. This research chose to further develop and validate these techniques with a specific application towards discriminating biomechanical features of OA level gait, and hence summarising how these biomechanical features change following TKR surgery.

Chapter 2 - Literature Review

2.1 Osteoarthritis and Mechanical Loading

Osteoarthritis (OA) is a progressive disease of synovial joints which is associated with cartilage and bone damage and general joint degeneration; leading to joint pain, stiffness and functional disability (Lane *et al.*, 2011). In the United Kingdom, a third of people aged 45 and over have sought treatment for OA. Of those, over a half sought treatment for OA of the knee (ARUK, 2013). Estimates suggest that the number of people within the UK with knee OA is expected to increase from 4.7 million in 2010, to 5.4 million in 2020, and reaching 6.4million by 2035 (ARUK, 2013). Obesity is a primary risk factor for OA, with very obese people being 14 times more likely to develop the condition than those with a healthy body weight (Coggon *et al.*, 2001). Considering the UK has an ageing population with high functional expectations, alongside growing rates of obesity (UKHF, 2014), it is clear that OA of the knee is a growing problem.

As opposed to viewing OA as a single disease, there is a growing agreement that it actually represents the net effect of a collection of diseases with different causes and potential treatments (Lane *et al.*, 2011). Kraus *et al.* (2015) conducted a review on the terminology surrounding the classification and diagnosis of OA, and identified a need for the standardisation of the definition of OA.

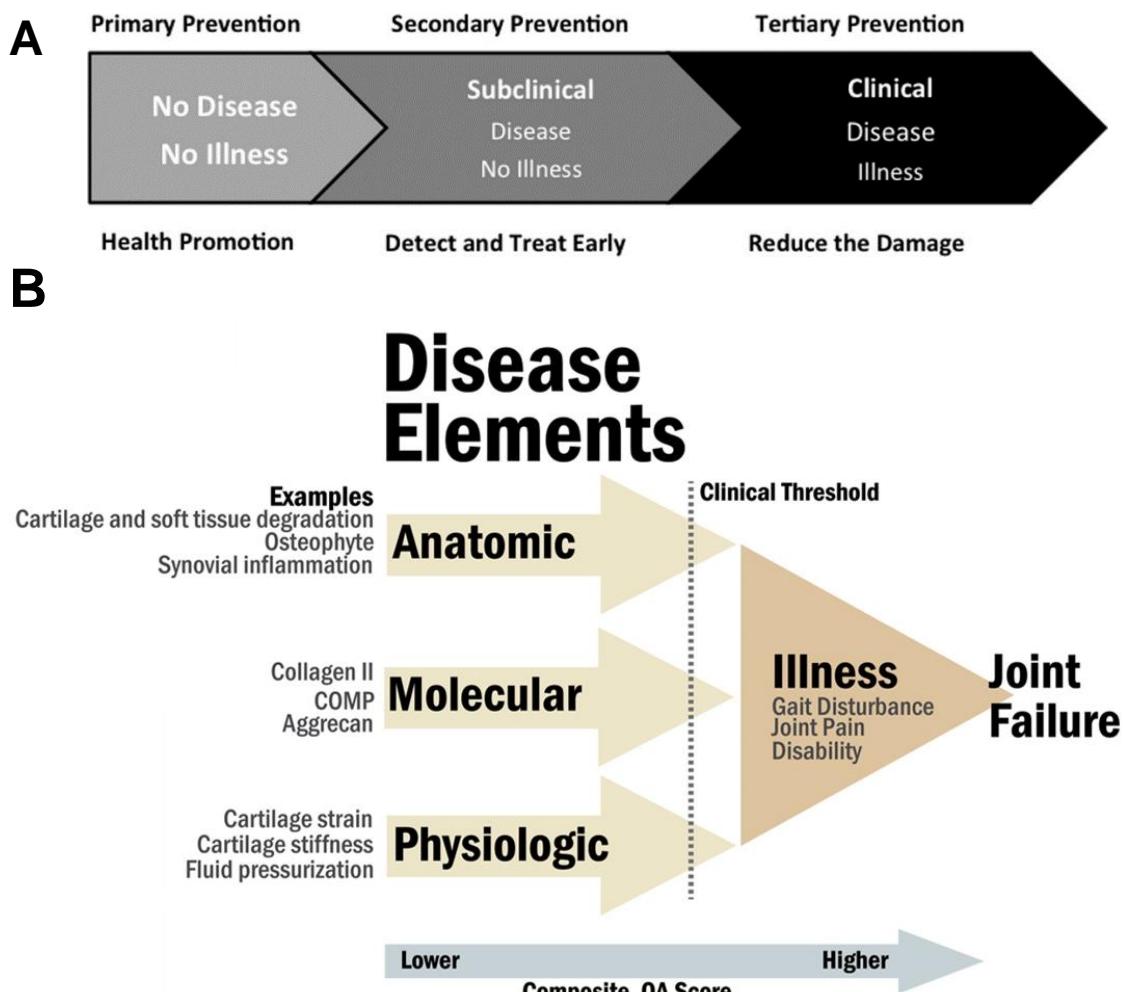
Of the four primary draft definitions collated by (Kraus *et al.*, 2015), the main elements are as follows:

1. OA is a complex disease involving movable joints, which is difficult to diagnose and define.
2. OA subjects are a very heterogeneous group, with large variations in clinical symptoms and outcomes.
3. The specific causes of OA are unknown – it is believed to occur as a result of both mechanical and molecular events.

4. OA is characterised pathologically by cell stress, extracellular matrix degradation and tissue remodelling due to maladaptive repair response, including pro-inflammatory pathways and disruption of the homeostasis of catabolic and anabolic processes.
5. The initial stages of the disease are characterised by abnormal joint tissue metabolism, which eventually leads to macroscopic changes such as joint inflammation, cartilage degeneration, and osteophyte formation, particularly around the joint margins.
6. The clinical condition is characterised by joint pain, tenderness, crepitus, movement limitations, inflammation and occasional effusion.

Within this review, Kraus *et al.* (2015) criticise the United States Food and Drug Administration for having an outdated disease classification system which lacks consideration of molecular causes and instead defines diseases primarily on the signs and symptoms. The authors propose that it is useful to consider the aspects of disease and illness separately. From a preventative perspective, the disease can then be split into three stages, as shown in Figure 2.1A. The first phase is when there is no disease present, and all modifiable risk factors should be minimised. The 'disease elements' have been categorised into 'anatomic', 'molecular', and 'physiologic' (see Figure 2.1B). A subclinical phase of the disease is defined as the period for which the disease is present but clinical symptoms have not yet developed. The challenge at this stage is the detection – there is a period of the disease where there may be some regenerative ability of the articular cartilage; it is currently, however, very difficult to detect OA before irreversible damage has already occurred (Madry *et al.*, 2016). The clinical stage of the disease is considered to be when illness develops and is the primary focus of this thesis. At this stage of the disease, some of the changes may be irreversible. However, there are numerous efforts to slow disease progression including the use of pharmaceuticals (Black *et al.*, 2009), both surgical (Hui *et al.*, 2011) and non-surgical (Raja and Dewan,

2011) biomechanical interventions, and weight management regimes (Christensen et al., 2007).



Composite OA Score

Figure 2.1 A taxonomy proposed by and reprinted from Kraus *et al.* (2015) for the classification of OA.

- A) The three stages of preventative medicine proposed by Katz and Ali (2009) applied to OA prevention and treatment. The goal of primary prevention is to modify risk factors in order to minimise the inception of disease. It is proposed that within the subclinical phase there is presence of disease but not illness. The secondary prevention involved the early detection of this phase and the prevention of progression. The clinical stage of the disease is defined as the initiation of “illness” at which clinical symptoms develop.

B) The taxonomy proposes a composite score of OA which involves all three major domains of the disease elements, alongside their clinical symptoms of ‘illness’. It is anticipated that a clinical threshold would be identified at the transition from disease to illness (i.e. the point at which symptoms occur).

2.1.1 Envelope of Function

While it has also been shown that traumatic or excessive joint loading can lead to cartilage degeneration and OA development, increasing evidence suggests that moderate joint loading at normal physiological levels is necessary to maintain healthy cartilage (Bader *et al.*, 2011). Scott F Dye, a long-practicing orthopaedic surgeon, published a theory of joint function which aims to model how joint homeostasis may be affected by changes in joint loading and load frequency (Dye, 1996). Figure 2.2 shows a visualisation of Dye's envelope of function. The theory suggests that relatively low-load activities, such as walking, can be repeated at a high-frequency without affecting joint homeostasis. The same is said for low-frequency, high-load activities. High-frequency, high-load activities, however, can be considered "supra / sub-physiological" and could cause structural degradation of joint tissues. Equally, if the only activities a person carried out were very low-load, such as sitting, this could be considered a sub-physiological under-load and could, therefore, lead to tissue degeneration such as atrophy.

The theory also helps to visualise how, in pathological subjects, common activities of daily living (ADLs) may be causing supra-physiological loads and therefore instigating long-term tissue damage. Post-operatively, it may be hoped that loading during ADLs is restored to levels within the 'zone of homeostasis', hence preserving the joint.

2.1.2 Joint Mechanics and Biological Changes

There are a great number of studies that support the relationship between joint mechanics and biological changes in joint tissues. Some examples relating to underloading are Behrens *et al.* (1989) which found changes in articular cartilage synthesis in joints which were immobilised, and Vanwanseele *et al.* (2003) which found in a longitudinal analysis that patients with spinal cord injuries had a higher rate of cartilage thinning than that observed in OA. The largest portion of this field, however, is in relation to the pathological mechanisms of cartilage destruction due to injury or repeated overloading: for summaries of this literature, the reader is directed to

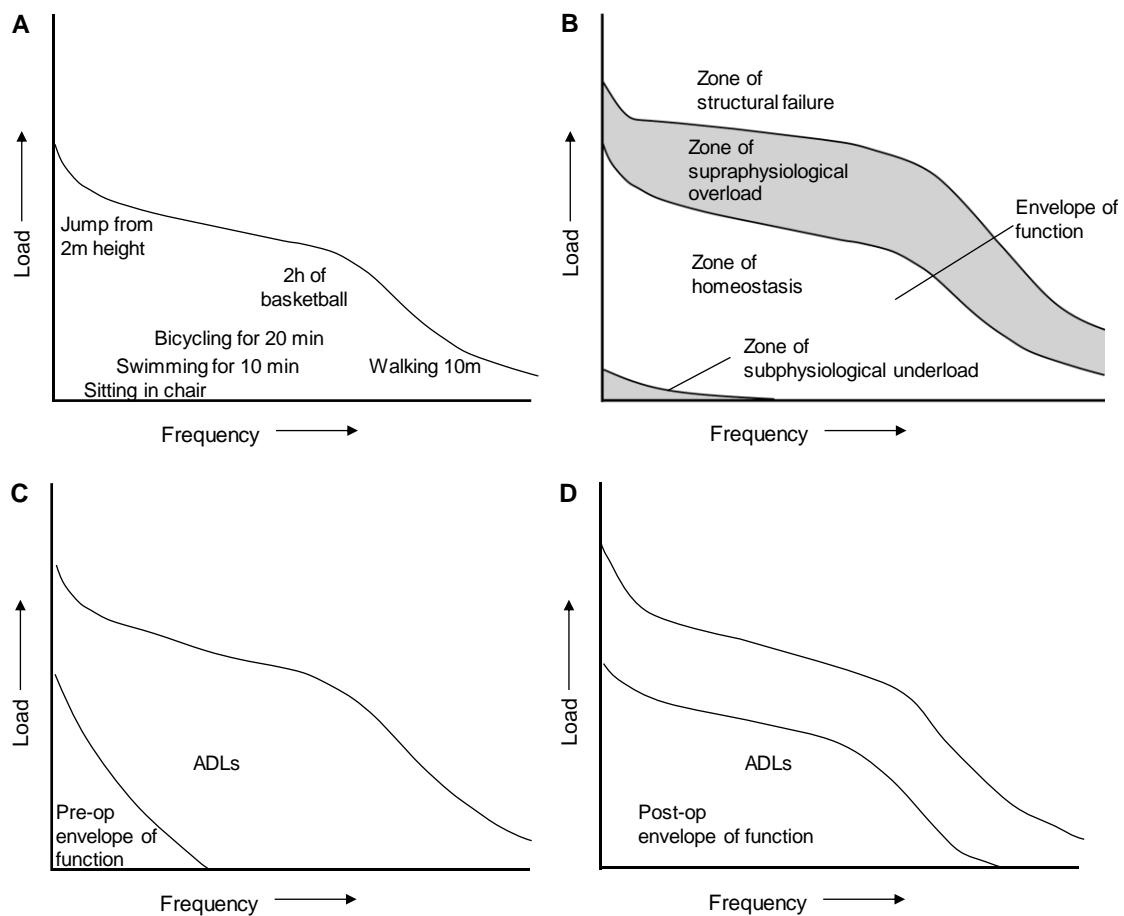


Figure 2.2 Dye's proposed envelope of knee function (Dye, 1996). A) The load-frequency relationship of some common activities. All come below the envelope of function, apart from the 3m Jump. B) The proposed envelope of function, displaying zones of underload, homeostasis, overload and structural failure. C) The potential effect of joint pathologies such as osteoarthritis on the envelope of function. Common Activities of Daily Living (ADLS) the zone of supraphysiological overload. The zone of homeostasis is much smaller. D) The proposed restoration of the envelope of function following an intervention. Common ADLS now fall within the zone of homeostasis

Buckwalter *et al.* (2013) and Kurz *et al.* (2005). There is strong evidence to support a causal link between joint loading and OA.

The progression of OA also leads to changes in loading of both the affected and unaffected joints. It is intuitive to suggest that someone with a pain or instability in a joint may move differently to compensate for this, or that they may avoid this activity entirely. It is also intuitive to suggest that this may affect their quality of life to some degree. What is much harder to ascertain is the effect this abnormal movement or activity avoidance

has on OA progression. When studying the way in which OA affects the way someone moves, it can be difficult to distinguish whether the abnormal loading is a cause of the OA, an effect of the symptoms, or a combination of both.

A good example of this is the well-cited study by Sharma *et al.* (1998) which found a significant positive correlation between OA disease severity and the peak External Knee Adduction Moment (EKAM) during gait. The EKAM occurs as the ground reaction force passes medial to the knee joint centre (see Figure 2.3), and is frequently used as a surrogate measure of contact forces within the media compartment of the tibiofemoral joint in order to assess the load reducing effects of orthopaedic interventions (Kutzner *et al.*, 2013). An equally well-cited study by Hurwitz *et al.* (2002) found that the EKAM during gait was much more closely correlated to static malalignment than disease severity. A systematic review of the relationship between malalignment and the development and progression of OA suggests that knee malalignment is both a risk factor for OA progression, and that malalignment can be caused and further increased by knee OA due to loss of cartilage and bone height (Tanamas *et al.*, 2009).

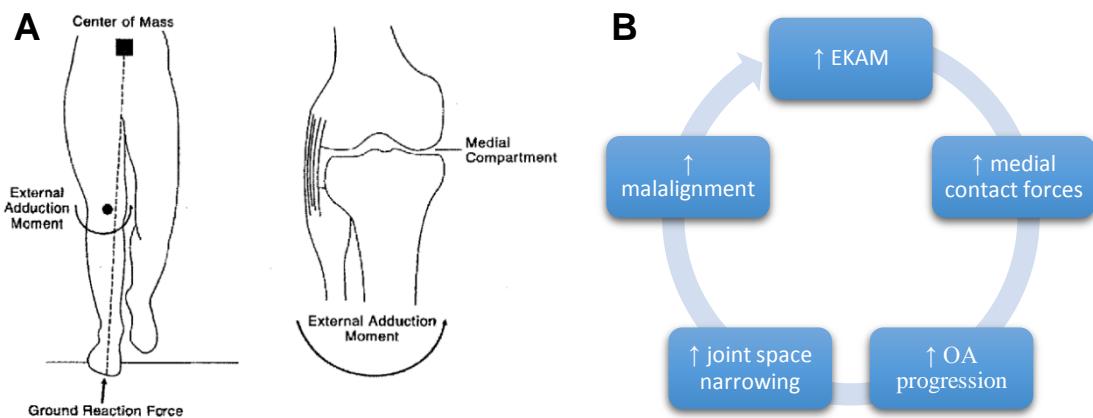


Figure 2.3 A) A simplified illustration of the calculation of the EKAM during gait. The ground reaction force passes medially to the centre of the knee, causing a frontal plane moment about the knee. This moment acts anticlockwise at the tibia, potentially causing increased contact forces (reprinted from Sharma *et al.* (1999))

B) A depiction of the potential cause and effect relationship between the EKAM and OA and knee malalignment.

Summary - OA is a complex heterogeneous disease process which is the result of mechanical and molecular events. This results in a cascade of further changes at a mechanical and molecular level, which makes it difficult to directly identify causal relationships. Kraus *et al.* (2015) calls for a new taxonomy for the classification of OA which includes all primary elements of the disease to arrive at a composite score. This could be useful for both the detection of OA and for monitoring the effectiveness of interventions, however, it is not clear how these multifactorial elements to the disease would be weighted.

2.1.3 Conservative Management of Knee OA.

Members of the Osteoarthritis Research Society International (OARSI) have, on multiple occasions, reviewed the evidence-base of conservative management of knee OA to provide expert consensus guidelines (Zhang *et al.*, 2007, Zhang *et al.*, 2008, Zhang *et al.*, 2010, McAlindon *et al.*, 2014). The summary of the current version of these guidelines is shown in Figure 2.4. There are a core set of treatments which are deemed appropriate for all individuals with knee OA, which together form a rather holistic approach to improving outcomes. The emphasis on weight management, exercise, strengthening, self-management and education all appear to be echoed by other recent guidelines such as those of the American Academy of Orthopaedic Surgeons (AAOS) (Brown, 2013), the European League Against Rheumatism (EULAR) (Fernandes *et al.*, 2013), and the American College of Rheumatology (ACR) (Hochberg *et al.*, 2012).

Core treatments - The supporting evidence behind these recommendations are beyond the scope of this thesis, however, outcomes of treatment modalities were commonly measured using self-reported measures of pain, function, physical activity, and general well-being. The AAOS guidelines are non-specific in their recommendations regarding strengthening and exercise. This may be due to the heterogeneity of regimes prescribed across the multiple research studies under consideration (e.g. content, duration, frequency) and the absence of a clearly superior regime. This again follows the trend within other guidelines. The EULAR group reached a consensus that mixed programs

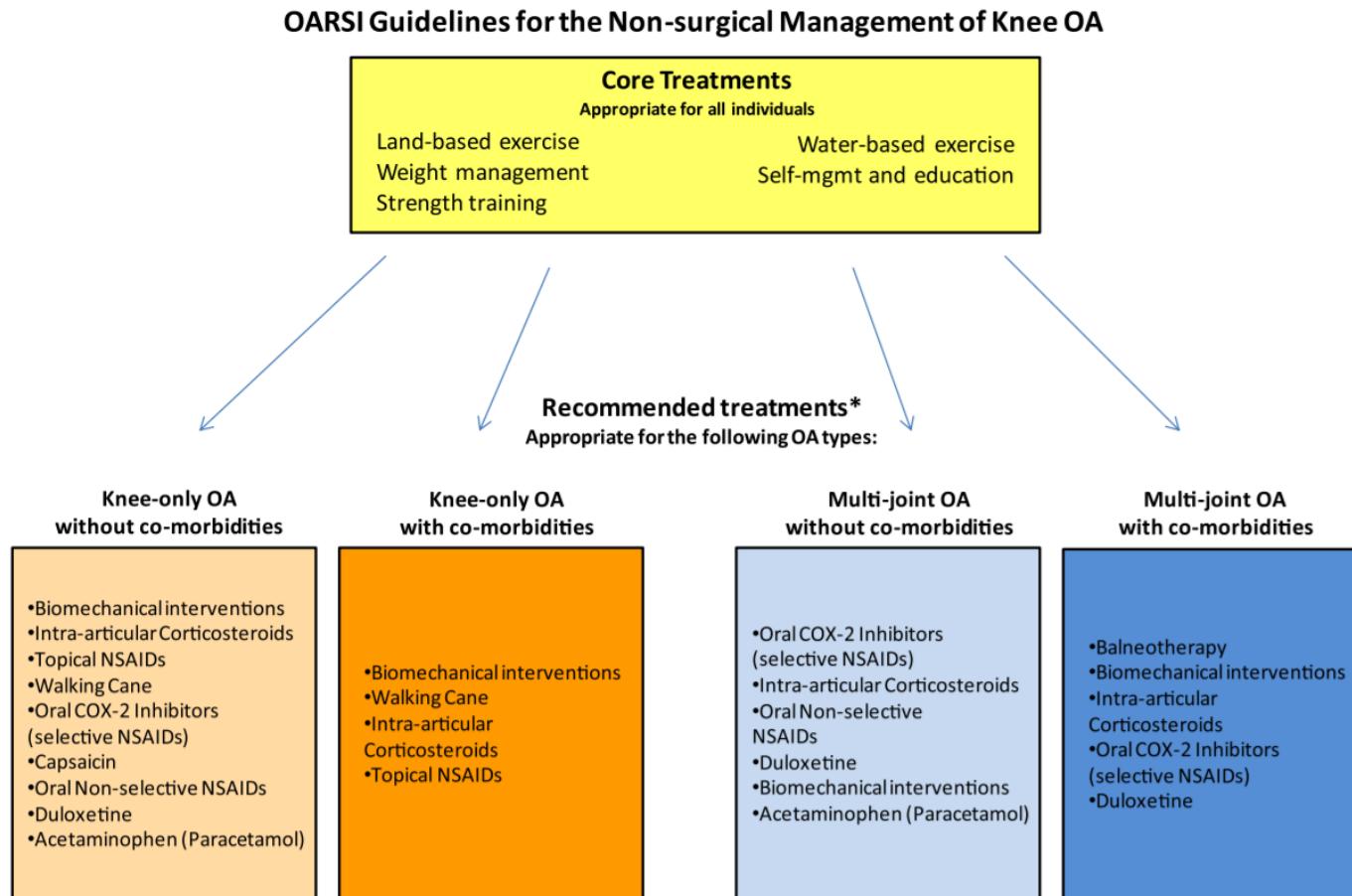


Figure 2.4 The summary of the expert consensus OARSI guidelines for the non-surgical management of knee OA., reprinted from McAlindon *et al.* (2014).

should be recommended with a focus on improving muscle strength, aerobic capacity, and joint range of motion (Fernandes *et al.*, 2013). While a mixed exercise program may appear an intuitive endorsement in lieu of evidence-based recommendations of specific targeted interventions (e.g. quadriceps strengthening), of the six mixed programs included in the review of Escalante *et al.* (2010), only one group achieved a significant reduction in pain. The authors highlighted the need for additional evidence to support mixed exercise programs for conservatively managing knee OA.

Biomechanical interventions - It can be seen in Figure 2.4 that biomechanical interventions are recommended for treating knee OA, irrespective the four identified sub-classifications. The quality of evidence for these recommendations were rated as 'fair', and were supported by a systematic review of the efficacy of knee braces and foot orthoses (Raja and Dewan, 2011), alongside randomly controlled trials assessing insoles (Bennell *et al.*, 2011, van Raaij *et al.*, 2010), knee braces (van Raaij *et al.*, 2010), and a variable-stiffness shoe (Erhart *et al.*, 2010). Both knee braces and foot orthoses are intended to offload one of the compartments of the knee (Raja and Dewan, 2011), and therefore may be more suitable for patients with OA only affected one compartment. Inserted insoles attempt to achieve this through changing the mechanical alignment of the calcaneus, and hence altering the mechanical alignment of the lower leg (Toda *et al.*, 2001). Several studies have demonstrated a reduction in the peak knee adduction moment of around 6% when using a lateral wedge of 5° (Kerrigan *et al.*, 2002, Kakihana *et al.*, 2005, Shimada *et al.*, 2006). The mechanism of action might also be attributed to the lateral shift of the centre of pressure relative to the foot (Hinman *et al.*, 2012). As opposed to reducing the external knee adduction moment, valgus knee bracing aims to reduce compression within the medial compartment by applying an external valgus moment, which counteracts the effects of the varus moment (Raja and Dewan, 2011). The National Institute for Health and Care Excellence (NICE) guidelines which inform clinical practice within the UK currently recommend that people with OA alongside

'biomechanical joint pain or instability' should be considered for assessment for insoles, joint supports or braces (NICE, 2014).

Pharmacological interventions - The recommended pharmacological interventions are focussed on the management of pain, with the ACR guidelines recommending the use of acetaminophen, oral or topical nonsteroidal anti-inflammatory (NSAID) drug, or tramadol for patients unable to obtain pain relief from over the counter equivalents (Hochberg *et al.*, 2012). Both the ACR and OARSI guidelines also advocate the use of intraarticular corticosteroids to relieve pain in knee OA patients, however, the AAOS guidelines deemed the evidence inconclusive, and hence clinical judgement should be exercised. There are numerous contraindications for the use of analgesics relating to comorbidities such as a history of gastrointestinal bleeding, arterial disease, hypertension, etc. On this grounds, the NICE guidelines have highlighted a need for more research into the long-term outcomes of treatments for OA in the elderly, for whom NSAIDs are often not appropriate (NICE, 2014).

2.2 Total Knee Replacement

Knee arthroplasty is a common procedure for patients with moderate to late-stage OA. The procedure involves the replacement of joint surfaces with orthopaedic prostheses which are specifically designed to restore functional movement to the joint. According to the latest report from the UK National Joint Registry (UK-NJR), over 103,000 replacement procedures were recorded in the UK in 2014 – an increase of 12.4% from the previous year (UK-NJR, 2015). Of all 772,674 knee replacements recorded in the database, 96% were specifically due solely to a diagnosis of knee OA (UK-NJR, 2015).

The knee is made up of three compartments: the medial and lateral tibiofemoral, and the patellofemoral. The choice of whether to replace one, two or all three compartments is dependent on expert opinion and the quality of the joint surfaces. The severity of OA within the medial and tibiofemoral compartment is most frequently classified radiographically using the Kellgren-Lawrence (KL) scale (Emrani *et al.*, 2008). If either solely the medial or lateral tibiofemoral compartment is to be replaced, this is considered a partial or Unicondular Knee Replacement (UKR). If, however, both the medial and lateral compartments are replaced, this is considered a Total Knee Replacement (TKR). The evidence as to whether to also resurface the patella during TKR surgery remains controversial, with a recent meta-analysis concluding patella resurfacing may be associated with better follow-up after five years, however, more evidence was required to further prove this (Chen *et al.*, 2013).

2.2.1 Choice of TKR Design Within the UK

Over 90% of knee arthroplasties performed in the UK are TKRs – a proportion which has shown no signs of shifting over the last ten years (UK-NJR, 2015). The percentage of these which have used implants designed to be fixated using bone cement has steadily increased over this period, to 97% in 2014. This is likely due to much higher costs, high rates of early loosening in early designs (Berry *et al.*, 1993), and lack of evidence for long-term clinical benefits of uncemented implants (Matassi *et al.*, 2013).

In 2014, of the cemented implants used, 71.6% were designs which retain the posterior cruciate ligament (PCL), the others being designs requiring the resection of the PCL. The clear majority (88%) of the latter were posterior-stabilised designs, which compensate for the absent PCL by introducing an intercondylar post and cam, which guide the knee through flexion resulting in an increase in anterior-posterior stability. A recent meta-analysis concluded that cruciate-retaining and posterior-stabilised TKRs have similar clinical outcomes (Li *et al.* (2014), and hence choice may be due to the preference of the surgeon or the pre-operative condition of the PCL.

Of these cruciate-retaining and posterior-stabilised designs, 97% were fixed-bearing; meaning that the polyethylene tray is fixed in the tibial baseplate. The remaining are mobile bearing designs, which allow a small amount of motion of the polyethylene component relative to the tibia. The primary proposed advantage of this is reduced aseptic loosening and wear of the polyethylene insert. However, a recent systematic review of 41 studies concluded there were no clinically relevant improvements in outcome (Van der Voort *et al.*, 2013).

In summary, the majority of TKRs in the UK appear to use one of two primary design types: cemented, PCL retaining, fixed-bearing implants (67%), or cemented, posterior sacrificing, fixed-bearing implants (23%). The aforementioned review, however, appears to find no significant differences in clinical outcomes between designs.

2.2.2 TKR Outcomes

Despite advancements in design and surgical technique, and the apparent consistency of clinical outcomes between designs, it is commonly reported that around one in five subjects are dissatisfied with their outcome (Baker *et al.*, 2007, Bourne *et al.*, 2010); in comparison to closer to one in 15 in hip replacement recipients (Anakwe *et al.*, 2011). Several studies have assessed the primary factors of this dissatisfaction, which are summarised below:

Post-operative pain – Numerous studies report post-operative pain, or level of pain relief from surgery, to be one of the main influences of dissatisfaction following TKR surgery

(Scott *et al.*, 2010, Baker *et al.*, 2007, Hamilton *et al.*, 2013). Interestingly, patients often expect greater reductions in pain when compared to improvements in function (Mahomed *et al.*, 2002).

Improvement in function – While generally not as strong a predictor as pain, functional improvement has shown to correlate with satisfaction in some (Baker *et al.*, 2007, Noble *et al.*, 2006), although not all studies (Scott *et al.*, 2010). As biomechanical function is limited by pain, they are intrinsically linked and it is therefore very difficult to analyse function as an independent variable, particularly when relying on Patient-Reported Outcome Measures (PROMs).

Expectations not met - It is not surprising that, in the majority of dissatisfied patients, pre-operative expectations were not met (Noble *et al.*, 2006), and / or patients had higher pre-operative expectations (Baker *et al.*, 2007, Scott *et al.*, 2010, Gandhi *et al.*, 2008). Interestingly, some studies have found as many as 50% of dissatisfied patients appear to have no specific adverse symptoms from their knee (Noble *et al.*, 2006, Kim *et al.*, 2009). This could indicate that a large proportion of dissatisfied patients may have had unrealistic pre-operative expectations, or it may indicate that PROMs and clinical assessment may not be detecting or representing the pathological symptoms the patient is experiencing.

Quality of care – The quality of care and overall experience within the hospital has been shown to have a role in patient satisfaction (Hamilton *et al.*, 2013, Scott *et al.*, 2010). This highlights an issue when using satisfaction as an outcome measure, as the quality

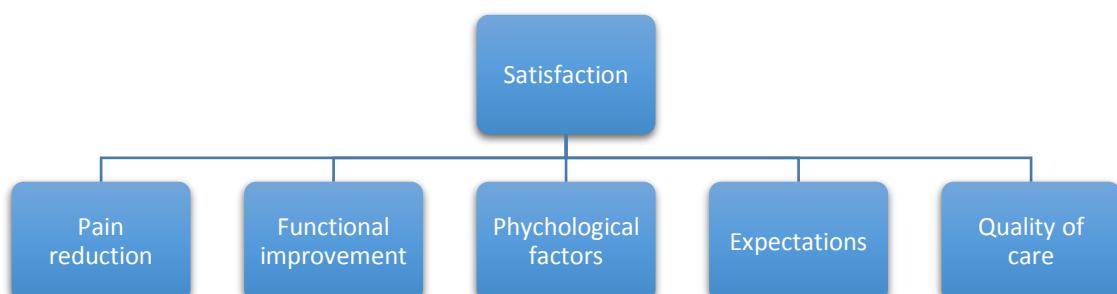


Figure 2.5 The primary categories of factors which appear to correlate to post-operative improvement following TKR surgery

of care may be highly variable, even within a single hospital, and is a factor the orthopaedic surgeon may have limited control over.

Psychological factors – Conditions such as depression, poor mental health, and a pessimistic explanatory style are positively correlated to dissatisfaction following TKR (Scott *et al.*, 2010, Gandhi *et al.*, 2008, Singh *et al.*, 2010). Depression is known to affect the experience of pain and perception of ability (Scott *et al.*, 2010), which again highlights a challenge when analysing PROMs.

2.2.3 Outcome Measures

Patient satisfaction is a common outcome measure for any intervention. However, as discussed above, it is the cumulative effect of several known and unknown factors. The reviews mentioned in Section 2.2.1 determined and compared clinical outcomes of prostheses mainly using:

- Clinical outcomes, such as post-operative complications (Li *et al.*, 2014), range of motion (ROM) (Li *et al.*, 2014), and radiological evaluation (Van der Voort *et al.*, 2013)
- Revision rates at long-term follow-up (Li *et al.*, 2014, Van der Voort *et al.*, 2013, Matassi *et al.*, 2013)
- Patient-reported outcome measures, such as The Knee Society pain score (Li *et al.*, 2014) and The Western Ontario and McMaster Universities Arthritis Index (WOMAC) score (Van der Voort *et al.*, 2013).

The UK-NJR also heavily reports revision rates in order to compare outcomes for TKR designs and patient demographics (UK-NJR, 2015). There are several challenges to using revision rates as an outcome measure for guiding surgical technique, patient selection, and rehabilitation. The revision rates of UKRs is much higher than that of TKRs in the UK (UK-NJR, 2015). However, a recent study suggests it is a particularly poor outcome measure when comparing these two surgeries due to differing patient indications and a completely different surgical decision process (Goodfellow *et al.*, 2010). The same argument applies, for example, when comparing different PCL retaining and

sacrificing designs which have differing patient indications, or comparing revision rates in older patients of whom surgeons might be less willing to revise due to functional expectations and surgical complications. It appears that improvements in patient-reported outcome measures are reflected in revision rates, particularly in younger patients (Price *et al.*, 2010). To assess revision rates, a long-term follow-up is required, which also adds significant practical challenge.

The use of PROMs can reflect how a patient perceives elements of their physical function before and after TKR surgery. There is, however, growing evidence that this often isn't reflected in objective measurements of functional performance (Maly *et al.*, 2006). In fact, it seems that patients report their improvements in physical function to be higher than it seems during objective assessments (Stratford and Kennedy, 2006, Worsley, 2011, Naili *et al.*, 2016).

Stratford and Kennedy (2006) investigated how patient-reported function was related to objective measures of function following TKR surgery and how this correlated with pain. Pain and function were assessed using the WOMAC questionnaire. Functional ability was objectively assessed using the following timed tests: 40m walk, ten step stair ascent/descent, sit-to-stand from a chair, and distance travelled during a six-minute walk. The researchers discovered that a reduction in pain was the primary predictor of the subject's perceived functional improvement, as opposed to objective functional performance measures.

Functional performance measures for OA subjects often involve the use of multiple limbs, as single-limb support can be too challenging. Mizner and Snyder-Mackler (2005) found that the quadriceps strength was strongly related to objective functional performance in TKR subjects, however, this relationship was stronger in the uninvolved limb. Functional performance tests which use the time taken or distance travelled as the primary outcome measure are likely strongly influenced by the other limb. OA and subsequent TKR surgery affect the biomechanics of not only the affected limb but also the unaffected joints, including those on the contralateral leg (Metcalfe *et al.*, 2013, Watling, 2014).

2.2.4 Rehabilitative Factors

In 2003, the National Institute of Health Consensus panel reported that the use of rehabilitation services before and after TKR surgery was perhaps the most understudied aspect of their care (Rankin *et al.*, 2004). The report acknowledged that there was strong theoretical justification that short and long-term outcomes would be improved through the treatment of preoperative and post-operative impairments e.g. joint contractures, abnormal movement patterns and joint mechanics, muscle weakness, and atrophy.

A more recent systematic review and meta-analysis of the effectiveness of physiotherapy exercise following TKR surgery compared 18 randomised control trials including a total of 1739 patients (Artz *et al.*, 2015). The study found evidence to suggest patients receiving physiotherapy exercise had improved physical function and reduced pain at 3-4 months in comparison to those receiving minimal physiotherapy. Benefits at 6 months were inconsistent, and primarily observed in the studies which were rated higher quality. However, no differences in pain and function were observed between outpatient physiotherapy and home-based exercise provision. The authors concluded that evidence was insufficient on long-term benefits of post-operative rehabilitation were limited and that further research is needed.

Pre-operative physiotherapy, sometimes termed ‘prehabilitation’, aims to increase the functional capacity of the patient before undergoing surgery. This is hoped to increase the patient’s ability to withstand the immediate effects of the surgery itself, as well as the post-operative rehabilitation phase (Ditmyer *et al.*, 2002). Some studies have supported a link between pre-operative functional ability and strength, and post-operative outcome (Dennis *et al.*, 2007) (Jordan *et al.*, 2014), however both the systematic reviews of Ackerman and Bennell (2004) and, more recently, Jordan *et al.* (2014) concluded that there is not enough evidence to support the effectiveness of pre-operative treatment by a physiotherapist.

2.3 Human Motion Analysis

Human Motion Analysis (HMA) is a technique which involves the objective quantification of human motion, including joint kinematics (e.g. joint angles), joint kinetics (e.g. external moments), temporal-spatial parameters (e.g. stride length), and muscle activity. This allows for a much more thorough objective quantification of how function has changed at both the unaffected and affected joints following surgical intervention.

There are various techniques for HMA, which all vary in terms of accuracy, precision, practicality and cost. The most common clinical application for HMA has been in the management of patients with walking disorders, causing gait analysis to become a routine part of patient management in certain centres.

Motion capture using opto-electronic stereophotogrammetry (MOCAP) is the most common method for quantifying both the kinematics and the kinetics (Fernandez *et al.*, 2008). This method has previously been used at Cardiff University to quantify the function of OA subjects (Jones *et al.*, 2006, Beynon *et al.*, 2006, Metcalfe *et al.*, 2013) and assess their post-surgical recovery (Jones *et al.*, 2006, Jones and Holt, 2008, Watling, 2014, Whatling, 2009).

Joint kinematics are assessed during MOCAP by using markers which are tracked in 3D space by cameras. The Qualisys system at Cardiff University uses retroreflective markers, which reflect infrared light (IR) emitted by the cameras. Within each camera, there is also an IR sensor which captures this reflected light. If the motion analysis laboratory is free from other sources of IR light, then the cameras will only see the markers, hence the complex object classification algorithms seen in HMA within the computer vision field are not necessary. There will, however, also be some level of unwanted IR light sources and reflections within a laboratory. This is easily addressed using preventative methods, camera masking, or pixel intensity thresholds.

The actual movement of bones relative to one another cannot be directly measured *in vivo* using markers. Instead, anatomical landmarks are palpated and used to estimate

the position and orientation of the underlying bone. This provides clear, repeatable and clinically interpretable axis definitions to the segments. For example, the distal end of the tibia is often defined using the medial and lateral malleolus, and the proximal end using the femoral epicondyles (see Figure 2.6A).

As a person moves, the soft tissues are continually moving relative to the bone due to skin movement, muscle contraction, and inertial effects. This results in inaccuracies in the assumption that marker movement directly corresponds to bone movement. The anatomical landmarks used to define the segment axis system also happen to be prone to large levels of soft tissue artefact (STA) during motion. It is, therefore, common to use tracking markers, which are placed on the subject at locations with less STA, such as

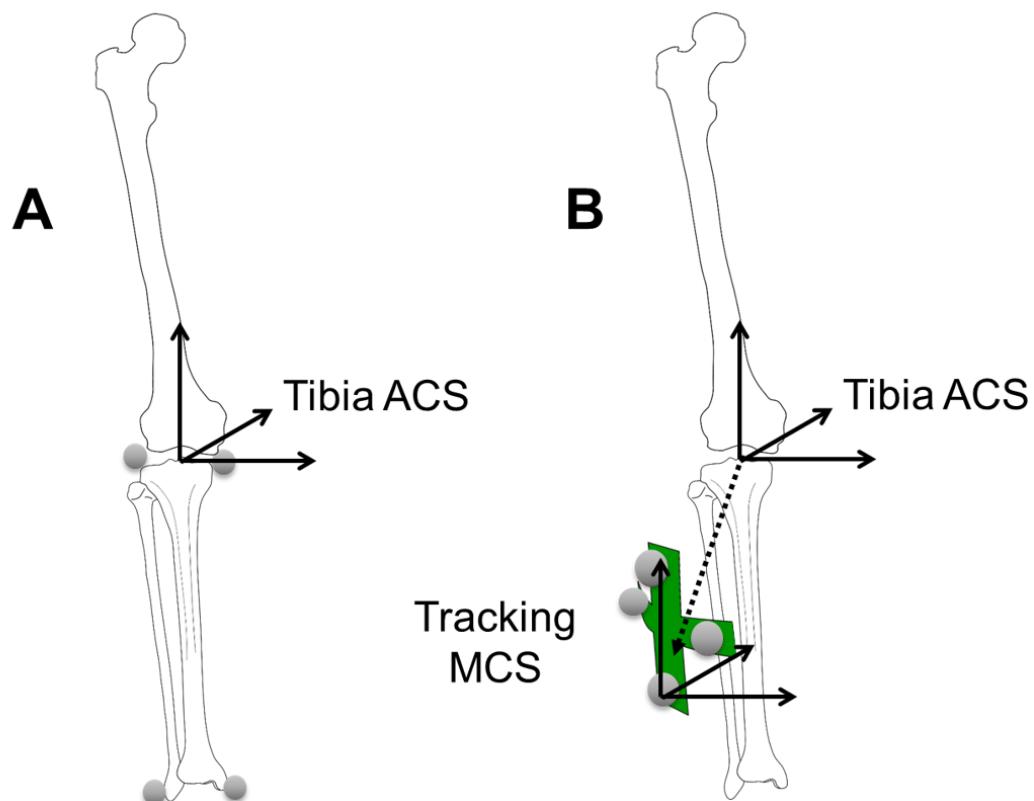


Figure 2.6 A) Illustration of how markers (grey circles) on the femoral epicondyles and the medial and lateral malleolus can be used to define an Anatomical Coordinate System (ACS) for the tibia during a static trial

B) Illustration of how the position of a rigid tracking cluster placed laterally on the tibia might be used to reduce errors due to soft tissue artefact. The tracking Marker Coordinate System (MCS) is defined relative to the tibial ACS during the static trial. During motion trials, only the position and orientation of the tracking MCS need to be collected, and the position and orientation of the tibial ACS can then be inferred.

the lateral shank and thigh (see Figure 2.6B). Generally, at least three tracking markers will be used per anatomical segment, which allows the creation of a tracking segment. The rotation of the tracking segments relative to one another does not produce a clinically interpretable joint angle. The position and orientation of each tracking segment relative to the corresponding anatomical segment is recorded during a static calibration trial. During the movement trials, it is assumed that the position and orientation between the tracking segment relative to the true anatomical segment axis remain constant, and anatomical segment orientation can, therefore, be inferred solely through measuring tracking marker segments.

For a thorough overview of the possible errors incurred during MOCAP, the reader is redirected to a comprehensive four-part review (Cappozzo *et al.*, 2005, Chiari *et al.*, 2005, Leardini *et al.*, 2005, Della Croce *et al.*, 2005). In summary, as the technology involved in MOCAP has advanced, the methodological errors have quickly far outweighed the instrumental errors. The primary methodological errors are STA, as previously mentioned, and the failure to accurately model the anatomical axis of the bone using anatomical markers. The latter can be due factors such as marker placement error, high amounts of subcutaneous fat due over bony landmarks, or that elements of anatomic axes, such as the hip joint centre, cannot be palpated. STA is particularly high for the thigh and can result in rotational errors greater than 12 degrees in calculations of internal/external rotation and ab/adduction of the knee (Peters *et al.*, 2010, Garling *et al.*, 2007).

Kinetic data is calculated using a force plate/platform. These plates measure the equal and opposite Ground Reaction Force (GRF) caused by the foot in contact with the floor during motion. The human body is being modelled as a system of rigid links with six degrees of freedom at each joint (unless inverse kinematics are being applied). A free body diagram can be described to estimate the reaction forces and moments that act about these links. In addition to the consideration of the GRF, the mass and inertia of the body itself contributes to reaction forces and external joint moments. To estimate these

effects, the inertial properties of body segments can be estimated during inverse dynamic analysis. This involves the use of cadaveric data, such as that provided by De Leva (1996) which provides linear regressions of the centre of mass and the radius of gyration in each plane for segments relative to parameters, such as leg length.

2.4 Data Reduction

The collection of HMA data results in an extensive amount of temporal information. Generally, gait variables are normalised using 101 data points to a percentage of stance phase or the entire gait cycle. To allow a meaningful statistical analysis to be performed, these temporal waveforms must be summarised using a smaller number of discrete variables. This has resulted in an extensive application of data reduction techniques to HMA data (Chau, 2001a). A common method of reducing data is to define discrete parameters of the waveform, as shown in Figure 2.7. For example, during the swing phase of gait, the knee must flex to achieve toe clearance as the limb progresses forward. A reduction in this angle might be related to an indication of an increased risk of trips or falls.

Choosing which discrete parameter to calculate, however, is subjective and may be discarding valuable information. While consistent peaks and troughs may be identifiable in healthy subjects, often the waveforms of pathological subjects will have completely different characteristics. Furthermore, by completely discarding the rest of the waveform, important information regarding inter-subject variability can be lost (Gaudreault *et al.*, 2011).

Deluzio *et al.* (1999) demonstrated that PCA was a useful technique in the reduction of temporal biomechanical data. The study found that principal component scores were sensitive to gait changes associated with knee OA, as well as changes following a partial knee replacement. PCA has since been successfully applied at Cardiff University to help distinguish between OA and non-pathological (NP) subjects and hence objectively measure changes in gait parameters following TKR surgery (Jones *et al.*, 2008, Whatling *et al.*, 2008, Whatling, 2009, Metcalfe *et al.*, 2013).

Principal Component Analysis is a multivariate data analysis technique which applies an orthogonal transformation of an n dimension dataset of potentially correlated variables, in order to arrive at a new n dimension dataset of linearly uncorrelated variables. The

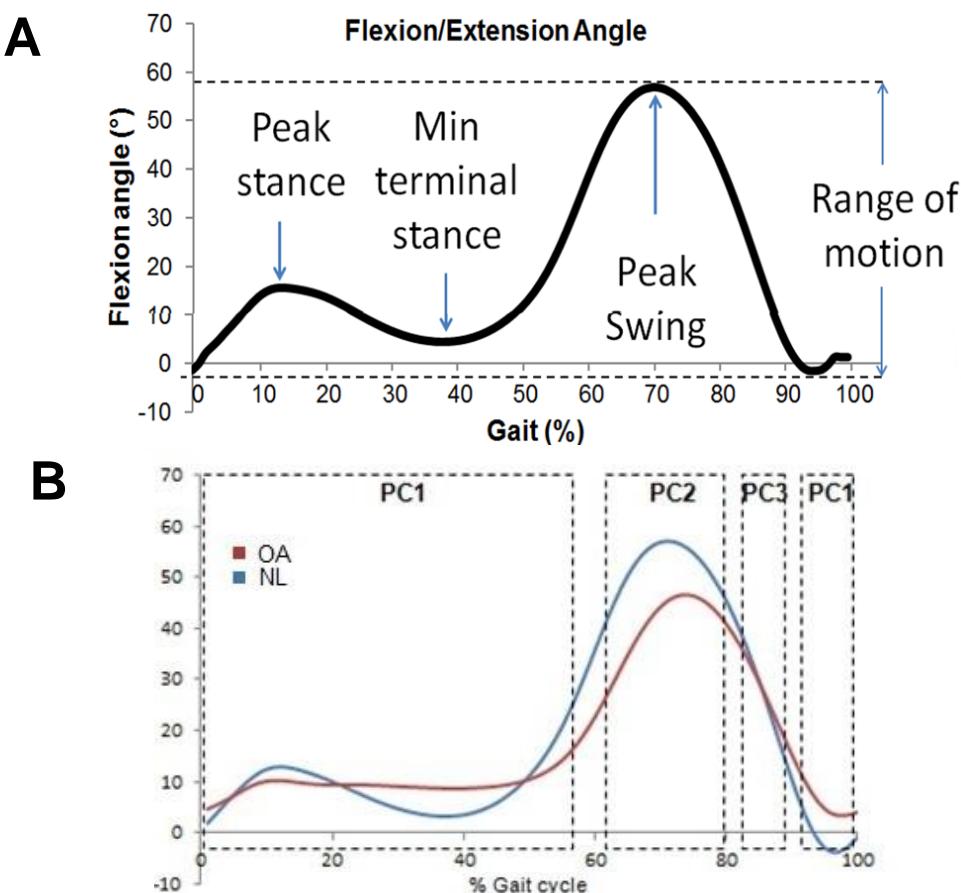


Figure 2.7 A) An example of how a knee flexion/extension angle during gait might be reduced into discrete parameters which can be easily interpreted.

B) An example of the how the results of principal component analysis (PCA) might be interpreted. The three principal components which represent the greatest total variance have been selected. The areas highlighted by the dashed lines represent the proportion of the gait cycle for which a principal component represents greater than 50% of the variance.

first dimension of the new dataset will represent the greatest amount of variance in the dataset, and so forth until the n th dimension, which will often end up representing an extremely small amount of the total variance. It then becomes possible to reduce the dimensionality of the dataset by only considering, for example, the first five dimensions.

2.4.1 Computing Principal Components

PCA is a relatively straightforward multivariate analysis technique. The steps are listed below but explained in much greater detail in Section 3.4.

1. Standardise the data – such that it has zero mean and a unit variance
2. Calculate the correlation coefficient matrix

3. Calculate the eigendecomposition of the correlation matrix to compute at the eigenvectors and eigenvalues
4. Multiply the eigenvectors by the square root of the eigenvalues to arrive at the factor loadings
5. Multiply the eigenvectors by the standardised data points to arrive at the principal component (PC) scores for each subject

If, for example, 101 data points have been used to normalise a gait waveform to 0-100% of the gait cycle, this method will calculate 101 eigenvectors, each with 101 dimensions. Each eigenvector will have a corresponding eigenvalue which represents how much of the total variance of the dataset that eigenvector represents; e.g. if the first eigenvalue was 0.78, and if we were then to reconstruct all the waveforms using just that first eigenvector/principal component, 78% of the initial variance between subjects would be represented.

The purpose in this instance was to reduce the dimensions of the dataset, and therefore not all 101 PCs will be retained. One objective criterion for PC selection is to use Kaisers rule (Kaiser, 1960). This rule suggests that all principal components with an eigenvalue of less than one should be discarded. Another reasonable technique is to define a target variance that would ideally be represented. The minimum number of PCs that are required to meet that threshold can then be used for further analysis.

A further potential selection technique is to use the factor loadings. The factor loadings can be thought of as the correlation coefficients of the new data. The correlation coefficient between two variables is often denoted as the r value, and the amount of the total variance that correlation represents is generally denoted as the r^2 value. If a correlation is greater than 0.71 or less than -0.71, its r^2 value is greater than 0.5 and it, therefore, represents greater than 50% of the variance. Each principal component has a factor loading for each point of the gait cycle, indicating how much of the total variance that principal component represents at that point of the gait cycle. Comrey and Lee

(2013) suggest that the >0.71 and <-0.71 range be used as a threshold for consideration of PCs.

2.4.2 Further Techniques

The technique of individually computing PCs for waveforms takes advantage of the high amounts of correlation that individual points of a single waveform have with each other. There is also a high amount of interdependency between individual gait variables and there are therefore potential advantages to performing PCA to all waveforms in a single 'state space'. For example, knee flexion is required during swing phase to achieve ground clearance, but hip circumduction, a combination of ab/adduction and internal/external rotation can be adopted as a gait compensatory strategy to achieve toe clearance. In subjects adopting this strategy, changes in knee flexion/extension waveforms would likely be highly correlated with changes in hip flexion/extension, ab/adduction, and internal/external rotation. If these waveforms were all considered in a single PCA, a large amount of this correlated variation might be representable using a single principal component.

The application of PCA to multiple joint angles within the same 'state space' has been reported in the literature (Boyer *et al.*, 2012). Other researchers have also employed a slightly different technique which applies PCA to the time normalised marker coordinate data. The marker coordinates of the pelvis are generally subtracted from the markers at each frame, such that markers' coordinates are represented as distances from the pelvis marker (von Tscharner *et al.*, 2013, Federolf *et al.*, 2013). These techniques often still include the whole GRF waveforms within the state space.

2.5 Classification / Data Summation

Interpretation of HMA data is challenging due to the complex independences of biomechanical variables (Schwartz and Rozumalski, 2008). Data reduction has previously been discussed, however, as dimensionality is reduced, data is also discarded. If for example, we wanted to describe changes in the biomechanics of the ankle, knee, and hip following TKR surgery, joint translations are going to be ignored due to STA, leaving three rotations and three moments about each joint – a total of 18 temporal waveforms. If we imagine that each waveform is reduced to two discrete parameters using either subjective parameterisation or PCA, there are now already 36 discrete variables for consideration. OA is a bilateral disease (Metcalfe *et al.*, 2013) and it is, therefore, important to also track changes of the other limb. It has also been shown that subjects with medial compartmental OA often adopt a lateral trunk lean – perhaps to reduce coronal plane loading of the affected joint (Hunt *et al.*, 2008). Trunk kinematics should, therefore, be considered. OA subjects also tend to compensate for decreased stability by co-contracting the surrounding muscles (Zeni *et al.*, 2010, Hortobágyi *et al.*, 2005), therefore, biomechanically, the movement may appear stable despite pathological co-contractions potentially increasing joint loads and leading to further cartilage loss (Fallah-Yakhiani *et al.*, 2012). Pathological subjects might consequently be incorrectly identified as being healthy or improved following an intervention without the inclusion of an electromyographic (EMG) analysis. This demonstrates the vast complexity of summarising important features of human biomechanics.

As discussed previously, OA is a complex and very heterogeneous disease with very different pathological routes to common clinical disease characteristics. Multiple researchers attempt to define subgroups or ‘phenotypes’ of OA patients based on clinical presentation (Knoop *et al.*, 2011, Van der Esch *et al.*, 2015), genetic factors (Van Meurs and Uitterlinden, 2012), patterns of cartilage degeneration (Snelling *et al.*, 2014) and risk factors (Bierma-Zeinstra and Verhagen, 2011). The goal of this research is to develop more targeted interventions which are driven by knowledge of specific OA

phenotypes. Changes in physical function are also very heterogeneous within OA subjects, as opposed to a linear progression of severity. For example, one subject's OA could be primarily due to a traumatic incident, such as a meniscal tear, another's primary due to obesity, and another's due to ageing. The first might present with a seemingly high level of function, yet chronic instability, the second might have severe challenges performing even low-load activities of daily living, and the third might struggle primarily with higher load activities, such as stair-climbing.

Due to this heterogeneity in the potential biomechanical adaptations in OA subject, vital information might be discarded if only a select few biomechanical variables were considered further. This also poses challenges due to conflicting and corroborating biomechanical data information. For example, a surgical intervention might have restored knee flexion moments towards that of a matched control subject, yet this may have resulted in changes in previously NP hip and ankle function. As mentioned earlier, Kraus *et al.* (2015) call for disease elements of OA to be summarised in a composite score to aid disease detection and monitoring following interventions, but this is clearly not a simple task for biomechanical information.

2.5.1 Gait Indexes

One simplistic solution to summarising gait data is to summate how far the subject deviates from that observed in healthy controls. One of the criticisms of this approach is that it ignores the inter-correlated nature of gait variables. An example which illustrates this has been modified from Schutte *et al.* (2000) and is displayed in Figure 2.9. From the figure, it can be seen that, when interpreting the distance from the standard deviation of highly correlated variables individually, this could result in the misinterpretation of gait data for subjects 2 and 4, who actually fall within the healthy hypothetical distribution of increasing peak flexion angle with increasing walking speed. Schutte *et al.* (2000) introduced a method called the 'normalcy' index, later becoming known as the Gillette Gait Index (GGI), which aimed to overcome this issue. The method is briefly summarised below and illustrated in Figure 2.8:

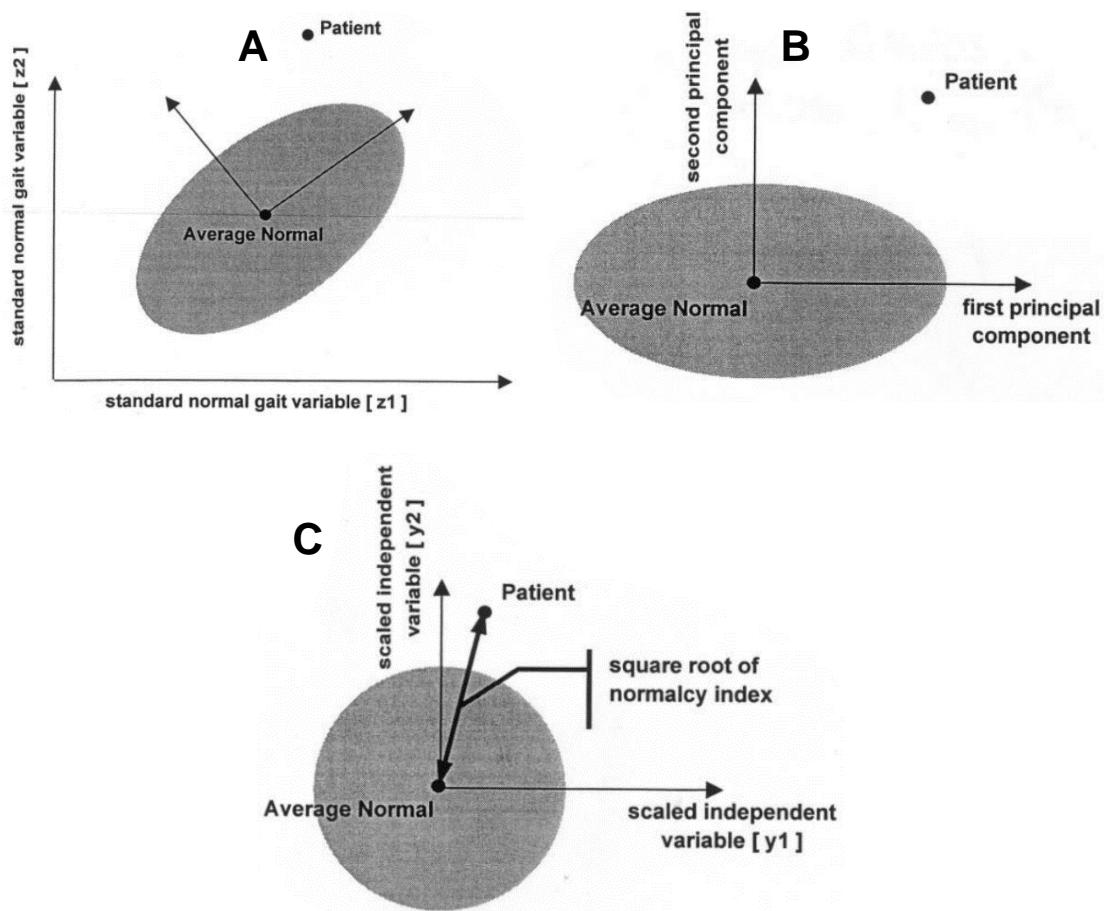


Figure 2.8 Graphic representation of the calculation of the normalcy index for two variables, reprinted from (Schutte *et al.*, 2000).

- The ellipse represents the normal healthy distribution of two correlated gait variables within two standard deviations (STDs) away from the mean. The patient is represented as a distance at an amount of STDs from these means. The primary axis of variation (principal component) is displayed (pointing top right) and the orthogonal axis to this is displayed (pointing top left).
- Principal component analysis is used to define a new axis system, and the data points have been projected onto this axis system by calculating their PC scores. The two variables are now uncorrelated; however, the variation is not equal.
- The two axis systems have been divided by the square root of their eigenvalues, such that the variance is now equal. The Euclidean length of the patient from the centre of this new coordinate system is now calculated and named the 'normalcy index'.

1. Principal component analysis is performed on the healthy data to define a new uncorrelated axis system, following the same steps listed in Section 2.4.1.
2. Project the data onto this new axis system, by multiplying the standardised values by the eigenvector (effectively calculating the PC scores)

3. Divide each axis (and variable value) by the square root of the eigenvalue associated with that PC. This essentially re-standardises the data such that each axis has equal variation.
4. Calculate the Euclidean length of the patient from the centre of this new coordinate system, which is named the ‘normalcy index’.

The most obvious criticism of this method is that, by dividing the PC axes by the square root of the eigenvalues, the contribution of PCs representing a high degree of variance is lessened, and those representing low variance is increased. The PCs which represent the smallest amounts of variance may be very sensitive to noise and have no contextual relevance, yet are being equally weighted in the normalcy index. This might explain why the normalcy index of a single subject can vary drastically between different sets of healthy control data (McMulkin and MacWilliams, 2008).

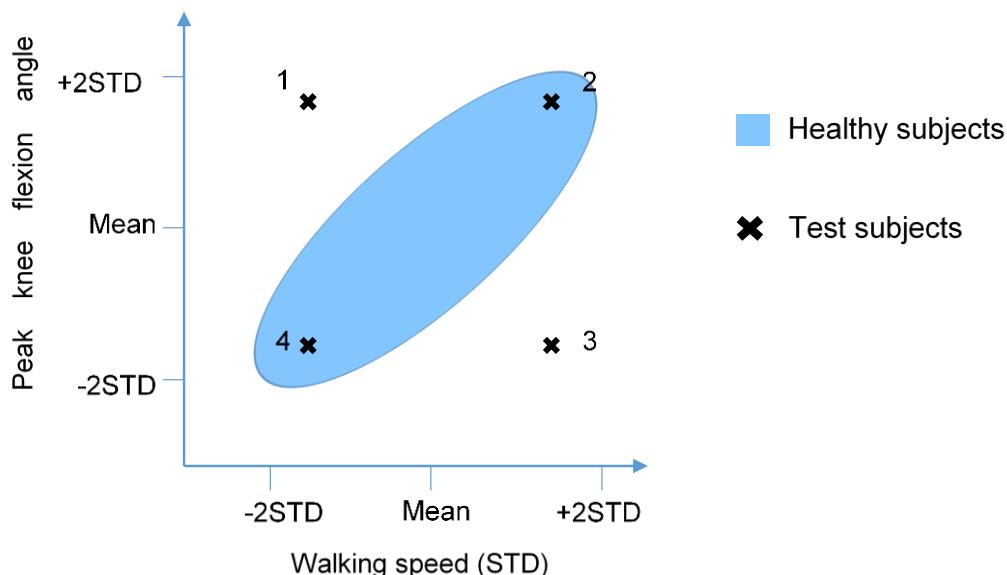


Figure 2.9 An illustration, adapted from Schutte *et al.* (2000), of how the interpretation of ‘normalcy’ of two gait variables can be misleading when there is high correlation between these variables within a healthy cohort.

The blue ellipse represents the normal healthy distribution of the peak knee flexion angle relative to walking speed within two Standard Deviations (STDs) away from the mean.

Hypothetical test subjects 1-4 have been highlighted as a black cross. While all four test subjects are all the same fraction of STDs away from the mean of both variables, subject 2 and 4 actually fall within the distribution found in healthy subjects, whereas subject 1 and 3 fall well outside this range.

In the demonstration, the normalcy index is calculated using 16 temporal-spatial and kinematic discrete variables which were subjectively selected based on the experience of the authors. Schutte *et al.* (2000) raise this issue themselves when presenting the technique, and even suggest the inclusion of kinetic parameters might more accurately reflect patient outcomes. Despite this, these same 16 variables have been used extensively by researchers to calculate the normalcy/GGI index (Assi *et al.*, 2009, Wren *et al.*, 2007, Hillman *et al.*, 2007, McMulin and MacWilliams, 2008). Also, the normalcy index seems to have been used primarily in the research field of cerebral palsy, hence children are often used as control subjects (Cretual *et al.*, 2010). It appears that the GGI may also be valid in adult populations (Cretual *et al.*, 2010), but it was again noted that these 16 variables may not be the optimal biomechanical descriptors of gait.

Further criticisms of the GGI include the difficulty of interpreting and the lack of physical meaning of the multivariate components which make up the score, and the non-normality of the index (Schwartz and Rozumalski, 2008). A new gait summary measure has been proposed by Schwartz and Rozumalski (2008) in an attempt to overcome these problems. The calculation involves the creation of a matrix, or state space, which includes all temporal gait parameters under consideration. A singular value decomposition (SVD) is then calculated on the matrix, a technique which has a lot of similarities to PCA. The SVD creates a new ‘orthogonal basis’ which can be used to reconstruct the data. Much like the PC selection process mentioned in Section 2.4.1, only some of the feature components will be considered for analysis. Also, each feature component accounts for a decreasing amount of variance to the previous. The analysis of Schwartz and Rozumalski (2008) included the use of threshold criteria for the minimum representation of total variance and a technique which measures the similarity of the reconstructed data to that of the original.

2.5.2 Artificial Intelligence

Artificial Intelligence (AI) has been applied to biomechanical data in order to classify healthy and pathological function for at least 25 years (Chau, 2001b). The goal of a

classification is to iteratively learn a mathematical relationship between input variables, e.g. discrete biomechanical variables and the target output, e.g. healthy or OA function.

Classification techniques can be broadly categorised as either supervised or unsupervised (Bishop, 1995). Both techniques will derive their relationships based on the training data. With supervised techniques, however, the training data has known class labels; for example, in this context, the data may have a class label of '0' if they are known to be healthy, and '1' if they are known to have OA.

Unsupervised classification techniques infer the classes based on the data itself (Bishop, 1995). For example, let's suppose there were two distinct compensation strategies that OA subjects use to avoid pain and instability. These might be distinct from healthy subjects, hence an unsupervised classification technique might arrive at three class labels i.e. 'healthy', 'OA1' and 'OA2'. While unsupervised classifier architectures have been applied to gait analysis, such as hidden Markov models (Cheng *et al.*, 2008) and self-organising maps (Barton *et al.*, 2006), this thesis is focussed on the supervised classification of labelled data. Further sub-classification or phenotyping of OA subjects using unsupervised techniques may be clinically informative, however, is beyond the scope of this thesis.

2.5.3 Supervised Training

Supervised classification architectures are the most common AI technique within gait analysis research (Lai *et al.*, 2009). This is likely because the training data can be initially labelled using expert knowledge, e.g. using a pre-existing accepted clinical classification of pathology. This preserves the clinical validity of the output class labels, however, inhibits the direct identification of new sub-groupings or phenotypes of a pathology.

The most prevalent of these techniques within gait studies is that of Artificial Neural Networks (ANNs) (Schöllhorn, 2004). ANNs are a type of machine-learning model which were inspired by the biological neural networks of the brain. In a broad sense, ANNs represent a series of interconnected "neurons" which make decisions to model the input data. These neurons are organised into different layers. The input layer is the one directly

connected to the input data, and the output layer is connected to the final output of the classification. The processing layers in between these two are referred to as “hidden” layers. Every neuron in the input layer is connected to every neuron within the hidden layer, which is likewise connected to every neuron in the output layer (see Figure 2.10). These connections are assigned weights, which are updated iteratively as the ANN is trained – most commonly using a backpropagation algorithm (Bishop, 1995).

A notable application of ANNs is the study of (Lafuente *et al.*, 1998), where kinematic, kinetic and temporal parameters during gait were passed to a three-layer backpropagation ANN to classify between four groups; NP, ankle OA, knee OA and hip OA. In total, 7% of the 148 pathological subjects were correctly classified, and 87% of the 88 NP subjects. A comparison was also made to the performance of a less complex, Bayesian classification technique, finding a 5% improvement in accuracy. The authors did, however, note that the interpretation of the classification rules was far more challenging due to the complexity of the relationships between input variables and an output within the three-layer network. The research study concluded at this time that additional effort was necessary to translate the inference model into explicit rules which can be readily understood and adopted.

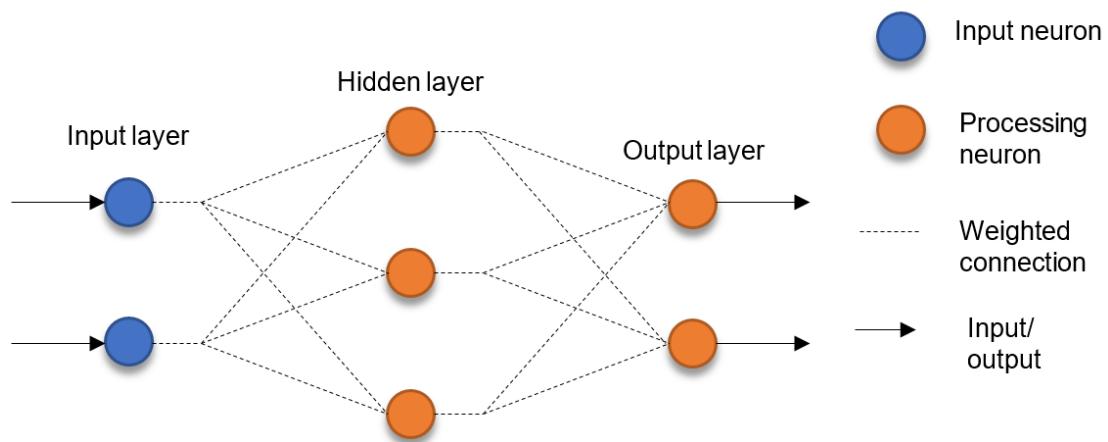


Figure 2.10 The layout of a very simple neural network consisting of only two inputs and outputs, and one hidden layer consisting of three neurons. Input neurons (blue) simply pass the data to the hidden layer. The processing neurons (orange) contain mathematical functions which sum the weighted inputs to the neuron, and then apply an activation function. More recent developments of AI within gait analysis research were well-reviewed by (Lai *et al.*, 2009). The author's concluded that despite high performances being demonstrated

by numerous researchers adopting various architectures, however, the uptake of AI techniques in applications such as rehabilitation monitoring, disorder detection, and risk management had been limited. This lack of adoption of techniques was attributed to the “black-box” nature of many AA techniques, where relationships between the model and the disorder are difficult to interpret or graphically visualise. It was also noted that supervised classifiers are limited by ignorance to pathologies not learnt by the AI system, and are further challenged by the heterogeneous nature of gait abnormalities associated with a particular disorder.

It appears that the theoretical basis of classification architectures such as ANNs is sound, there are limitations relating to clinical interpretability which inhibit easy adoption by clinicians. This is perhaps due to the iterative nature of AA architectures which contribute to the complexity of the model. Figure 2.11 shows a very simplistic model of a generic AI classifier. The relationship between the input data and classification output is usually dependant on some control parameters. Often, these initial values are arbitrarily chosen, as the system should theoretically converge to the same optimal solution (Bishop, 1995).

Within a supervised classifier, the classification outcome is compared to the original class labels of the input data, and the classification accuracy can be calculated. This

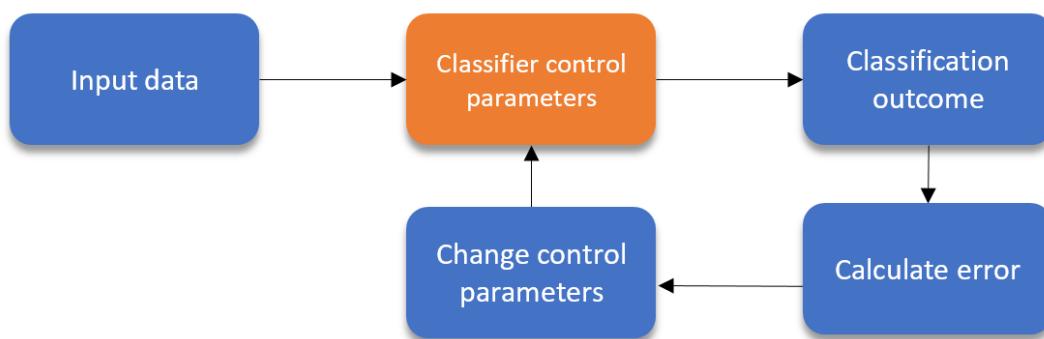


Figure 2.11 A very simple depiction of the learning process of a supervised classifier. The labelled training data makes up the input data, and defines the target classification outcome. The classification control parameters shown in orange defines the mathematical relationship between the inputs and outputs of the classifier. Here an error is calculated as the difference between the classification outcome and the ideal outcome. A small change is then made to the classification and the cycle repeats, until some target outcome has been reached.

represents just one type of objective function; often referred to as a ‘loss’ or ‘cost function’ within the machine-learning field (Alpaydin, 2014), which can be used to evaluate the performance of the classifier. The outcome of this objective function is then fed back through the network to affect how the control parameters are adjusted. The classification is then re-run with the new control parameters. This is then repeated until some target has been met, e.g. a classification accuracy of 95%. This is somewhat different to the application of the DST theory classifier introduced by Beynon et al. (2000) and Jones (2004), for which the control parameters are explicitly defined as opposed to optimised iteratively.

2.5.4 Dempster-Shafer Theory

The Dempster-Shafer Theory (DST) of evidence is founded on the work of Dempster (1968) and Shafer (1976). The technique enables decision making in the presence of ignorance (Safranek *et al.*, 1990), which is particularly pertinent when classifying OA function due to the aforementioned quantity of corroborating and conflicting evidence. The DST classification technique and its application to the classification of the gait biomechanics of OA subjects, and monitoring of functional recovery following TKR surgery has been described in detail by Beynon (2002) and Jones *et al.* (2006), both based at Cardiff University. From hereafter within this thesis, Jones *et al.* (2006)’s specific implementation of the DST classification will be referred to as simply “DST classifier”.

The three main stages of the classification technique, when applied to the classification of OA subjects, is summarised:

1. Input variables are converted into confidence factors.
2. Each confidence factor is converted into a degree of belief of either healthy, OA or uncertainty. This is referred to as the Body of Evidence (BOE).
3. The BOE for each variable is combined using Dempster’s rule of the combination of evidence. The Combined Body of Evidence (CBOE) represents the calculated

probability that each subject belongs to either the osteoarthritic or the healthy class, and also the uncertainty.

So, in other terms, the value of each input variable is converted mathematically into evidence within steps 1 and 2, and then this evidence is then summarised within step 3. The conversion of individual input variable into evidence has some similarities with that of neural networks.

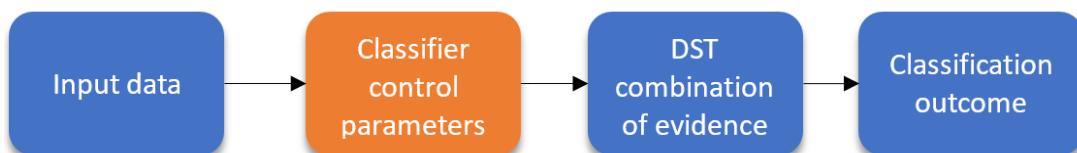


Figure 2.12 A simple depiction of the DST classification process, intended as direct comparison with Figure 2.11. The classification control parameters shown in orange are clearly defined by the input data, and not affected by classification outcome. There is therefore only one cycle, or ‘epoch’. The control parameters are therefore defined *a priori* to reaching a classification outcome.

Within step 1, input variables are converted into a body of evidence by using a sigmoid function. The sigmoid function is commonly used as an activation function within neural networks to summate the inputs of a neuron, and are also commonly referred to as logistic functions within other research fields.

The function can be defined in its simplest form as:

$$\text{sig}(x) = \frac{1}{1 + e^{-v}} \quad (2.1)$$

Where v is the input variable. Jones (2004) gave four examples of different sigmoid functions which illustrate how the standard equation above can be modified to convert a variable value into a ‘confidence factor’ (see Figure 2.13). It can be seen from this figure that two attributes of the sigmoid curve are being modified – the steepness (A vs B) and the value of v for which the confidence factor is equal to 0.5 (C vs D). It was argued that the steepness of the sigmoid curve should be related to how well the variable discriminates between the two groups, e.g. OA vs NP. If all OA subjects had a value of $v > 40$, and NP all were <10 , i.e. there was a very strong separation, the sigmoid curve of Figure 2.13B might be most appropriate. It has also been proposed that the midpoint

of the sigmoid curve should not introduce a bias towards either group (Beynon, 2005).

These two attributes of the sigmoid curve are altered by Jones (2004) using the following equation:

$$(cf) = \frac{1}{1 + e^{-k(v-\theta)}} \quad (2.2)$$

Where θ defines the value of x at which $cf=0.5$, k defines the steepness of the activation function, and v is the input variable value.

Jones explored two definitions of the value of k , both suggested by Beynon and Buchanan (2004). One definition suggests that k is the Pearson's correlation coefficient of the input variables and their category label. In practice, this means that if there is more separation between the two groups, the higher that variable is weighted; and the closer to the boundaries [0,1] the confidence factor will be. Another definition explored was:

$$k = \pm \frac{1}{\sigma} \quad (2.3)$$

Where σ is the standard deviation of the input and the sign is dictated by whether there is a positive or negative association within the data, i.e. determined by whether the Pearson's correlation coefficient is a positive or negative number.

Following the previous work of Beynon (2002), Jones defined the value θ as the mean variable value of all the subjects, so as not to introduce a bias towards either of the two groups. The definitions of k and θ and their implications within the final classification are explored further within Section 3.5.

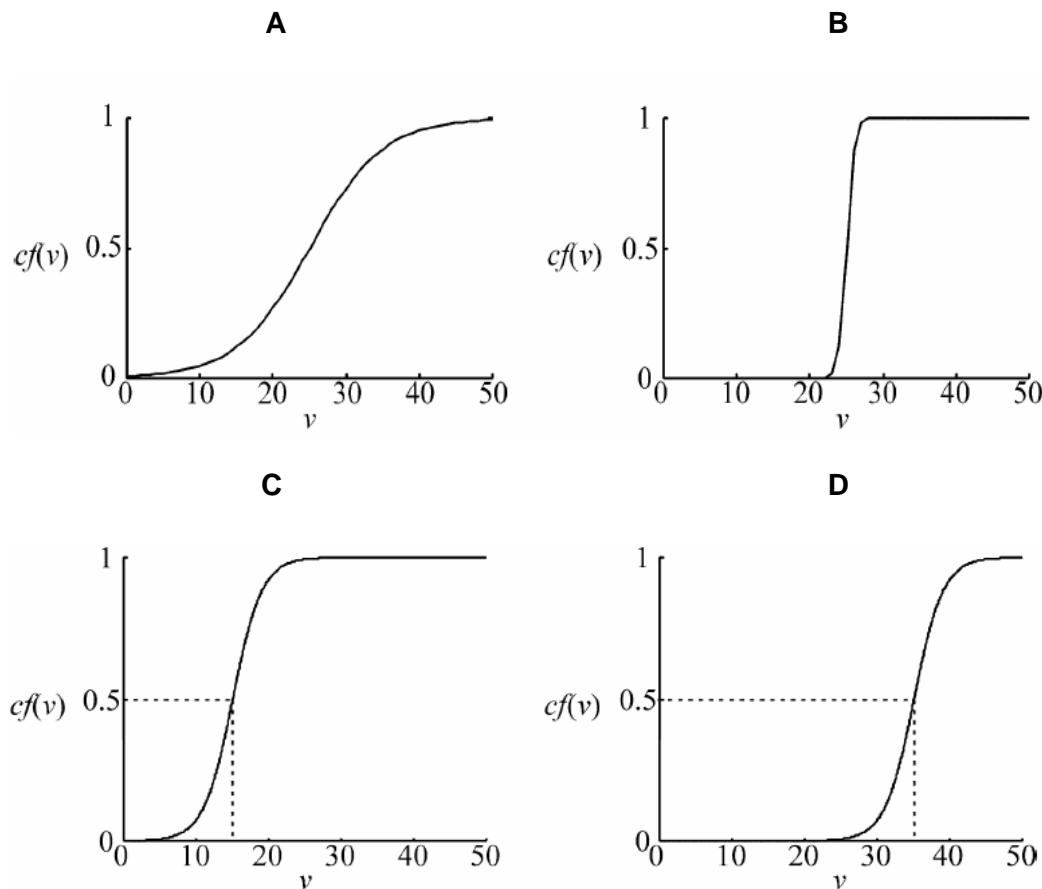


Figure 2.13 Defining the steepness and midpoint of the sigmoid curve using the control factors k and θ . Comparing A and B, the value of k is larger, and hence the sigmoid curve is more steep. Values v are therefore more likely to convert to confidence factors approaching 0 or 1. Between C and D, the steepness is the same, however value of θ is larger and hence the midpoint of the curve at which $cf(v)=0.5$ is shifted to the right.

2.5.5 Conversion to a Body of Evidence

Within the DST classifier, the confidence factor, which is within the range $[0,1]$, is converted to a belief in healthy, a belief in OA and an uncertainty by assigning upper Θ_U and lower Θ_L boundaries of probability. This technique is described in more detail in Section 3.5.3. Within Jones (2004), varying levels of upper and lower boundaries of uncertainty are investigated and it is argued that, for the specific dataset used, the optimal boundaries are $\Theta_U = 1$ and $\Theta_L = 0.8$. Using these probability boundaries, the maximum belief value any one piece of evidence can contribute is 0.2 (20%), and the uncertainty is fixed between the range of 0.8 and 1 (80-100%). The argument presented by Jones (2004) to choose these boundaries was grounded on the optimal classification for the presented dataset, and therefore not universal for other classification tasks. It will

become apparent in the next section that the final level of uncertainty is very much related to the number of variables within the training set.

2.5.6 Dempster's Combination of Evidence

Once the BOE has been calculated for each variable, belief values are then combined using the DST theory of combination of evidence, which can be expressed mathematically as:

$$(m_i \oplus m_j)(\{OA\}) = m_c(\{OA\}) = \frac{m_i(\{OA\})m_j(\{OA\}) + m_j(\{OA\})m_i(\Theta) + m_i(\{OA\})m_j(\Theta)}{1 - (m_i(\{NL\})m_j(\{OA\}) + m_i(\{OA\})m_j(\{NL\}))} \quad (2.4)$$

$$(m_i \oplus m_j)(\{NL\}) = m_c(\{NL\}) = \frac{m_i(\{NL\})m_j(\{NL\}) + m_j(\{NL\})m_i(\Theta) + m_j(\{NL\})m_i(\Theta)}{1 - (m_i(\{NL\})m_j(\{OA\}) + m_i(\{OA\})m_j(\{NL\}))} \quad (2.5)$$

$$(m_i \oplus m_j)(\Theta) = m_c(\Theta) = 1 - m_c(\{OA\}) - m_c(\{NL\}) \quad (2.6)$$

Where m_i and m_j are the BOE of two different variables, with m_c being the combined BOE resulting from $m_i + m_j$, and $m_i(OA)$ being the degree of belief of OA within the first body of evidence. Perhaps a more intuitive explanation can be found following these steps:

1. Combine each belief value from m_1 with that from m_2 , following these rules.
 - a. Two beliefs of $m(OA)$ multiply to results in a belief $m(OA)$
 - b. Two beliefs of $m(NL)$ multiply to result in a belief $m(NL)$
 - c. A belief of $m(OA)$ multiplied by a belief $m(\Theta)$ results in a belief $m(OA)$
 - d. A belief of $m(NL)$ multiplied by a belief $m(\Theta)$ results in a belief $m(NL)$
 - e. A belief of $m(\Theta)$ multiplied by a belief $m(\Theta)$ results in a belief $m(\Theta)$
 - f. A belief of $m(OA)$ multiplied by a belief $m(NL)$ becomes unassigned.
2. Sum all of the resultant $m(OA)$, $m(NL)$ and $m(\Theta)$ individually
 - a. Notice they no longer sum to form 1 (100%), as there is now some unassigned belief
3. Divide all the belief values by $(1 - g)$, where g is the unassigned belief resulting from step 1.e. You could also divide them by their cumulative total, which is the same value. This step scales the belief values such that they add to one. The

normalisation factor g can be interpreted as a “*measure of contradiction or inconsistency*” of the combined evidence.

It appears that uncertainty can only ever decrease when combining new elements of evidence using this technique. The final level of uncertainty within the classification must, therefore, be highly related to the number of input variables within the classifier, and the upper and lower uncertainty boundaries. i.e. the smaller the lower uncertainty boundary, for example, the more input variables would be required to reduce uncertainty below, say, 50%.

This theory of combination of evidence is not without its criticism. In fact, the method of discarding conflicting evidence and then using a normalisation factor has been shown some contexts to produce very erroneous results (Zadeh, 1984, Dezert *et al.*, 2012). These results occur when the normalisation factor is very high. It is, therefore, important to take this into account when determining the validity of the combination (Beynon *et al.*, 2000).

As the lower boundary of evidence is currently set at 80%, this effectively limits the size of the normalisation factor to:

$$g = m_i(\{NL\})m_j(\{OA\}) + m_i(\{OA\})m_j(\{NL\}) \quad (2.7)$$

$$g = (0.2)(0.2) + (0)(0) = 0.04 \quad (2.8)$$

2.5.7 Comparisons with Neural Networks

Being able to understand and interpret the way an input relates to an output is desirable within applied research, particularly when monitoring health outcomes. Interpretation is not only necessary in order to learn about underlying mechanisms, it is also necessary for the validation of any findings which challenge current opinion (Intrator and Intrator, 2001). While neural networks have been explored extensively in monitoring health outcomes, interpretability of results is challenging (Intrator and Intrator, 2001). This is because, unlike in the DST classifier, each activation function is passed a weighted

summation of every input. Therefore, when looking back at how, for example, ten variables arrived at a classification with one hidden layer of five neurons, all ten variables are passed to each of the five sigmoid functions. Contextualising and interpreting the clinical relevance of each hidden layer neuron becomes a difficult task. With the DST classifier, only one input is passed to each sigmoid function. It is therefore very clear to map how each variable contributes to the combined body of evidence.

Both ANNs and the DST use sigmoid activation functions, although in slightly different ways. Within an ANN, the definition of k is not necessary as the inputs have associated weights which mathematically have the equivalent function. Also, within ANNs, a ‘bias node’ can be added. A bias node acts as an additional input to the neuron, always has an input of 1, and itself has an attributed weight which can be iteratively updated. Let’s say that a neuron has two inputs, with two weights, the standard sigmoid function would equal 0.5 when the combined inputs multiplied by their weights was zero. Now, if a bias node was introduced with a weight of -5, the input variables multiplied by their weights would now have to also equal five such that $\text{sig}(x)=0.5$. Notice how the ‘bias node’ has a similar function to that of θ within the DST classifier.

Figure 2.14 shows a schematic of how the DST classifier appears when illustrated in the same fashion as is common with neural networks. Notice how each input is passed to a separate activation function. Both Jones *et al.* (2008) and Parisi *et al.* (2015) have found the DST classifier to perform favourably in comparison to neural networks, although comparisons are made difficult by the different cross-validation algorithms used. These results are promising considering the much more simplistic relationship between the input variable and final classification within the DST classification. It also aids clinical interpretation that the control parameters use traditional statistical parameters which are defining *a priori* to classification itself. Considering the recommendations of Lai *et al.* (2009), further work is still necessary to improve the ability to graphically visualise the trained classifier, and hence aid clinical interpretation of the modelled relationship between input variables and classification output.

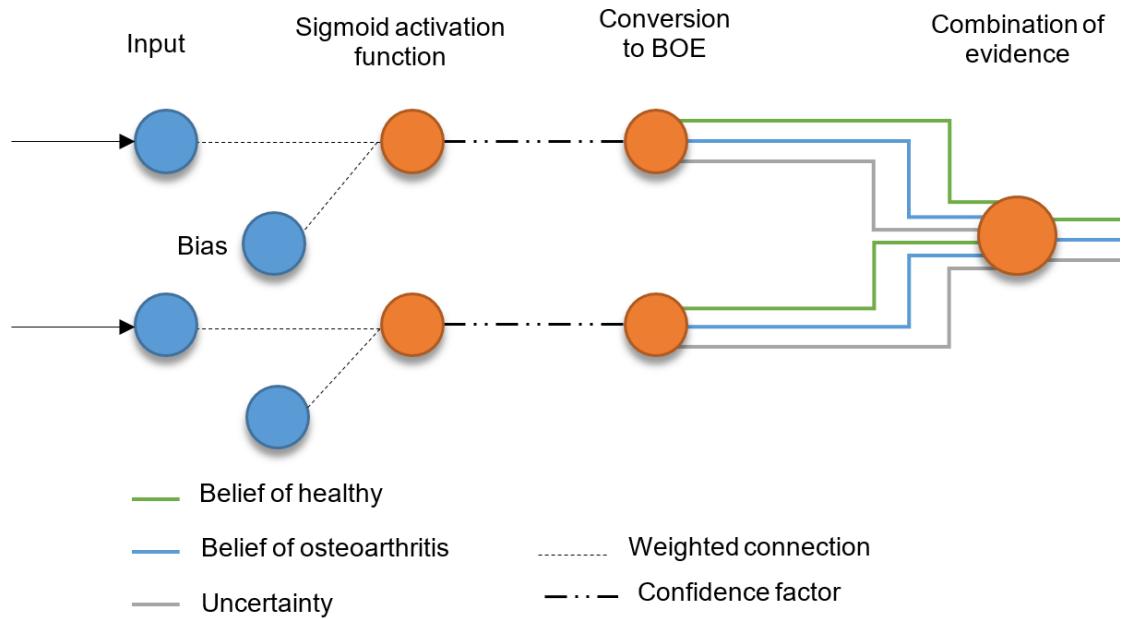


Figure 2.14 A schematic of how Jones (2004)'s DST classifier would look if it was illustrated in the same fashion as a neural network. In this illustration, the DST classifier has only two inputs.

The blue circles represent discrete variables which are merely passed to the processing neurons (orange circles). An input variable is weighted either using the Pearson's correlation coefficient or the standard deviation (Jones, 2004). A bias, θ , is applied using the average value for all the subjects. The output of the sigmoid function is in the range [0, 1] and is converted to a body of evidence (BOE). The BOE from each input variable is combined using the DST theory of combination of evidence.

Chapter 3 - Objective Assessment of Knee Function During Gait

3.1 Introduction

This section details the methods used to assess the knee function of non-pathological (NP) and osteoarthritic (OA) subjects. The data used throughout the entirety of this thesis was collected as part of ongoing research at the School of Engineering, Cardiff University. At the time of writing, the collection of gait biomechanics pre and post-TKR surgery has been ongoing for at least 16 years. Throughout this period, a number of different trained operators have carried out the data collection, following a Standard Operating Procedure (SOP). Due to the natural progression of research methods and technology, the SOPs have evolved over time. The adaptations to the data collection protocols are described in detail by Whatling (2009) and Watling (2014) and represent additions to, as opposed to modifications of, the original protocol described by Holt *et al.* (2001). There have also been hardware changes over this period, the most notable being the upgrade of the motion capture cameras from eight ProReflex cameras (Qualisys, Sweden), to nine Qqus 3 cameras (Qualisys, Sweden). The resultant effect of these hardware changes on the accuracy and precision of the motion capture data is not quantified within this thesis but is discussed in Section 7.3. The previous key studies utilising the Cardiff Protocol to define knee function are listed in Table 3.1, and the relevant additions to the protocol are also listed.

The foundations of the ‘Cardiff Protocol’ for three-dimensional analysis of the tibiofemoral joint are defined within Holt *et al.* (2001). This technique uses a bespoke MATLAB script in order to measure knee joint kinematics using the Grood and Suntay technique (Grood and Suntay, 1983). The approach was adopted by Dr Lianne Jones within her PhD work in order to assess knee function, before and after TKR surgery, in comparison to NP control subjects (Jones, 2004).

Table 3.1 Key previous work at Cardiff University on the functional classification of biomechanical knee function, *Only stair gait was processed, **Not collected for all NP subjects

	Joints included			Biomechanical variables			Anatomical landmarks definition		No. subjects included	
	Knee	Hip	Ankle	Kinematics	GRF	Joint kinetics	Pointer	Marker	OA	NP
Jones (2004)	✓			✓	✓		✓		22	20
Whatling (2009)	✓	✓*	✓*	✓	✓	✓	✓	✓*	32	30
Watling (2014)	✓	✓	✓	✓	✓	✓	✓**	✓	25	23

The next significant development to the Cardiff Protocol was the calculation of knee joint moments (kinetics), which was applied by Dr Gemma Whatling. Initially, Whatling adapted the pointer method, defined by Holt, in order to calculate both knee kinematics and kinetics of stair gait (Whatling, 2009). Whatling adapted the marker set based on the CAST marker protocol (Benedetti *et al.*, 1998, Cappozzo *et al.*, 1995), which included additional markers to define foot and pelvis segments. As opposed to further adapting the MATLAB script developed by Holt (Holt *et al.*, 2001), Whatling instead used Visual3D (C-Motion, USA), an advanced biomechanics analysis software, to assess the biomechanical changes at adjacent joints; the hip and the knee.

Dr Daniel Watling then went on to adopt the latter Visual3D methodology in order to classify OA knee function during level gait. As the pointer method was not necessary when using the Visual3D method, pointer data was not always collected for NP subjects, yet was consistently collected for OA subjects.

As part of the research team within the Arthritis Research UK Biomechanics and Bioengineering Centre, I have continued to collect motion analysis data using the methods described by Jones, Whatling and Watling. The aforementioned data collections also incorporate a number of different elements, such as additional activities of daily living (ADLs) and electromyography; however as this data is not used within this

thesis, the methodologies shall not be further described. While there are notable benefits of including the ankle and hip joint within the biomechanical analysis, the inclusion of only knee kinematics allows for a much larger dataset. It is proposed that this large dataset could be processed and utilised in order to achieve a large enough training body of OA and NP subjects, such that the following research aims can be addressed:

Aim 1: Explore the validity of the classifier control variables

Aim 2: Explore the sample size required to classify osteoarthritic subjects accurately.

Aim 3: Assess the reliability of the LOO cross-validation technique as an estimate of classification accuracy.

Aim 4: Assess which biomechanical gait features best discriminate between NP and OA gait.

For a smaller number of subjects, the force plate locations have also been calibrated, and hence joint moments can be calculated by adapting the knee kinetic calculations adopted by Dr Whatling. Furthermore, Dr Jones did not consider the mediolateral GRF within the classification of OA subjects due to technical difficulties during several the data collections. The following question will therefore also be addressed:

Does the inclusion of mediolateral GRF force and knee joint moments have a significant impact the ability to classify osteoarthritic subjects?

3.2 Data Collection

3.2.1 Non-pathological Subject Recruitment

The recruitment of NP volunteers was approved by the Research Ethics Committee for Wales and Cardiff University Health Board. Volunteers were recruited via email and poster advertisements throughout Cardiff University and the wider South Wales community. The criteria for inclusion in the study as a NP volunteer was as follows:

- No self-reported OA, or pain in the foot, ankle, knee, hip or back.
- No known difficulty performing ADLs.
- No history of musculoskeletal conditions which required medical treatment e.g. ligament or meniscal tear.
- No other musculoskeletal, neurological or visual condition which might affect the way they move.
- An ability to give informed consent.

Any volunteers who expressed an interest in participating were given an information sheet. If the volunteer understood the information sheet and was happy to proceed, they were asked to sign a consent form.

3.2.2 Osteoarthritis and TKR Patient Recruitment

The recruitment of NHS patients with severe osteoarthritis was approved by the Research Ethics Committee for Wales and Cardiff and Vale University Health Board. These patients had all been listed for TKR replacement surgery.

The criteria for inclusion in the study as a patient volunteer was as follows:

- An ability to walk 10m without a walking aid.
- An ability to give informed consent.
- No unrelated musculoskeletal, neurological or visual condition which might severely affect the way they move.

Before taking part in any aspect of the study, patient volunteers were given a patient information sheet. If the patient volunteer was still interested in taking part, they were asked to sign a patient consent form.

3.2.3 Gait Assessment

The gait analysis protocol matched that described by Watling (2014), and is, therefore, shall for conciseness only be described in brief. Further information and rationale behind the marker placement procedure is given within Section 4.2 .

Calibration - Before the volunteer arrives, the infrared cameras were calibrated such they could calculate their position and orientation relative to one another. Part of this procedure involved the definition of the origin and orientation of an orthogonal global coordinate system (GCS). The GCS was defined by placing an L-Frame, where the long hand of the frame represents the x-axis, the short hand the y-axis, and the vertical and mutually orthogonal axis is called the z-axis. The definition of the GCS defined the coordinate system within which all the data was then described by the camera system.

With the L-Frame in place, the camera system was then calibrated by waving a calibration wand; which contained two markers, through the intended volume of capture. The camera wand was passed through the volume for 45 seconds, as each camera recorded the movement of the two markers relative to the L-Frame. The cameras match the identified trajectories of the two markers, in order to calculate the disparity between their individual views (how differently the markers of the wand and the L-Frame appear within their 2D view). This measures disparity across the frames of the calibration trial were used to calculate the position and orientation of each camera relative to the GCS. The 3D coordinates of the markers are reconstructed by the multiple cameras, and a residual error is calculated which reflects the discrepancies between these 3D reconstructed points. If cameras showed residual errors greater than 1mm, the calibration procedure was repeated.

The position of the force platforms relative to the GCS was calculated by using two metal plates with markers attached, which enabled the calculation of the 3D coordinates at each corner of the force platform. This is discussed further in Section 3.3.4, and shown in Figure 3.4. This step crucial in identifying accurate COP coordinates, and hence calculating knee kinetics.

Informed consent – The participants were given an information sheet at least 48 hours before their first assessment, which explained the purpose of the study, what would be expected from them, and how the data would be anonymised and stored. Once it was established that the participant has read and understood the information sheet, they were asked to sign a consent form, which was signed by one of the lead researchers.

Questionnaires: The participant was asked to fill in relevant questionnaires regarding their knee pain and function. Additional questionnaires were added during the course of the study, following ethical amendments, which are described further in Section 5.3.3.

Clothing: The participant was asked to change into suitable clothing, having been advised within the patient information sheet to bring a loose-fitting pair of shorts and t-shirt. The assessment was carried out without footwear.

Anthropometrical measures: Height, weight, knee width, depth and knee girth measurements were taken. The measurement of height and weight is important for both the calculation of BMI, but also the normalisation of joint moments and GRFs. The added anthropometrical measures of knee width, depth and girth followed the protocol of Jones (2004). The anthropometrical measures were included in the final classification presented by Jones, and hence were considered in order to compare initial classifications within Section 3.6.1.

Assessment preparation: Retroreflective markers were attached to the subject using hypoallergenic wig tape. The full marker set is discussed in Section 4.2.1. For several subjects, EMG analysis was also performed however this data isn't presented in this thesis. The placement of electrodes didn't affect the marker locations, therefore it is

assumed the addition of EMG didn't affect the calculation of kinematic and kinetic parameters.

Assessment of gait: Participants were asked to walk barefoot at a self-selected pace over a 10m walkway. The force platforms were located within the middle third of the walkway. If clean force platform readings weren't recorded, the participant was sometimes asked to start slightly further back, and then the measurement was repeated. This whole process was repeated until there were at least six clean force platform readings for each leg. The participant was not made aware of the existence or location of the force platforms; however, they may have been conscious that something was different about this area of flooring. This is discussed further in Section 7.1. Following the assessment of gait, the subjects were also asked to perform other ADLs such as stair-climbing and sit-to-stand. This data, however, is not presented in this thesis. For further information, the reader is directed to Chapter 2.3 of Watling (2014), within which a comprehensive description is provided.

3.3 Optimising the Pointer Method Pipeline

Previous work at Cardiff University using the DST classifier to define OA knee function has had relatively small cohort sizes. This is in part due to the progression towards Visual3D processing methods, for which the previously collected data cannot be readily used. As previously mentioned, since the original DST classifier development of Jones (2004), there has been a large number of subjects collected using the pointer method, and hence the opportunity to process this data and add these subjects to the training body of OA and NP subjects. It is this data which is used within this chapter in order to critique and develop the current Cardiff protocols for the reduction and classification of biomechanical data.

The pointer method used is an in-house custom built set of MATLAB scripts (Holt *et al.*, 2001) in order to calculate tibiofemoral joint kinematics of tracked data using the joint coordinate system proposed by Grood and Suntay (1983). One of the challenges in using this technique to process a large number of subjects is that the current pipeline for data processing is time-consuming, and lacks the flexibility of reprocessing data in order to compute new variables or to modify how variables are calculated. The current data processing pipeline is summarised below and shown visually in Figure 3.1.

1. Run input ‘variables.m’ (calculates kinematics, ground reaction force)
 - a. Path of stored data must be set
 - b. Force plate columns must be checked and then set
 - c. Subject ID must be selected
 - d. Height and weight must be input
 - e. Good walking trials must be checked from paperwork and input
 - f. Good force trials must also be checked and input
 - g. Gain and camera frequency must be input
 - h. Force plate frequency must be checked from file headers and input
 - i. Results must be organised appropriately.

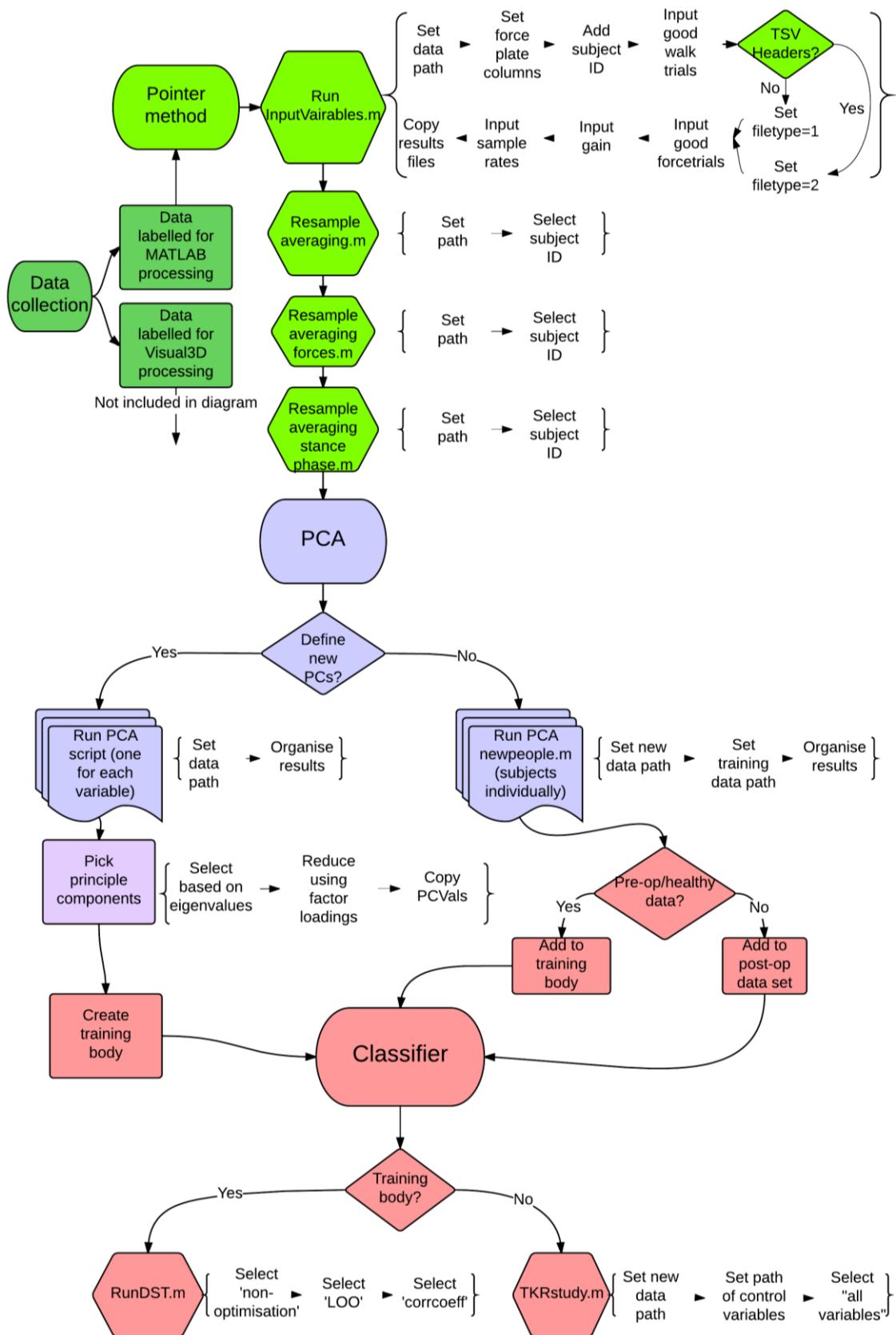


Figure 3.1 Schematic of the data processing pipeline. Light green represents the calculation of joint kinematics and kinetics using the pointer method, blue represents the application of principal component analysis to reduce temporal data, and red represents the application of the Cardiff Classifier.

2. Run 'resampleaveraging.m' (calculates a mean of the six kinematic trials)
 - a. Set path
 - b. Select patient's initial
3. Run 'resampleaveragingforces.m' (calculates a mean of the six kinetic trials)
 - a. Set path
 - b. Select patient's initial
4. Run 'resampleaveragingstancephase.m' (resamples the forces over stance phase and averages them)
 - a. Set path
 - b. Select patient's initial

Good force trials are defined as gait cycles for which there was a clean heel strike and toe-off on a single force platform. These are visually checked and noted during the motion analysis session and these notes should be referred to when processing data.

The resultant knee kinematics and kinetics are then stored as tab-delimited text files.

Previous work has applied PCA in order to reduce temporal data. Before this step is possible, it is necessary to collate all the waveforms into tables for all the patients e.g. all flexion/extension angles for all of the patients in a single table.

The processing of data using this method is very time-intensive. This is particularly a problem when working with in-house scripts because small changes may be necessary when making further developments; however, this would require all of the data to be reprocessed from scratch. A solution to this is to introduce software which saves the user inputs during processing, such that data can be automatically reprocessed with minimal user input.

While data is quality checked before processing, it is worthwhile quality checking during the calculation of motion analysis data. For example, the tracking marker labels can often switch during motion trials. This may not be obvious when simply viewing the QTM files but becomes more apparent when viewing the calculated kinematics, in which sharp

peaks will occur. A further problem is the required time to collate the data in a format which is suitable for post-processing.

3.3.1 Design Criteria for the Development of the Pointer Method Pipeline

Minimal user input – There are currently many requests for user input during data processing. As part of good data management, this information should already be stored within a central subject database. The MATLAB software could access this centrally, without the need of user prompts.

Data organisation – Raw and processed data should be stored in a methodical way, using consistent file formats such that it is possible for the MATLAB software to locate this data, based on an ID number, within the central database. This also enables the potential for automated assimilation of patient data for further processing, such as PCA.

Inverse dynamics – Previous work by Whatling (2009) adapted the pointer method software to include the calculation of knee joint moments during stair ascent and descent. The software should be adapted to calculate the knee joint moments during level gait. The procedure for calibrating the location of the force platforms has since changed, and the software should be adapted to use the raw calibration data which is stored within Qualysis Tracker Manager files.

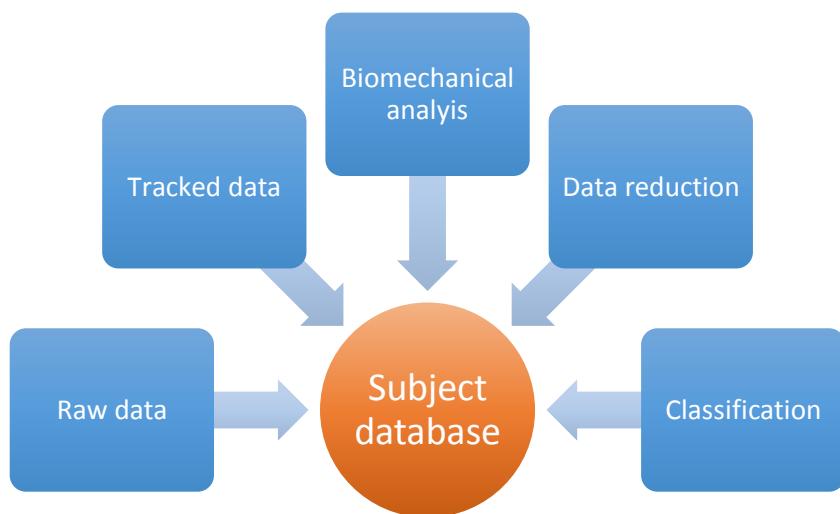


Figure 3.2 A graphical representation of the proposed ideal data management solution. Data should be stored in a central methodical way such that they both receive and send necessarily processing information to the same central subject database.

3.3.2 Patient Spreadsheet and Data Organisation

To enable the data processing pipeline to receive information from the subject database, it must be stored in a way which can be accessed through MATLAB. It is possible to import data from both MS Access and MS Excel databases/spreadsheets; however, the use of Access databases requires the use of the Mathworks Database Toolbox™. Due to the simplicity of Excel and the author's familiarity, it was decided that a subject spreadsheet would be created within Excel. Table 3.2 shows the proposed layout of the spreadsheet. This layout contains all of the information which would normally be manually inputted during the MATLAB pointer method.

Table 3.2 Patient Spreadsheet layout

Visit ID	Leg	Date of Visit	Height	Weight	Visit Type	Gender	Age	Knee Outcome	Oxford Score	Gain	Camera Sample	Force Sample	First FP Column	Good Walking Trials				Good Force Trials				Rotations?	ForceOk?	Moments Ok	Notes

To facilitate the automation of data processing, the data must follow a naming protocol.

Data for each specific gait analysis visit was saved using the following format:

- **** \ "Patient type" \ "Patient identifier" \ "Initials" "Visit" \
- e.g. D:\Osteoarthritic \#Am\Am 1&2\

Where #Am is the subject initials, the hashtag indicating that they are the second subject with those initials, and 1&2 indicating this to be a pre-operative visit. A three-month post-op visit is 3&4, six months is 5&6, and 12 months is 7&8. This allows for a MATLAB script to be written which can automatically find the data of any subject within the main spreadsheet, whilst also organising the results accordingly. Within each of these files, there are the following folders:

Raw data – This is the raw untracked QTM data. This is useful to have if there are mistakes in tracking, or if something unusual has occurred with the processed

file which needs to be checked against the original. The data here will be raw QTM files.

Processed data – This is where the labelled marker data, which has been tracked within QTM, is stored. The data in this folder must be saved using a consistent naming protocol. This allows it to be located and accessed by an automated script, such that data can be collated for further analysis.

There are a set of scripts within the standard MATLAB library which are designed to read (xlsread) and write (xlswrite) Excel files to and from MATLAB. The xlsread command is used to load the cell data from Excel into a new array of cells.

In order to automate the process of data processing, it is useful to process each subject within a “for-loop”. The primary for-loop is created such that its value corresponds to the row number within the subject database. The user defines the origin of the main data folder e.g. D:\Osteoarthritic\ and the script defines the subject data path by using the “VisitID” column.

3.3.3 Knee Kinematics

Knee kinematics are calculated within the for-loop using ‘kneeangles.m’. This element of the software hasn’t been fundamentally changed and therefore, for a detailed explanation, please see Jones (2004). The process can be summarised as:

1. Calculate an LCS for the pointer, and define its transformation matrix relative to the GCS.
2. Using the known diameters of the pointer, calculate the position of the very tip of the pointer within the GCS. This identifies virtual landmarks; each can be used for anatomical calibration.
3. Define an anatomical coordinate system (ACS) for each segment using the virtual landmarks.
4. Define a coordinate system for the marker clusters (MCS), and define its transformation matrix relative to the ACS.

5. Track the movement of the MCS during the motion trial. By assuming the transformation between the ACS and MCS for the segment is constant throughout the motion trial.

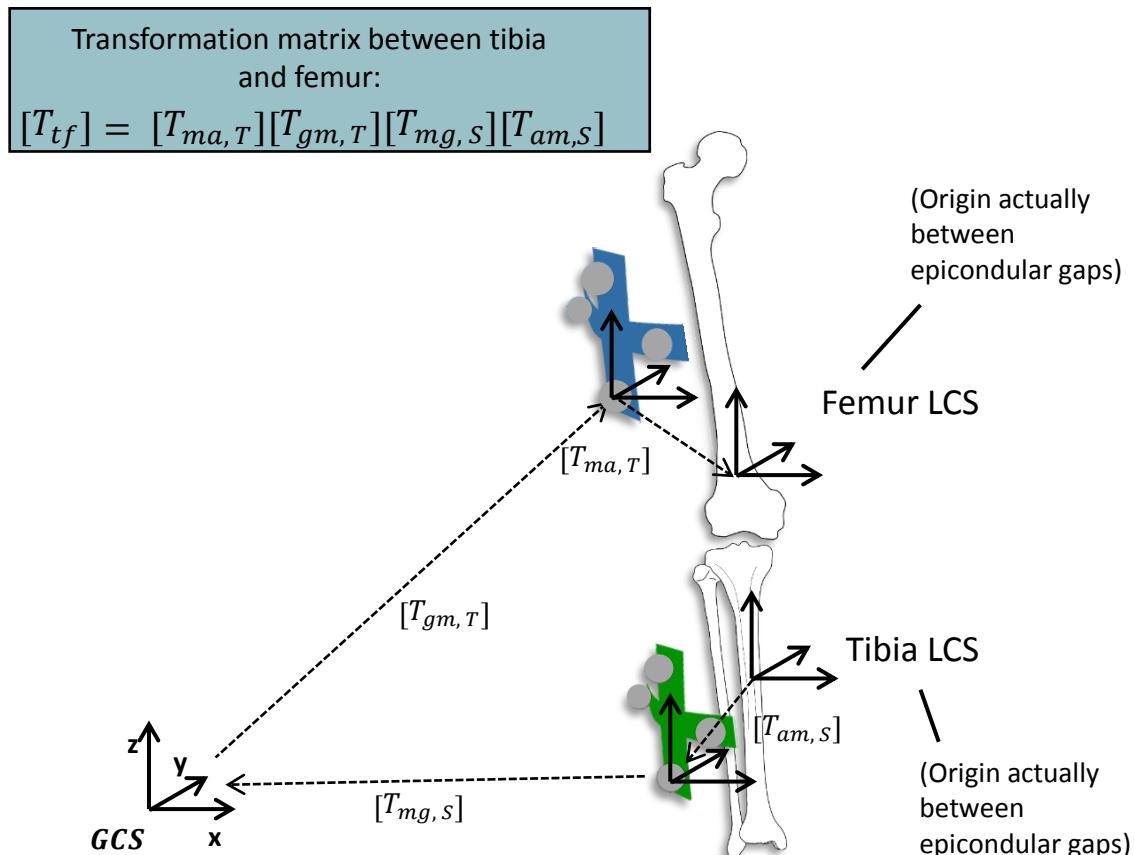


Figure 3.3 Illustration of how the 6 degrees of freedom of the tibia relative to the femur are defined. The anatomical marker coordinate systems are defined relative to the marker cluster coordinate systems. This transformation matrix ($T_{ma,T}$ and $T_{am,S}$) is assumed to be fixed. The marker clusters are tracked for every frame of the motion, therefore these transformation matrices ($T_{gm,T}$ and $T_{mg,S}$) are recalculated for every frame.

3.3.4 Calculating Knee Kinetics

Three-dimensional GRF's were calculated from the signals by modifying the previously developed MATLAB software (Holt *et al.*, 2001).

The relationship between the raw analogue signals and the forces and moments acting on the plate are defined by the calibration matrix of the force plate. At Cardiff University, the force platforms were upgraded from the original Bertec force platforms (Bertec Corporation, Ohio, USA) . The calibration matrix of these two plates is different, making

it necessary to insert code which made sure that any files after the 1st Jan 2012 used the alternative calibration matrix.

For the older force plates the following calibration matrix is applied:

$$C = \begin{matrix} 1741 & -8.5 & 3.1 & 8.7 & 9.4 & 0.9 \\ 6.8 & 1748.5 & -18.4 & -1.5 & 0.6 & -7.1 \\ 22.6 & -38 & 3684.2 & 71.3 & -47.1 & 19.5 \\ 0.6 & -119.7 & -3 & 1139.7 & -2.8 & 2.7 \\ 114.6 & -3.3 & 6.6 & 4.8 & 797.2 & 2.3 \\ 1.6 & -5.3 & -7.6 & -8.0 & -0.2 & 411.3 \end{matrix}$$

For the newer force plates, the following calibration matrix is given in the manual:

$$C = \begin{matrix} 1000 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1000 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2000 & 0 & 0 & 0 \\ 0 & 0 & 0 & 600 & 0 & 0 \\ 0 & 0 & 0 & 0 & 400 & 0 \\ 0 & 0 & 0 & 0 & 0 & 300 \end{matrix}$$

The new version of 'inputforces.m' takes the Visit Date from the data processing sheet, and uses the new calibration matrix if the serial date number is greater than 734869 – which corresponds to the 01/01/2012. The GRF data is calculated relative to the coordinate system of the force plate. To calculate the external moments acting about the knee, it is necessary to identify the location of the force plates within the lab GCS. The

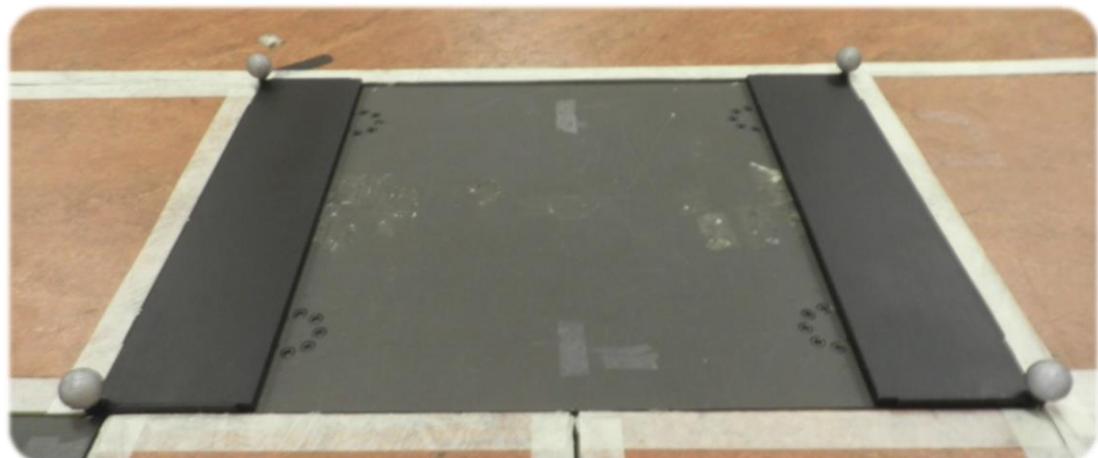


Figure 3.4 Frame used to define the location of the corners of the force plate relative to the lab GCS

corners of the force plate are identified using a metal frame, as shown in Figure 3.4. These markers are labelled within QTM and used to generate the force plate location. In doing so, QTM requires the offset distance between the surface of the force plate and the centre of the marker – which has been measured as 22.5mm.

The calibrated locations of the force plates are then stored in the QTM files until the next calibration. It is possible to export these as tab-delimited text files. A transformation matrix for the force plate is then defined as follows:

1. Define a vector v_1 from corner 3 to corner 2.
2. Define a vector v_2 from corner 2 to corner 1.
3. Define unit vectors u_1 and u_2 in the direction of v_1 and v_2 .
4. Define a unit vector u_3 as the cross product of u_2 and u_1 .
5. Define a unit vector u_4 as the cross product of u_1 and u_3 .
6. Define the centre of the force plate as the intersection of two diagonals.
7. Define the transformation matrix:

$$[T] = \begin{bmatrix} [R] & [O] \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.1)$$

$$[R] = [\hat{u}_4 \quad \hat{u}_1 \quad \hat{u}_3] \quad (3.2)$$

8. Transform the centre of pressure coordinates from the force plate coordinate system to the lab GCS:

$$COP_{GCS} = T^{-1} \begin{bmatrix} COP_x \\ COP_y \\ COP_z \\ 1 \end{bmatrix} \quad (3.3)$$

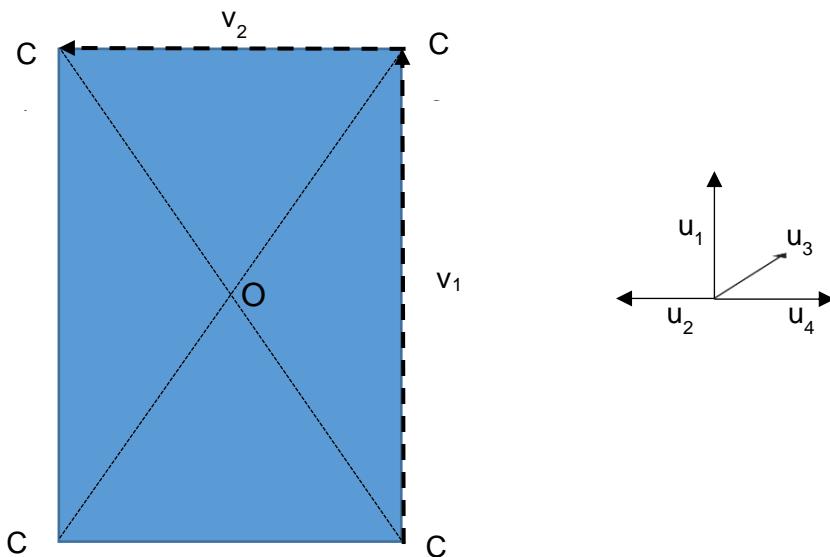


Figure 3.5 The calculation a local coordinate system of the force plate (blue). C1-C4 show the default numbering of the corners of the force plate in the tab delimited exported of QTM. Unit vectors $u_1 - u_4$ are shown next to the force plate. The coordinate system is created using u_4 (x), u_2 (y) and u_3 (z). The origin of the coordinate system is at the intersection of the two diagonals, O.

3.3.5 Estimating the Moments About the Knee

To estimate the external moment acting about the knee, it is necessary to consider the effect of both the GRF and the inertia of the segment. The inverse dynamics approach was modified from that used by Whatling (2009). The lower limb was modelled using only thigh and shank segments (as shown in Figure 3.6), as no pointer position data is recorded for the foot. Previous work by Whatling used the segment inertial parameters provided by Zatsiorsky and Seluyanov (1983). These use bony landmarks as reference points which don't correspond to the anatomical landmarks defined in this study. De Leva (1996) defined a series of updated centre of mass and radius of gyration values for the shank using the knee joint centre and the lateral malleolus. These have been used in this study.

The primary limitation of this approach is that it ignores the effect of the foot and ankle on the resultant moment about the knee. Previous work at Cardiff University has found differences in joint moments when calculated using this approach, as opposed to a full lower limb model within Visual3D (Whatling, 2009); this may be as a result of the

exclusion of the foot segment. Other limitations are that the regressions, which estimate inertial parameters based on segment length, use healthy, Caucasian, college-aged males, and may be less accurate when applied to different subject demographics. For example, subjects with OA frequently have high BMIs, which would increase the radius of gyration of the segments. The inertial parameters are also assumed constant during motion, which isn't always true, particularly for obese subjects whose soft tissues are likely to move during impact and motion.

By summing moments about the centre of mass, the external joint moment M can be calculated. In the sagittal plane, this can be defined as:

$$M_x = I_x \alpha_x + R_z(COM_y - KJC_y) + R_y(COM_z - KJC_z) + F_z(COM_y - COP_y) - F_y(COM_z - COP_z) \quad (3.4)$$

Where:

I_x = Moment of inertia of the shank segment about the x (coronal plane) axis.

α_x = Angular acceleration of the shank segment

R_z, R_y = Reaction forces at the KJC

COM_y, COM_z = Position of the COM in the GCS

KJC_y, KJC_z = Position of the knee joint centre in the GCS

COP_y, COP_z = Position of the centre of pressure of the GRF in the GCS

F_z, F_y = The components of the GRF

The moment of inertia I , is determined from Winter (1990) for each plane where:

$I = \text{Shank mass} \times (\text{radius of gyration})^2$

Shank mass was defined based on Zatsiorsky and Seluyanov (1983):

$LLmass = 0.433 \times \text{Subject weight}$

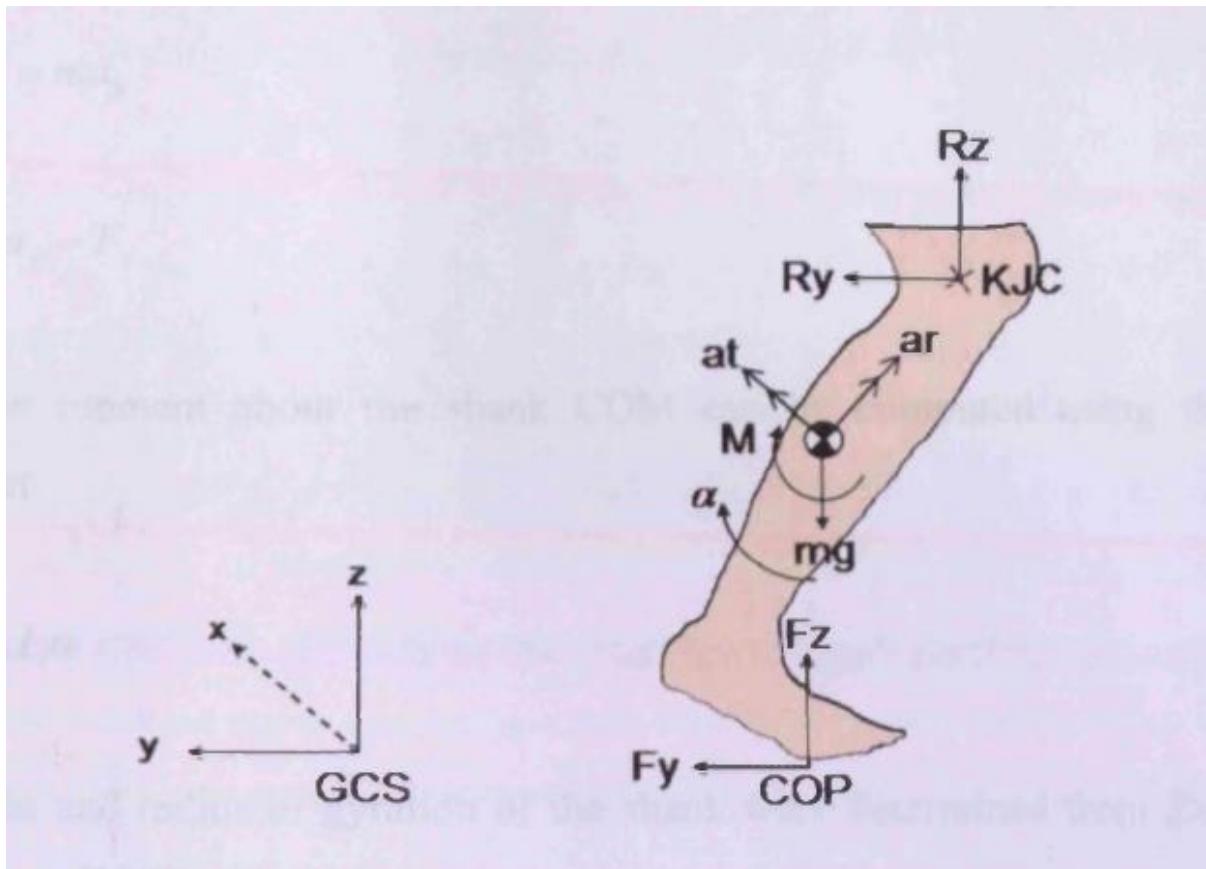


Figure 3.6 A free-body diagram for the right shank, showing joint reaction forces R_z and R_y , ground reaction forces F_z and F_y . The diagram also shows the tangential (at), radial (ar), and angular (α) accelerations acting about the centre of mass (checkered circle), contributing to a joint moment M . This is shown in the sagittal plane – the same calculations were performed for the coronal and transverse plane. Reprinted from (Whatling, 2009). As previously mentioned, the radius of gyration was calculated using the regressions from De Leva (1996). R_z and R_y can be found by resolving forces horizontally and vertically. As forces aren't in equilibrium:

$$\sum F = ma \quad (3.5)$$

Hence

$$R_z = ma_z - F_z + mg \quad (3.6)$$

$$R_y = ma_y - F_y \quad (3.7)$$

The position of the centre of mass of the shank is adapted from De Leva (1996) as 40.47% of the distance from the medial malleolus to the lateral epicondylar gap. This process can then be repeated for the coronal and transverse plane. The resultant

moments are then normalised by dividing by both the height and the mass. Moments are further normalised in the time domain by resampling to one hundred data points and expressing as a percentage point of stance phase.

3.3.6 Data Verification and Saving

When processing biomechanical data, there are several sources of errors. Some of these can be spotted visually if the user is familiar with both normative and osteoarthritic gait data. It is, therefore, important to visually check each kinematic and kinetic waveform during batch data processing. Some common errors which can be visually identified are:

- Marker labelling errors – Sometimes labelled markers' trajectories swap with each other during a motion trial. This obviously leads to sudden changes in joint angles and large segment inertial parameters affect joint moments.
- Force data is in different analogue channels – If the force data doesn't look correct it could be that the data for the second force platform is in different analogue channels. There have been times when different sets of channels have been used; the additional code has therefore been written, which cycles through the other possibilities until the correct analogue data is identified.
- Force platform incorrectly zeroed – Before heel strike and after toe-off, the GRF data should approach zero. If the user indicates the ground reaction force data doesn't appear correct but the correct channels appear to have been selected, then the code has been written which attempts to correct the force platform data. This is achieved by taking the mean of the last frames of gait cycle, at which point there should be no contact with the force plate, and offsetting the rest of the data by this value.
- The subject is walking in the opposite direction – In some cases, data for subjects walking along the walkway in the opposite direction to the lab axis was collected. To assess this, the code was written which assumes all subjects should have a posterior GRF on average during the first third of the gait cycle. If this is not the

case, the subject is assumed to be walking in the opposite direction, in which case both anteroposterior and mediolateral forces must be flipped.

All processed data is automatically saved following a file-naming protocol into a folder named “Matlab results” within the main data folder. This consistent data organisation makes it much easier to catalogue and collate in the later stages.

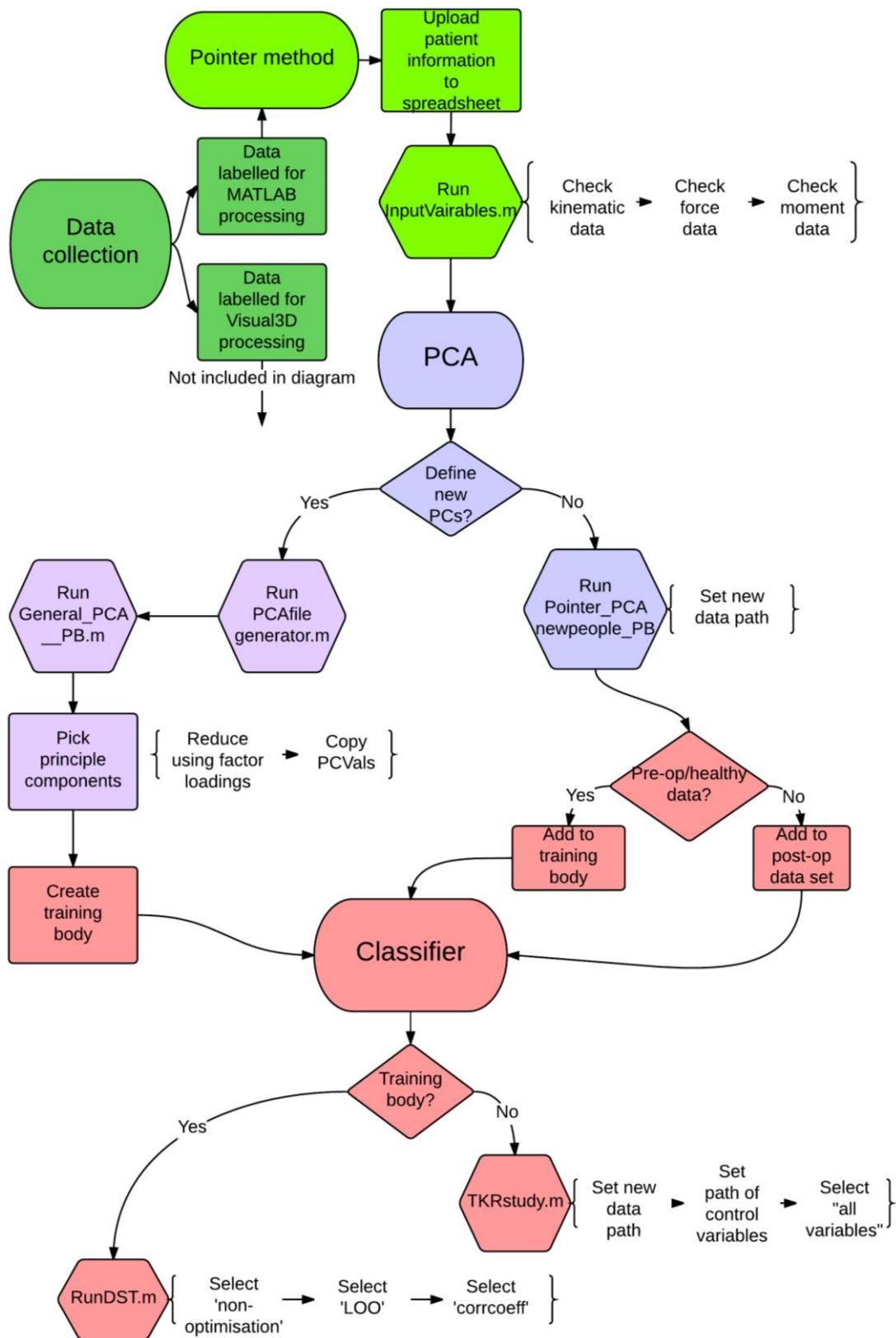


Figure 3.7 The new data processing pipeline Schematic of the data processing pipeline.

Intended for direct comparison with Figure 3.1. Light green represents the calculation of joint kinematics and kinetics using the pointer method, blue represents the application of principal component analysis to reduce temporal data, and red represents the application of the Cardiff Classifier. Notable differences are the reduction in steps within the pointer method, and the automatic assimilation of into spreadsheets for Principal Component Analysis.

3.4 Principal Component Analysis

Kinematic and kinetic waveforms are further processed using PCA to reduce the size of the dataset while retaining important temporal information. A previously developed MATLAB software has been adapted in order to calculate PCs. The steps involved in performing PCA have been introduced in Section 2.4.1. A visual way of interpreting principal component analysis shall be introduced, and then the steps involved in computing them shall be elaborated.

Figure 3.8 shows actual waveform data of knee flexion during gait for a mixture of NP and OA subjects (data from Chapter 3). As discussed in Section 2.4.1, one method of reducing this into fewer discrete variables might be to take maxima and minima, or the ROM during stance or swing phase; however, this method involves subjectively discarding temporal information.

The method adopted by Jones (2004) involves treating every percentage point of the gait cycle as an independent variable. For n data points, there are n dimensions – therefore, for visual illustration, the method will be first outlined using two data points in two dimensions. Figure 3.9 shows just two points of the waveform: the knee flexion angle at 40% and 60% of the gait cycle, plotted against each other for all subjects.

3.4.1 Standardisation

Generally, the first step to PCA is to standardise the data (Chau, 2001a). This can be calculated by removing the mean and dividing by the standard deviation, and is sometimes referred to as the ‘z-score’:

$$z = \frac{x - \mu}{\sigma}$$

By dividing by the standard deviation, we are effectively scaling the independent variables by different amounts. This increases the effect that small deviations, in areas of the waveform with low variation, have on the direction of the principal component, and

decreases the effect that large deviations in areas of the waveform with large variation have. Figure 3.10 shows the z-scores of the two variables plotted against each other.

3.4.2 Correlation Matrix

The next stage is to calculate the correlation matrix of the data (Chau, 2001a). The correlation matrix is related to the covariance matrix as follows:

$$\text{corr}(x, y) = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y} \quad (3.8)$$

As the data has been standardised, the standard deviation of both variables is one, hence the covariance matrix is the equivalent of the correlation matrix. In this example, the covariance matrix is:

$$C = \text{cov}(x, y) = \begin{bmatrix} 1 & 0.1808 \\ 0.1808 & 1 \end{bmatrix} \quad (3.9)$$

Therefore, the person's correlation coefficient between the two variables is 0.1808 – a very weak correlation.

3.4.3 Eigendecomposition

The next step is the Eigen decomposition of this matrix, i.e. the eigenvectors \mathbf{u} and the scalar eigenvalues λ which satisfy the following:

$$Cu = \lambda u \quad (3.10)$$

Where \mathbf{u} is the matrix of eigenvectors of \mathbf{C} , and λ is the diagonal matrix of eigenvalues. For an $n \times n$ matrix, each eigenvector has n dimensions, and each eigenvalue is a single scalar value. We can see that, when a matrix is multiplied by the eigenvector, it results in a vector in the same direction as the original eigenvector; however, it is scaled by a constant λ .

This equation can be rewritten as follows:

$$(C - \lambda I)\mathbf{u} = 0 \quad (3.11)$$

Where I is the identity matrix, which for a 2x2 matrix is:

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (3.12)$$

Therefore, in this example:

$$\begin{bmatrix} 1 - \lambda & 0.1808 \\ 0.1808 & 1 - \lambda \end{bmatrix} \begin{bmatrix} u_x \\ u_y \end{bmatrix} = 0 \quad (3.13)$$

If \mathbf{u} is non-zero, then $(C - \lambda I)$ must be singular, hence the determinant equals zero.

$$\begin{vmatrix} 1 - \lambda & 0.1808 \\ 0.1808 & 1 - \lambda \end{vmatrix} = 0 \quad (3.14)$$

$$(1 - \lambda)^2 - 0.1808^2 = 0 \quad (3.15)$$

$$(1 - 0.1808)^2 - 2\lambda - \lambda^2 = 0 \quad (3.16)$$

This forms the quadratic equation

$$\lambda^2 - 2\lambda - (1 - 0.1808)^2 = 0 \quad (3.17)$$

Which has roots

$$\lambda = 1.1808 \text{ and } 0.8192 \quad (3.18)$$

Subbing each of these values into back into Equation (3.13) gives us two possible solutions:

$$\begin{bmatrix} u_x \\ u_y \end{bmatrix} = \begin{bmatrix} u_x \\ u_x \end{bmatrix} \text{ when } \lambda = 1.1808 \quad (3.19)$$

$$\begin{bmatrix} u_x \\ u_y \end{bmatrix} = \begin{bmatrix} u_x \\ -u_x \end{bmatrix} \text{ when } \lambda = 0.8182 \quad (3.20)$$

By their very definition, there are infinite solutions for each eigenvector; however, what is important is their direction. It is usual to take the unit vector form of each vector.

$$u = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \text{ when } \lambda = 1.1808 \quad (3.21)$$

$$u = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \text{ when } \lambda = 0.8182 \quad (3.22)$$

The eigenvectors are ordered in descending order based on their corresponding

eigenvalues. In this case, the eigenvector $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ is the first principal component axis,

shown in Figure 3.10. The majority of variance is represented along this axis. The exact amount can be calculated from each eigenvalue as a percentage of all of the eigenvalues, i.e.:

$$\frac{\lambda_i}{\sum_{i=1}^n \lambda_i} = \frac{1.1808}{1.1808 + 0.9182} = 0.5904 = 59.04\% \quad (3.23)$$

The linear line of best fit using a least squares algorithm has also been shown in Figure 3.10. This highlights a very distinct difference between the linear regression and PCA. Both techniques minimise the cumulative distance of each point from the vector; however, in linear regression, this is minimised in the original y-axis (shown as d_1); whereas with PCA, the distance of each point orthogonal to the vector is minimised (shown as d_2).

3.4.4 Transforming Data Points

Figure 3.11 shows the data plotted on the new axis: the two eigenvectors, PC1 and PC2. Each point can be transformed to the axis system simply by multiplying by the eigenvector. This can be referred to as the PC scores. In this figure, the NP subjects have been plotted as blue circles and the osteoarthritic subjects as orange. From visual inspection, it appears that, while describing less variance, the second principal component is a better discriminator between OA and NP subjects. Interestingly this matches the findings of (Deluzio and Astephen, 2007). This highlights a challenge when discarding dimensions using PCA, because the dimensions which describe the most variance within a dataset aren't necessarily the ones which are best at discriminating between different classes within that dataset. It is also worth noting that steps need to be taken in order to contextualise the PCs such they can be interpreted.

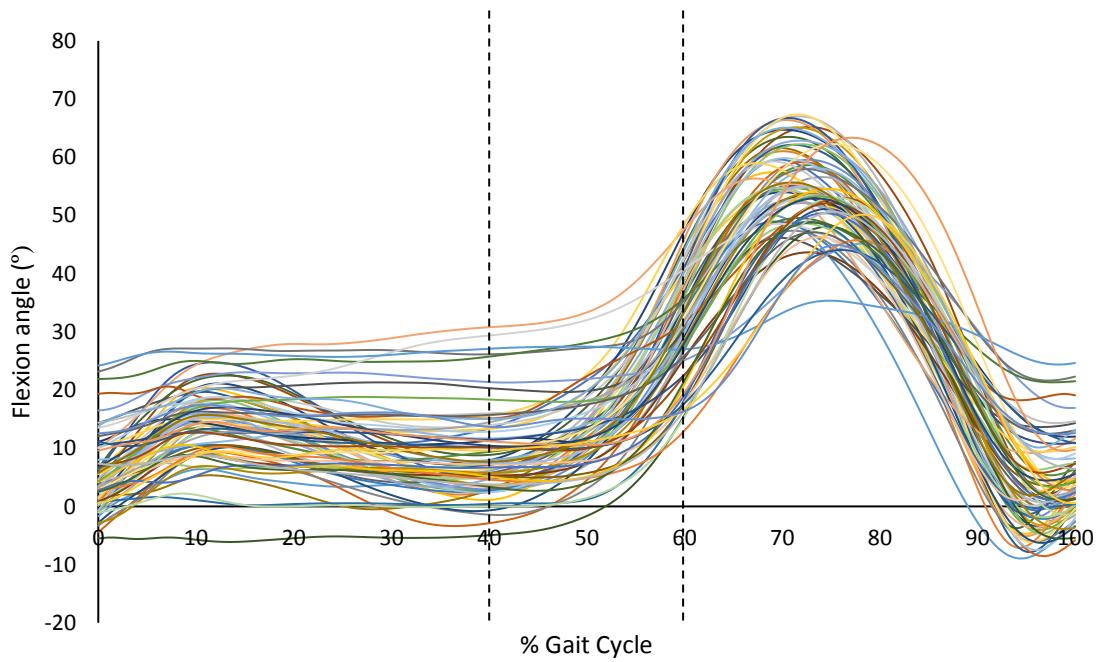


Figure 3.8 Example of knee flexion/extension angle during level gait a combination of OA and NP subjects. The angle at two discrete points, 40%, and 60% (dashed lines), are selected for the two-dimensional demonstration of principal component analysis.

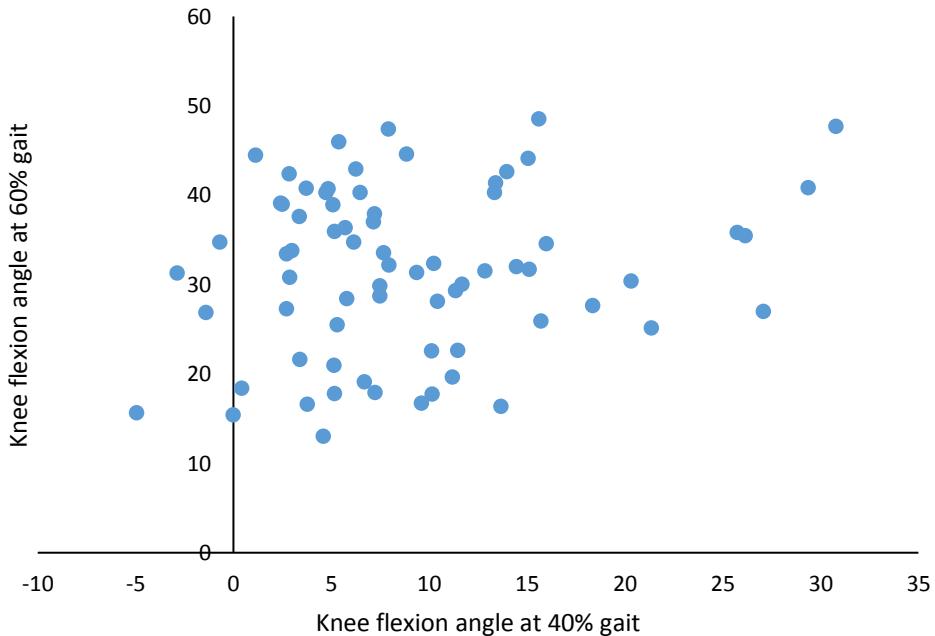


Figure 3.9 The relationship between the knee flexion angle at 40% of the gait cycle and the knee flexion at 60% of the gait cycle for a combination of OA and NP subjects.

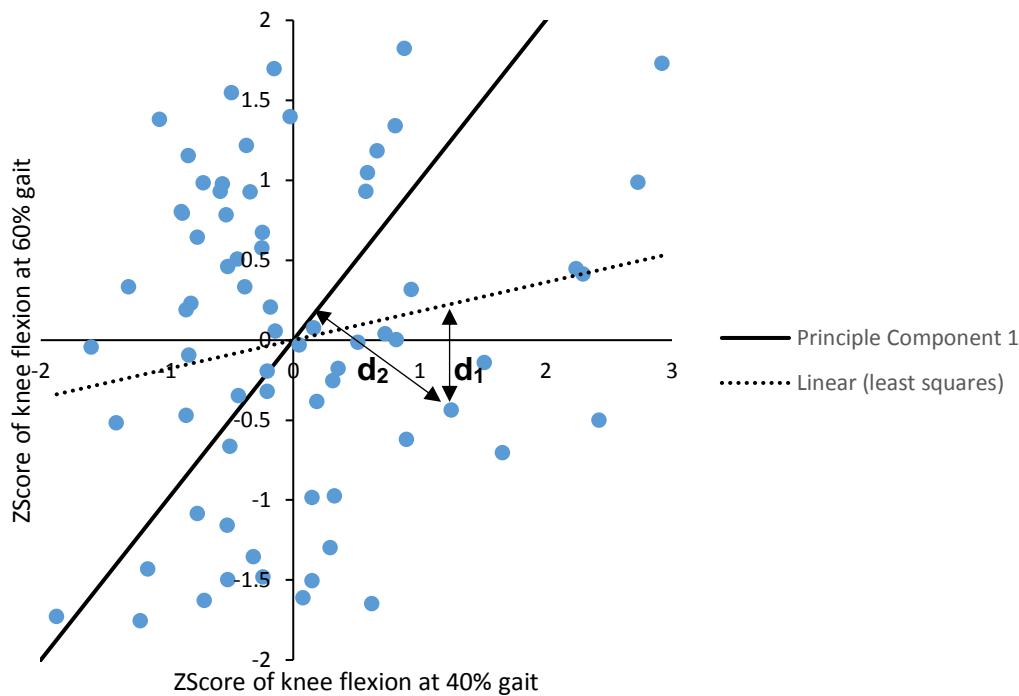


Figure 3.10 The relationship between the standardised variables from Figure 2.9. The linear least squares regression is shown as a dashed line, which minimises the total distance of each data point from the line in the y-direction only (d_1). The first principal component is also plotted as a solid black line and is intended as a direct comparison. As opposed to minimising the distance in the y-direction, it minimises the perpendicular distance between the line and the data points (d_2).

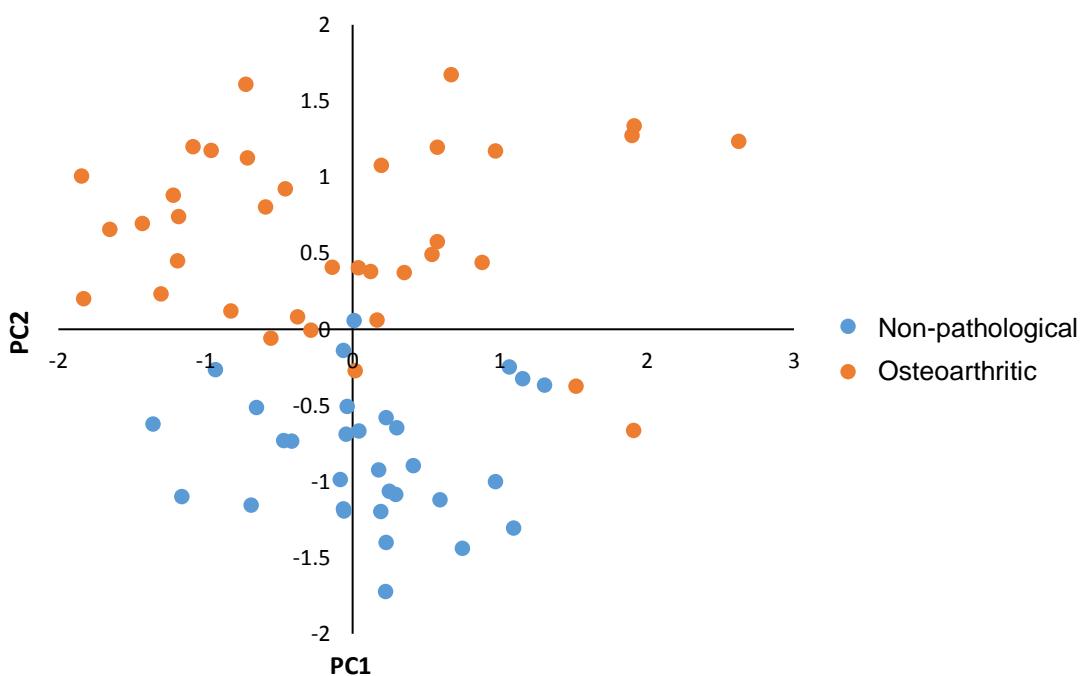


Figure 3.11 The data from Figure 3.12 replotted in the new axis system – PC1 and PC2. OA and NP subjects have now been plotted in separate colours, and it appears that the second principal component may be a better discriminator of OA.

3.4.5 Calculating Factor Loadings

Factor loadings, component loadings, or correlation coefficients are the correlation coefficients between the PC scores in the original axis with the PC scores in the new axis system. This can be calculated as:

$$\text{Loadings} = \text{Eigenvectors} \cdot \sqrt{\text{Eigenvalues}} \quad (3.24)$$

These are the equivalent of:

$$L = \begin{pmatrix} \text{corr}(s_1, z_1) & \text{corr}(s_2, z_1) & \cdots & \text{corr}(s_n, z_1) \\ \text{corr}(s_1, z_2) & \text{corr}(s_1, z_2) & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \text{corr}(s_1, z_n) & \cdots & \cdots & \text{corr}(s_n, z_n) \end{pmatrix} \quad (3.25)$$

Where $\text{corr}(s_1, z_1)$ represents the correlation coefficient of all of the scores s_1 along PC1 with the z-scores z_1 along the original standardised axis.

In this example

$$L_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{\sqrt{2}}{2} \\ 1 \\ \frac{1}{\sqrt{2}} \end{bmatrix} \sqrt{1.1808} = \begin{bmatrix} 0.768 \\ 0.768 \end{bmatrix} \quad (3.26)$$

$$L_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{\sqrt{2}}{2} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \sqrt{0.9182} = \begin{bmatrix} 0.678 \\ 0.678 \end{bmatrix} \quad (3.27)$$

$$L = \begin{bmatrix} 0.768 & 0.678 \\ 0.768 & -0.678 \end{bmatrix} \quad (3.28)$$

The amount of variance that each principal component represents of each of the standardised independent variables can then be calculated as the root of the z-score.

E.g. the first principal component represents $\sqrt{0.768} = 59.04\%$ of the variance of z_1 , and $\sqrt{0.768} = 59.04\%$ of the variance of z_2 . Note that these values are the same as the total variance in this 2D case, due to the nature of PCA on two standardised variables. In this special case, the first PC has a gradient of one, and variance represented is equal between the two standardised variables, z_1 and z_2 .

3.4.6 Expanding from 2D to N-Dimensions

The previous steps have shown step by step the calculations of a new, uncorrelated orthogonal axis system for two independent variables taken from the knee flexion angle waveform during gait. The first principal component represents the axis of primary variation within the standardised dataset, and the second is orthogonal to the first. In two dimensions, of course, there is only one vector which is orthogonal to PC1. If we were to include a third point of the waveform, and hence its z-score z_3 , it is easy to visualise a 3D scatter plot and a 3D eigenvector defining the primary axis of variance. Imagine the second PC is ‘constrained’ as being orthogonal to the first 3D vector; there is essentially one rotational degree of freedom for which this can occur. Imagine rotating it around this degree of freedom, until it defines the primary axis of variation along this axis. Now the third PC must be mutually orthogonal to the first two – for which there is only one solution. For an interactive 3D example of PCA the reader is directed to Powell (2015).

This is still only representing percentage points of the gait cycle, meaning that much data is being objectively discarded. PCA is therefore performed on all 100 data points. It isn’t possible to physically visualise a 100-dimension scatter plot, and therefore imagining a 100-dimension eigenvector which defines the primary axis of variance is less intuitive. Calculating the eigendecomposition of a matrix becomes exponentially more intensive as the number of dimensions increases, therefore, computer software such as MATLAB, which uses numerical methods, is required to find solutions.

3.4.7 Optimising the Calculation of Principal Components

The application of PCA using software developed by (Jones, 2004) was previously a very time-intensive process. Each waveform had to be individually collated for each patient from the output of the pointer method software and collated into a table. For n patients and v variables, it was, therefore, necessary to collate $n \times v$ individual waveforms into v different files, each containing the waveforms for n patients for a single variable. It was then necessary to perform the PCA code on each waveform individually. Additionally, if a new subject was collated and the user wished to calculate the PC scores of the subject

based on the original eigenvector definitions, it was necessary to run a second piece of code for all v variables individually. Therefore, to add five subjects, the code must be run $n \times v$ times.

It was therefore deemed appropriate to create software to automatically generate these tables.

The code ‘pcafilegenerator.m’, written in MATLAB, assumes that the file format outlined earlier is maintained. The code uses an identical patient spreadsheet format and compiles the waveforms of all the subjects listed on the spreadsheet into a single Excel file.

The Excel file contains a sheet for each variable. Within each sheet, each normalised waveform is compiled next to the subject ID for each subject. I.e. all of the knee flexion waveforms. This currently collates the ensemble average per subject; information of intra-subject variability is therefore lost.

The PCA code has also been adapted such that the whole process is repeated for each sheet of the MS Excel workbook. A results folder is automatically generated and an Excel workbook is created for each variable within the PCA. The prefix for the variable name is used (e.g. GRF_X_PCAresults.xlsx).

Within each workbook, eigenvalues, eigenvectors, factor loadings and filtered factor loadings are saved on separate sheets. Filtered factor loadings are binary filtered to ones or zeros based on the whether or not they exceed the threshold suggested by Comrey and Lee (2013) of >0.71 and <-0.71 .

3.4.8 Retention of Principal Components

The primary goal in the application of PCA is to objectively reduce the temporal dataset into much fewer discrete variables. For n independent variables, PCA results in n principal components. I.e. the dimensions of the transposed dataset remain same. It is often found, however, that the first few PCs represent the clear majority of variance within the dataset.

Figure 3.13 shows the relationship between the number of PCs retained and the total percentage of variance represented in the knee flexion angle waveforms displayed in Figure 3.8. Unlike the previous method adopted by Jones (2004), PCs weren't initially selected using Kaiser's rule (Kaiser, 1960). The rule, which retains factors (PCs) with an eigenvalue of greater than one, has historically been the most commonly used stopping rule (Jackson, 1991). It has, however, come under a vast amount of criticism, in part due to its tendency to retain too many components which often have no significance (Jackson, 1993, Ferré, 1995). In practice, this has very little effect on the resultant retained PCs when following the method presented by Jones (2004) because it is extremely rare for a PC to pass the second retention rule and to not have also passed the first.

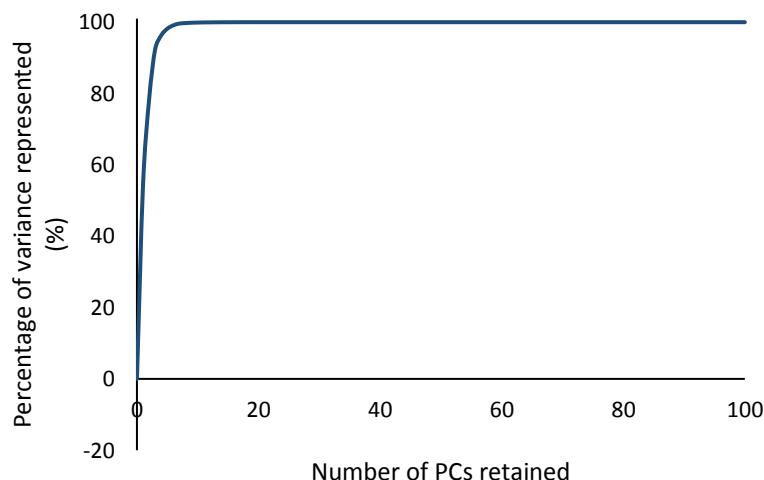


Figure 3.13 Example of the relationship between the number of PCs retained, and the cumulative total percentage variances represented by those PCs. It was commonly found throughout this study that the first 10 PCs represent almost 100% of the variance.

The second retention rule, adopted by Jones (2004), is to only retain PCs which have a factor loading above or below the threshold >0.71 and <-0.71 suggested by Comrey and Lee (2013). As discussed in previously, the factor loading is equivalently the r value between original variables and their new PC scores. The r^2 value is, therefore, the variance represented by that factor. This results in only factors which represent at least 50% ($\pm 0.71^2$) of the variance of one particular point in the gait cycle being considered.

Figure 3.14 shows the squared factor loadings plotted for the first five PCs against the original independent variables – the percentage points of the gait cycle. It can be seen from this graph that, using this threshold, PC4 would not be retained for analysis. PC1 would be interpreted as representing variance from 0-54% of the gait cycle as well as 96-100%, PC2 as 57-75%, and PC3 as 77-89%. This figure also shows the total variance represented at each point of the gait cycle when including these three PCs.

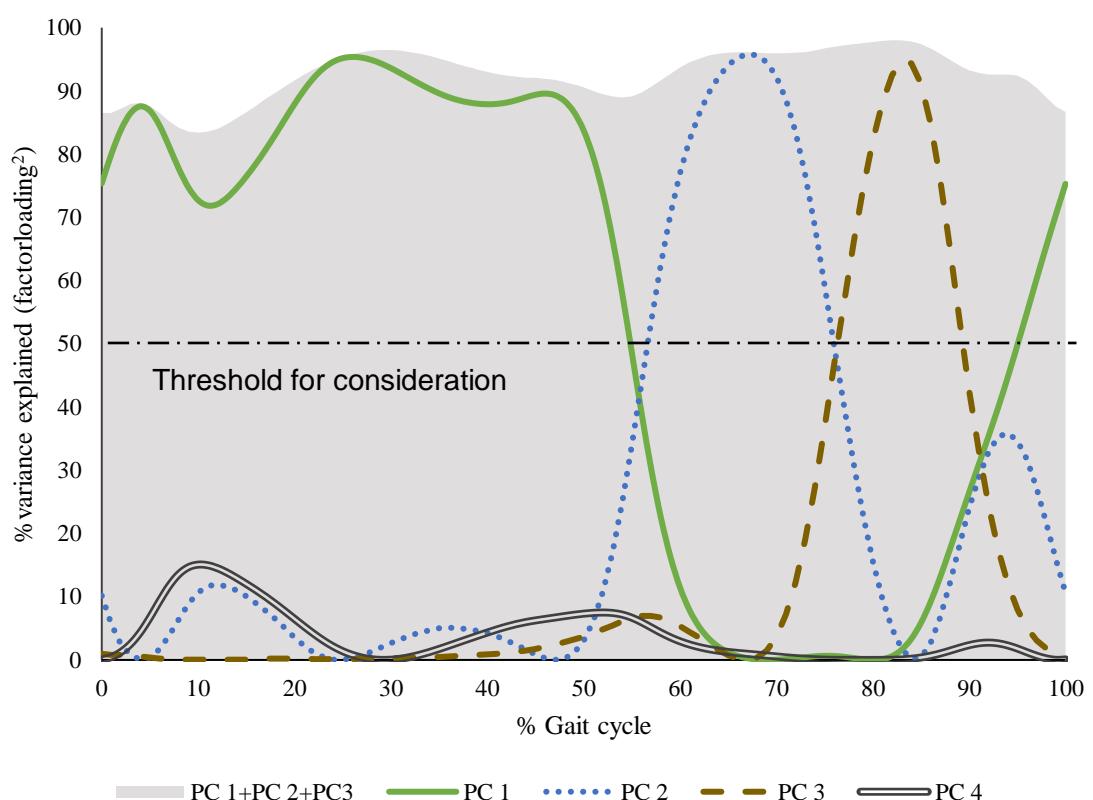


Figure 3.14 Example of the amount of variance explained, defined by the square of the factor loadings, by four PCs at each point in the gait cycle. It can be seen that PC4 did not pass the threshold for consideration. Also shown is the cumulative variance represented by PCs 1-3 at each point of the cycle.

3.4.9 Reconstructing Data Using PCs

Another method of interpreting the PCs is to reconstruct the data using only the retained components. To perform the reconstruction, it is necessary to transform the data back to the initial axis system. Figure 3.15 illustrates the process of reconstructing data from PC scores. To restore PC Scores back into the same domain as the z-scores, it is necessary to multiply the eigenvectors by the PC scores:

$$Z_r = \sum_{i=1}^n (E_i S_i) \quad (3.29)$$

Where Z_r is a vector of the reconstructed standardised waveform for one patient, n is the number of PCs retained, E_i is an eigenvector and S_i is the PC score associated with that eigenvector for that single subject.

It is then necessary to reverse the steps which created the zero mean unit standard deviation data. i.e.

$$v_r = Z_r \sigma + \mu \quad (3.30)$$

Where v_r is a vector containing the reconstructed waveform for one subject, σ is a vector containing the original standard deviation of each point of the gait cycle, and μ is a vector containing the original mean at each point in the gait cycle. Figure 3.16 displays the reconstruction of a flexion extension waveform for a number of subjects using three PCs. This aids in visual interpretation of what each PC actually represents. For example, the reconstruction using only PC1 appears that this component mostly represents changes in stance phase, and shows that subjects with a higher flexion angle at early stance phase are likely to maintain this increased flexion through to early swing phase. This could often be referred to as a fixed flexion gait and is common in subjects with OA. PC2 appears to represent mostly the variance during swing phase and shows that subjects with a reduced and delayed peak flexion angle during swing phase also tend to have less ROM during stance phase. PC3 represents mostly the variance at terminal swing, where in subjects with late peak knee flexions during swing are likely to have increased angular velocities in terminal swing.

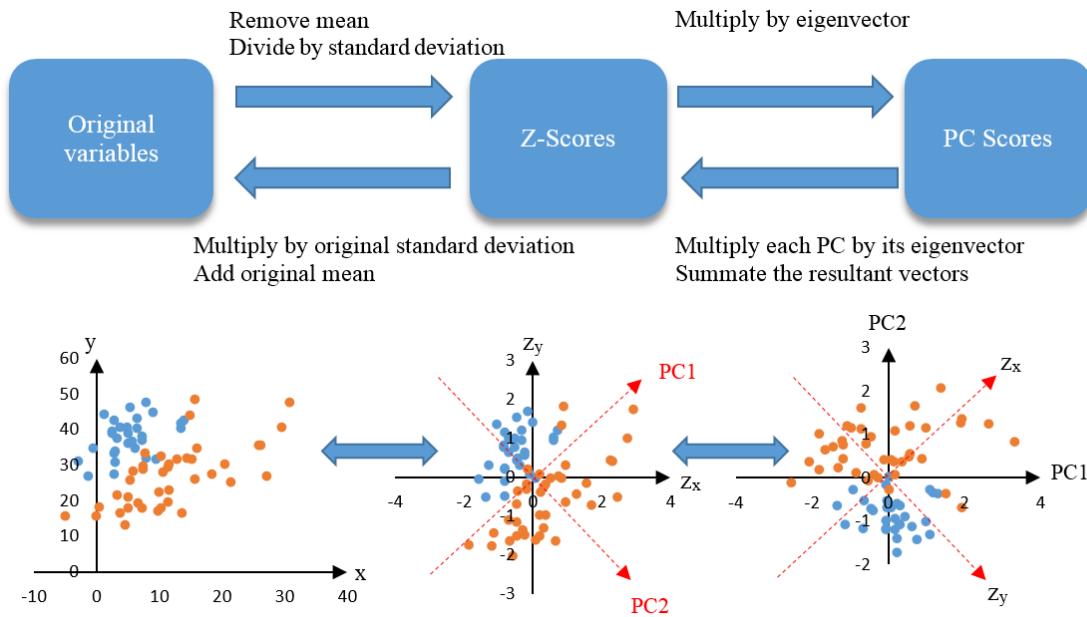


Figure 3.15 Illustration of the transformation from the original data, to z-scores, to PC scores, and how to transform or reconstruct the original data back from the PC scores. Below each stage is a 2D example of how this would look if there were only two independent variables within the PCA.

This demonstrates the value in using PCA reconstruction in order to help interpret components. Previous biomechanical studies have interpreted PCs by identifying subjects with the highest and lowest PC values and analysing their raw waveforms (Reid *et al.*, 2010, Kirkwood *et al.*, 2011, Deluzio and Astephen, 2007). The issue with this technique is that, say the subject had both a high value of PC1 and a high value of PC2, it wouldn't be possible, by looking at the raw waveform, to discern which PC characterises specific waveform features. This criticism has also been made by Brandon *et al.* (2013), whom instead opt to reconstruct the waveform of the subjects with the highest and lowest PC values using only the PC of interest.

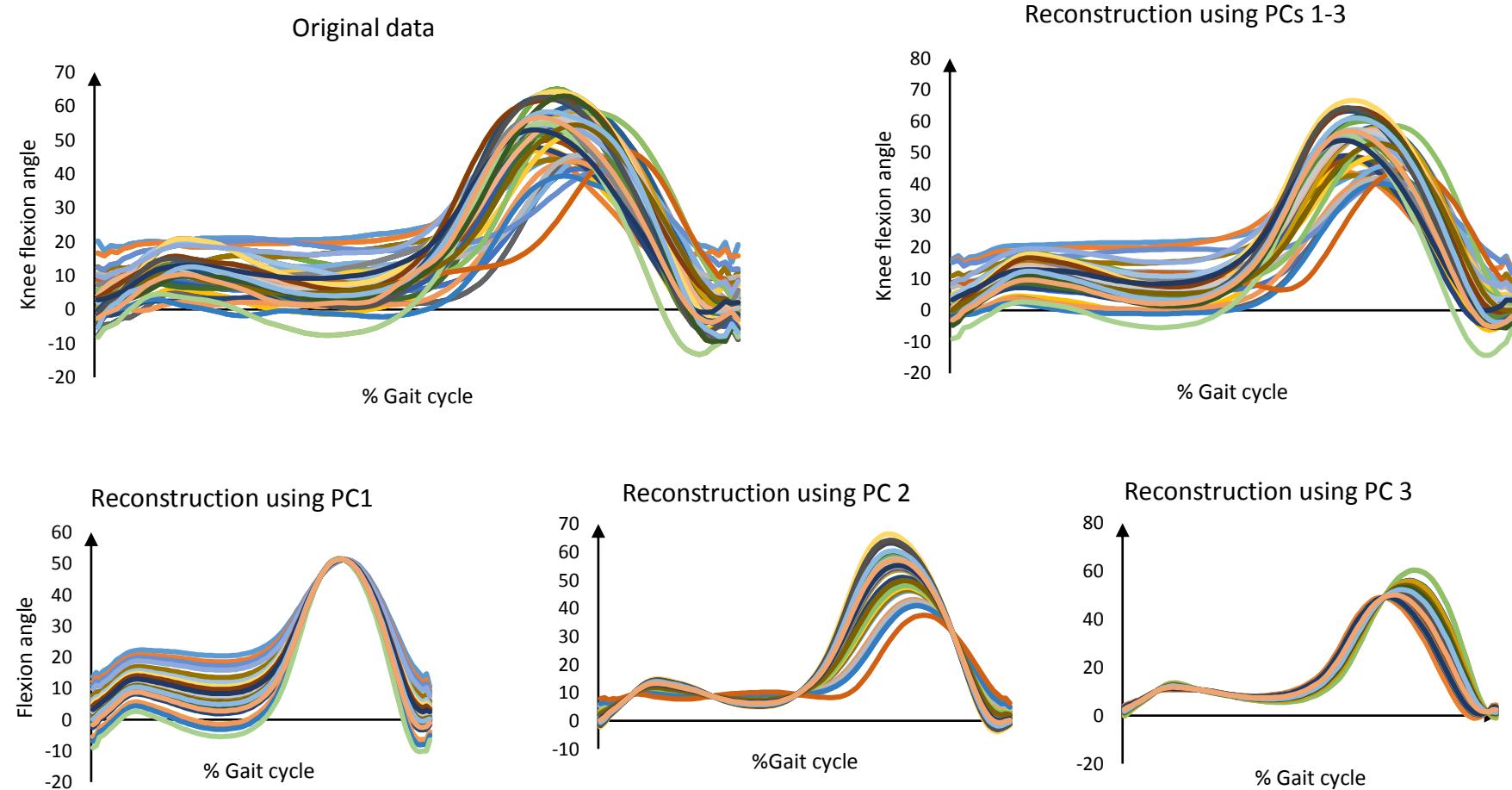


Figure 3.16 An illustrative demonstration of the application of PC reconstruction to knee flexion/extension waveforms. The original data (top left) for all subjects is plotted, alongside the reconstruction using PC1 (bottom left), PC2 (bottom centre), and PC3 (bottom right). The reconstruction of the original data using all three PCs (top right) displays the visual similarity between features in the original data and the reconstruction.

3.5 The DST Classifier

3.5.1 Defining K

As previously discussed, the DST classifier uses a modified activation function, for which the steepness is defined by k :

$$(cf) = \frac{1}{1 + e^{-k(v-\theta)}} \quad (3.31)$$

It was also discussed that the factor k is the equivalent of the weighting in a neural network. Jones (2004) explored two definitions of k , one relating to the correlation coefficient and one relating to the standard deviation (see Section 2.5.4). It was concluded that the best definition of k for classifying the test data set was the person's correlation coefficient between the variable value and the category:

$$k = \text{corr}(v_i, L) \quad (3.32)$$

Where v_i is the independent variable, L is the column of binary data labels (e.g. either 1 or 0 depending on OA or NP).

The value of k is equivalent in function to the weighting used within a neural network (see Section 2.5.7). The weighting can be considered to have the following primary function: variables that are more likely to be important in the classification can be assigned greater contribution to the classification result. Considering this function of the weighting, the use of the correlation coefficient makes sense. If there is not much separation between the two groups, k would closer to 0, and if the groups were very well separated, the correlation coefficient would approach 1 and hence the gradient is steeper.

Another function of the weighting within a neural network is that, when the inputs have different scales or units, the weighting should adjust for this during the backpropagation. Therefore, theoretically, if you divided one input variable by five, for example, the neural network would reach the same optimal solution; however, the weights attached to that variable would be five times larger.

One notable shortcoming in using the correlation coefficient to define the gradient of the sigmoid curve is that it does not adapt to different scales of input data. Figure 3.17A shows a sigmoid function defined based on synthesised example training data of NP and OA subjects for peak knee flexion angle during gait. Figure 3.17B shows how the sigmoid curve is defined if the variable were to be scaled by a factor of 1/5. It is clear that variable values now convert to very different confidence values now after the scaling.

To expand further, say the correlation coefficient was 0.5.

$$(cf) = \frac{1}{1 + e^{-0.5(v-\theta)}} \quad (3.33)$$

If we consider the value of v required such that $cf = 0.8$

$$0.8 = \frac{1}{1 + e^{-0.5(v-\theta)}} \quad (3.34)$$

$$1 + e^{-0.5(v-\theta)} = \frac{1}{0.8} \quad (3.35)$$

$$e^{-0.5(v-\theta)} = 0.25 \quad (3.36)$$

$$-0.5(v - \theta) \ln(e) = \ln(0.25) \quad (3.37)$$

$$(v - \theta) = \frac{\ln(0.25)}{-0.5} \quad (3.38)$$

$$(v - \theta) = 2.77 \quad (3.39)$$

Therefore, the confidence factor is 0.8 when the data value is 2.77 greater than the group mean. This is completely independent of the scale of the input data.

The other considered definition of k was 1/STD. This definition doesn't suffer from the aforementioned issue, as the standard deviation is scaled with the data. The use of the standard deviation effectively means that if the spread of the data is greater, the sigmoid curve will be less steep. A problem with this technique is shown in Figure 3.18. The two datasets shown have exactly the same standard deviation; however, the groups

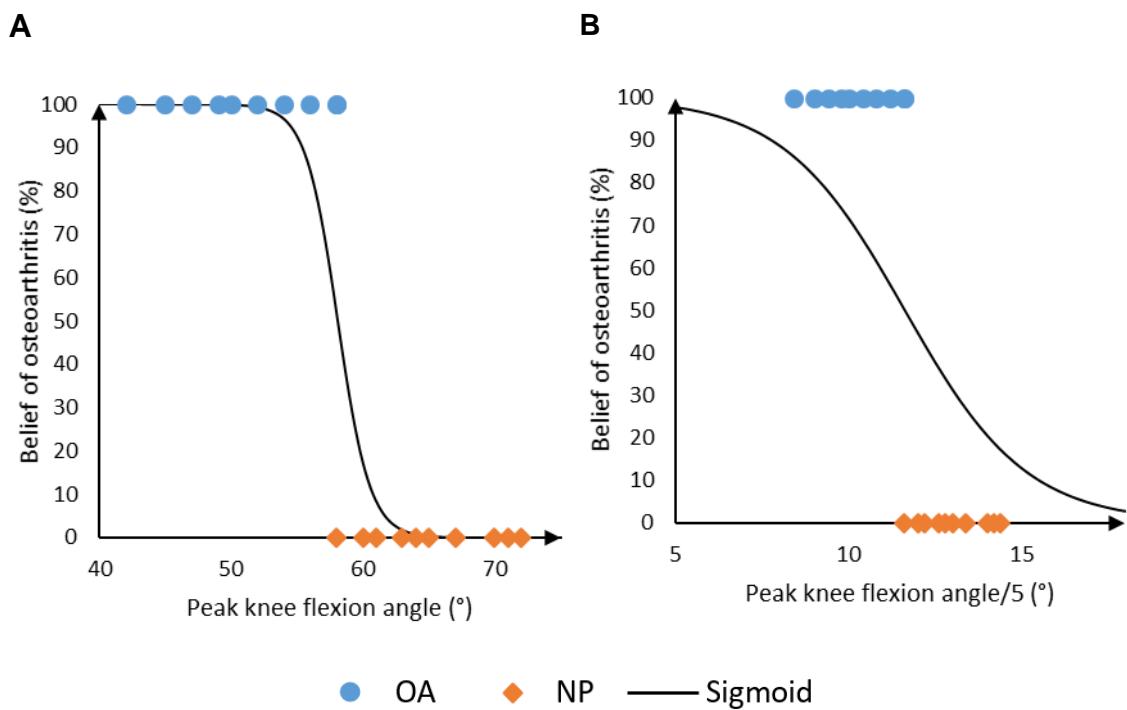


Figure 3.17 Demonstrating the issue with defining the sigmoid curve within the DST classifier using only the correlation coefficient. A) Shows the sigmoid curve defined using k_c for a synthesised example of peak knee flexion during gait for OA and NP subjects. B) Shows the very same data scaled by 1/5. It is apparent that these data points now convert to much lower belief values as a result of the change in scale.

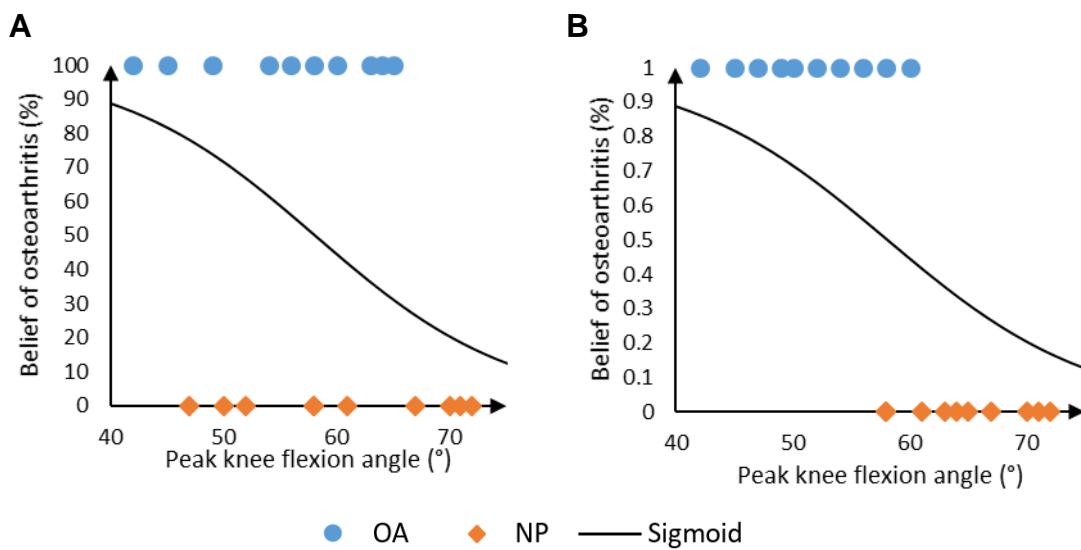


Figure 3.18 Demonstrating the drawback of using the standard deviation alone to define the steepness of the sigmoid curve. The datasets plotted within A and B have the same standard deviation, however, it is clear that the two groups are more separated within B. Intuitively, the sigmoid curve should, therefore, be steeper in this case.

are clearly more separated within Figure 3.18B than Figure 3.18A. This technique is therefore not considering the separation of the two groups, only the spread of the data as a whole. Hence, it is not achieving the primary function of the weighting. It is also worth noting that it is effectively achieving the same as the standardisation of data to have a unit variance.

$$(cf) = \frac{1}{1 + e^{-k(v-\theta)}} \quad (3.40)$$

$$k = \frac{1}{\sigma_v} \quad (3.41)$$

$$(cf) = \frac{1}{1 + e^{-\frac{(v-\theta)}{\sigma_v}}} \quad (3.42)$$

Where $\frac{(v-\theta)}{\sigma_v}$ is the equivalent of the difference of v from θ , as a fraction of the standard deviation. When θ is defined as the mean of the variable, \bar{v} , then $\frac{(v-\bar{v})}{\sigma_v}$ is the equivalent of converting the variable to unit variance.

A compromise to both techniques might be to combine them and calculate the correlation coefficient as:

$$k = \frac{\text{corr}(v_i, L)}{\sigma_v} \quad (3.43)$$

Where v_i is the independent variable, and L is the column of binary data labels (e.g. either 1 or 0, depending on OA or NP) and σ_v is the standard deviation of the independent variable.

One potential consequence this may have is that by multiplying by the correlation coefficient, the steepness is always reduced as the correlation coefficient falls without boundary $-1 \leq r \leq 1$. For example, within the training data used by Jones (2004), the average absolute value for k when defined using the correlation coefficient was 0.36, and when using the standard deviation was 0.28. Had the two methods been combined, the average steepness would have been around 0.1.

It is proposed that a constant should be introduced which accounts for this reduction in steepness which is, as a result of weighting, using the correlation coefficient. i.e.

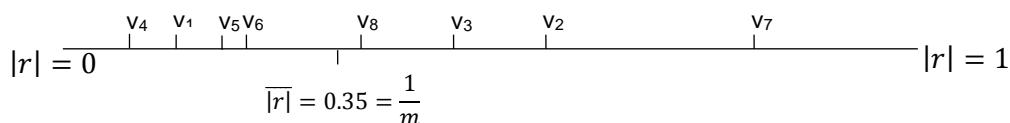
$$k = \frac{m \times \text{corr}(v_i, L)}{\sigma_v} \quad (3.44)$$

Figure 3.19 shows the absolute correlation coefficients for several variables in a synthesised example. It is evident that the average absolute correlation coefficient \bar{r} is 0.35. On average, by weighting using the correlation coefficients, the value of k would become $2.6 (\frac{1}{0.35})$ times smaller. If the constant m is defined as:

$$m = \frac{1}{\bar{r}} \quad (3.45)$$

Where \bar{r} is the average absolute correlation coefficient of all of the variables with their class label, the average weighting $\bar{r}m = 1$.

A



B

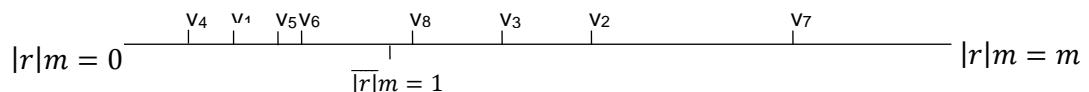


Figure 3.19 A) The magnitude of the Pearson's correlation coefficient for eight different variables plotted on a scale between 0 to 1. The average magnitude of r is 0.35. Here the value m is derived.

B) The same values are now each multiplied by m and replotted between 0 and m . The average magnitude of the correlation coefficient is now 1.

3.5.2 Defining Theta

The value theta within the sigmoid activation function of the Cardiff Classifier was briefly introduced in Section 2.5.4 and defines the point at which the confidence factor is 0.5. This can be proven as:

When

$$(cf) = 0.5 = \frac{1}{1 + e^{-k(v-\theta)}} \quad (3.46)$$

Thus:

$$e^{-k(v-\theta)} = 1 \quad (3.47)$$

$$-k(v - \theta) = 0, \quad k \neq 0 \quad (3.48)$$

Hence when $(cf)=0.5$

$$v = \theta \quad (3.49)$$

Therefore the variable θ is the value of v for which $(cf) = 0.5$ and the evidence supporting the belief of OA is equal to the evidence supporting the belief of NP function (Jones, 2004). Beynon (2005) stated that the value of θ should be chosen using the following criteria:

1. It lessens the distance to all the criterion measurements (all values for that variable)
2. It is not biased towards either class label

By this reasoning, Beynon chose to define θ as the mean \bar{v} of all the values for that variable, and hence under this recommendation was also used by Jones (Jones *et al.*, 2008). It holds to reason that, when there are an equal number of subjects within the two groups, this mean value achieves both criteria.

Figure 3.20 uses data from this chapter to demonstrate a problem with this definition of θ when the number of subjects in each group of labelled data is different. The figure shows data from 85 OA subjects and 38 NP subjects. The value of θ has been

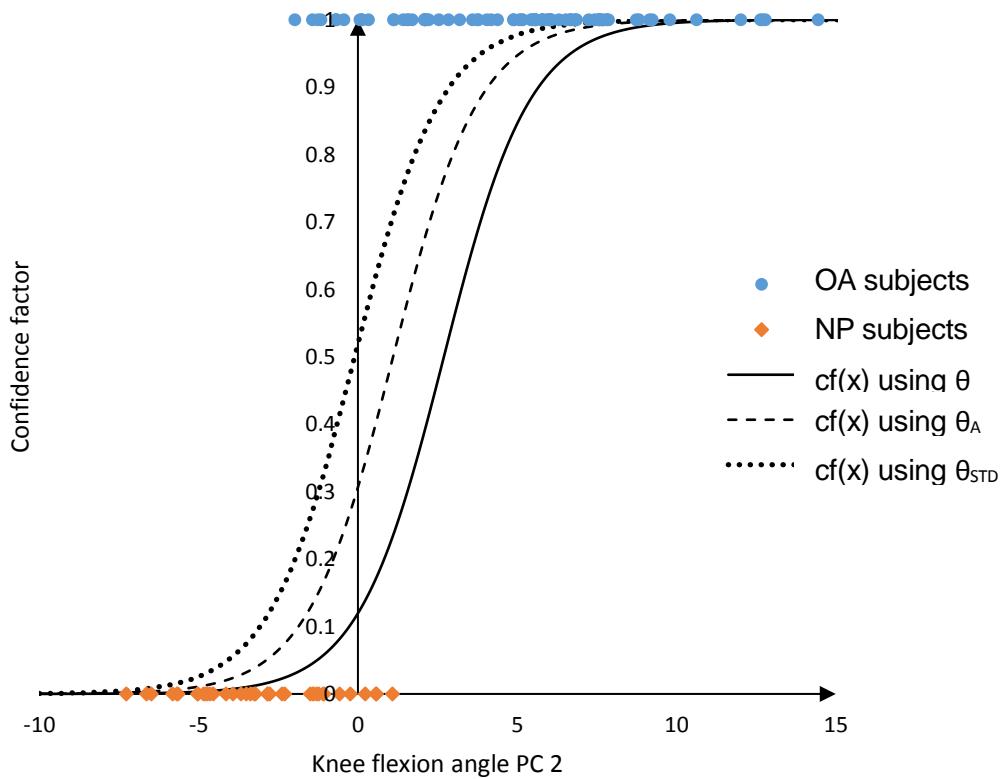


Figure 3.20 Comparing the effect of the new proposed definitions of the value of θ in the definition of the sigmoid curve. The sigmoid curve is plotted using the original theta value, θ , and the two proposed alternatives, θ_A and θ_{STD} .

defined as previously suggested. While using this value has achieved criterion 1, it has failed at achieving criterion 2: the point at which the confidence factor is 0.5 appears very much biased towards to the osteoarthritic class label. This is not surprising because the larger group will have a greater influence in the value of the mean \bar{v} .

It is hereby suggested that, in the case of imbalanced data class sizes, the point equidistance from the mean of each group should instead be taken. I.e.

$$\theta_A = \frac{\bar{v}_{OA} + \bar{v}_H}{2} \quad (3.50)$$

This definition now fulfils criterion 2, as the value of θ is equidistant from the mean of each group and therefore not biased towards either data class. Figure 3.20 shows the updated sigmoid curve using the new theta definition. Using the old definition, 25 OA subjects incorrectly have a confidence factor of less than 0.5, and 0 NP subjects have a

confidence factor above 0.5. Using the new definition, θ_A , only 12 OA subjects have a confidence factor below 0.5 and still 0 NP subject have a confidence factor above 0.5.

This adjusted definition no longer fulfils criterion 1 – the sum of the absolute distances of all the data points from θ has in fact increased. The average distance of all the points from the mean within each category however is much more fairly distributed; although not equal due to the difference in standard deviation.

A further suggested modification to the definition of the midpoint θ is to account for heterogeneity of variance between the two groups. It can be seen in Figure 3.20 that there is greater variance of this variable within the OA group in comparison to the NP group. Assuming that both distributions were approximately parametric, this would result in a greater misclassification rate in the group with the larger spread. If the midpoint were moved proportionally closer to the NP group with the smaller spread, the benefits of less misclassified OA subjects would outweigh the costs of more misclassified NP subjects.

A further definition of the midpoint is proposed:

$$\theta = \theta_{NP} + (\theta_{OA} - \theta_{NP})\left(\frac{\sigma_{NP}}{\sigma_{NP} + \sigma_{OA}}\right) \quad (3.51)$$

Where θ_{NL} is the average variable value of the NP subjects, θ_{OA} is the average of the OA subjects, σ_{NP} is the standard deviation of the NP subjects and σ_{OA} is the standard deviation of OA subjects.

3.5.3 Defining the Uncertainty Boundaries:

The next step in the classification is to convert the confidence factor representing each variable into a BOE using the Dempster-Shafer Theory. The BOE consists of three belief functions:

1. $m(OA)$ - The degree of belief in OA
2. $m(NP)$ - The degree of belief in NP (i.e. healthy) function
3. $m(\Theta)$ - The associated ignorance (uncertainty)

The relationship between the belief functions and the confidence factors followed the work of Safranek *et al.* (1990), where:

$$m(OA) = \frac{B}{1-A} cf(v) - \frac{AB}{1-A} \quad (3.52)$$

$$m(NP) = \frac{B}{1-A} cf(v) + B \quad (3.53)$$

$$m(\Theta) = 1 - m(OA) - m(NP) = \frac{1-A-B}{1-A} \quad (3.54)$$

Where A represents the dependence of the $m(OA)$ on the confidence factor, B represents the maximal support which can be assigned to either $m(OA)$ or $m(NP)$. Jones (2004) states that the values of A and B should be assigned based on knowledge of the upper Θ_U , and lower Θ_L boundaries of uncertainty. These were related to upper Θ_U , and lower Θ_L boundaries of uncertainty as follows:

$$A = \frac{\Theta_U - \Theta_L}{1 + \Theta_U - 2\Theta_L} \quad (3.55)$$

$$B = 1 - \Theta_L \quad (3.56)$$

It is, therefore, possible to assign different values for A and B depending on the variable. For example, the user might have previous knowledge of the level of ignorance within a particular variable. In the majority of the following research using the DST classifier, the same limits have been applied to each individual variable. This is likely because the classifier has been used as an objective tool, and the manual selection of A and B for each individual variable introduces a level of subjectivity. Furthermore, defining Θ_L and

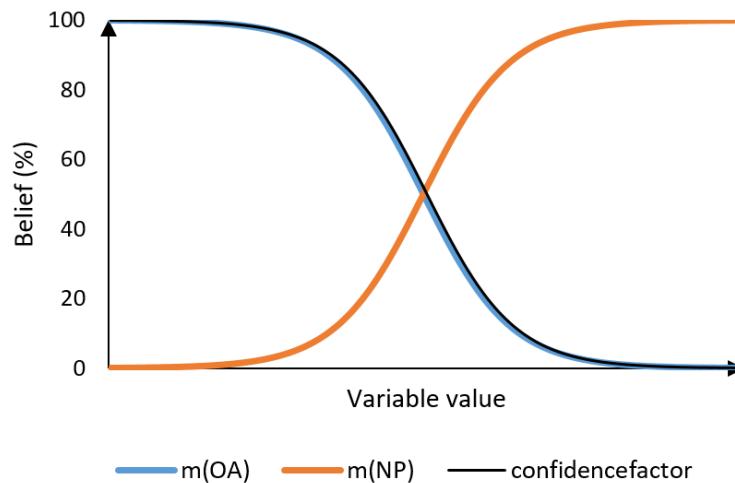


Figure 3.21 The conversion between confidence factor and belief functions $m(OA)$ and $m(NP)$ in the case of zero uncertainty/ignorance.

Θ_U based on expert knowledge is not an instinctive process. For this reason, the same values of A and B will also be applied to every variable.

Another way to understand the effect of the assignment of A and B is to first consider the situation where there is zero ignorance, as shown in Figure 3.21. In this instance, the belief $m(OA)$ is equivalent to the confidence factor, and $m(NP)$ is one minus the confidence factor. As previously discussed OA in a disease with a vast number of symptoms and is generally diagnosed through the collection of multiple pieces of evidence. Furthermore, all biomechanical variables are interdependent and each one considered alone only accounts for a small piece of a much larger picture. It is, therefore, unrealistic to expect that any one piece of evidence would result in a 100% belief of OA or NP.

Figure 3.22 introduces uncertainty, or ignorance. The maximal degree of belief any one piece of evidence can contribute is 70% and consequently, it would require further pieces of evidence to approach a belief of 100%. The upper Θ_U , and lower Θ_L boundaries of uncertainty have been set to 0.3, or 30% in this example; therefore, A=0 and B=0.7. Table 3.3 displays an example of two different input variables, gait velocity and peak flexion angle, we had two different items of evidence, gait velocity and peak flexion angle, for which the confidence values have already be found as 0.5 and 0.6 respectively.

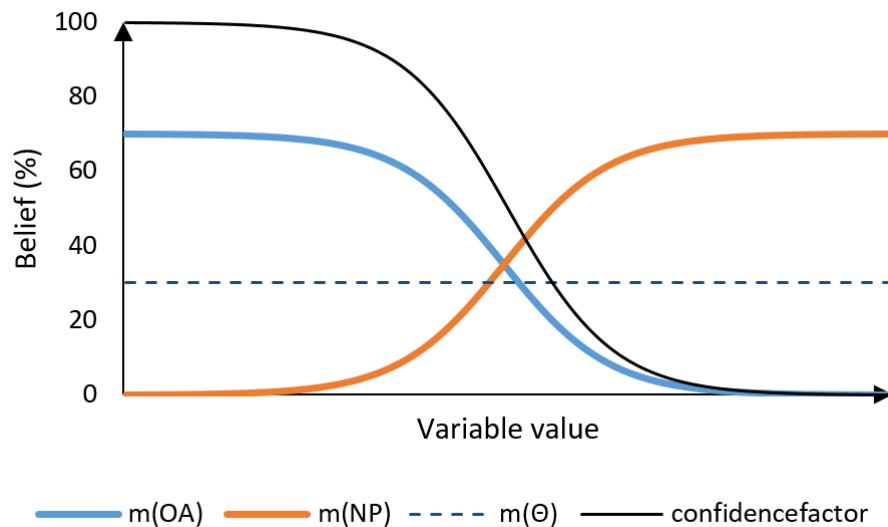


Figure 3.22 The conversion from confidence factor to a BOE when both upper and lower boundaries of uncertainty are 0.3 (30%).

Notice how the belief attributed to the peak flexion angle (which by itself would not favour either a belief of NP or uncertainty) has both increased the belief $m(OA)$ and the belief $m(NP)$, while decreasing uncertainty. Also, notice that while both values have increased, the ratio $m(OA):m(NP)$ has decreased. This result is challenging for the following reasons:

- i) There has been a marked decrease in uncertainty: it doesn't seem intuitive that a belief function which favours neither hypothesis would remarkably reduce the uncertainty of the classification. If there were multiple pieces of evidence which equally supported both OA and NP function, and these truly were the only two possibilities, in practice our ignorance or uncertainty would not decrease.

Table 3.3 Example of the combination of the bodies of evidence from two different variables.

	$cf(v)$	$m(OA)$	$m(NP)$	$m(\Theta)$
Gait velocity	0.6	0.42	0.28	0.3
Peak flexion angle	0.5	0.35	0.35	0.3
Combined	NA	0.5	0.38	0.12

- ii) The belief in $m(NL)$ increased by a relatively larger amount than $m(OA)$: if we were initially quite sure of OA function due to the gait velocity and we then find the peak knee flexion angle doesn't support our hypothesis, either way, it might seem intuitive that this would then decrease our relative belief of OA. However, what if there was a high amount of error in the measurement of peak knee flexion angle? In this situation, the sigmoid curve might have a very small steepness coefficient k , and hence a large range of values of v would result in a confidence factor approaching 0.5. If multiple pieces of evidence, where errors are too large to discern differences were additionally combined, not only would uncertainty approach zero, but both $m(OA)$ and $m(NL)$ would approach 0.5. This is counter-intuitive when there is the possibility that the aforementioned variables might just have a very high level of error.

Figure 3.23 shows the same example as the previous figure, with the addition of an upper boundary of uncertainty Θ_U , 0.5. Notice that the uncertainty increases as the confidence factor approaches 0.5, while both $m(OA)$ and $m(NP)$ change at an increased rate. Notice also that there is still a region where both $m(OA)$ and $m(NP)$ can be assigned simultaneously.

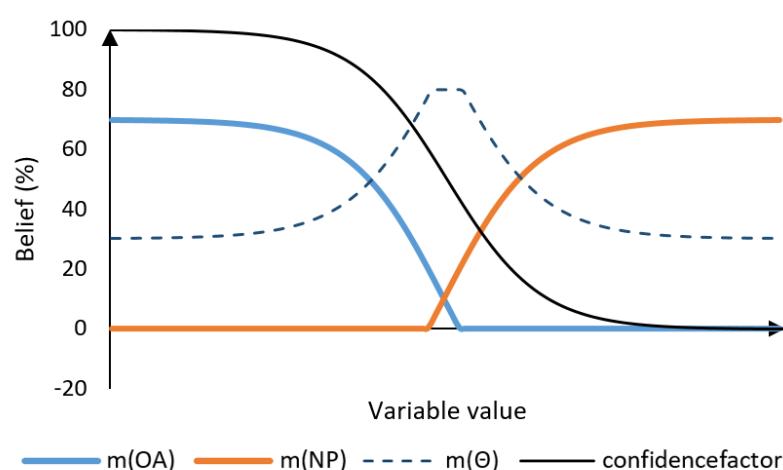


Figure 3.23 The conversion from confidence factor to a BOE when the upper and lower boundaries of uncertainty are 0.5 (50%) and 0.3 (30%) respectively.

As mentioned previously, in the absence of a clear objective way of selecting these boundaries, Jones (2004) investigated the effect that different uncertainty boundaries had on the final classification of a training body of NP and OA subjects. Jones argued that, for that specific dataset, the optimal boundaries were $\Theta_U = 1$ and $\Theta_L = 0.8$. These boundaries are plotted within Figure 3.24. It can be seen from the figures that the value of Θ_U affects the rate of change of $m(OA)$ and $m(NP)$ as $cf(v)$ approaches 0.5, and the value of Θ_L affects the maximum belief one piece of evidence can contribute.

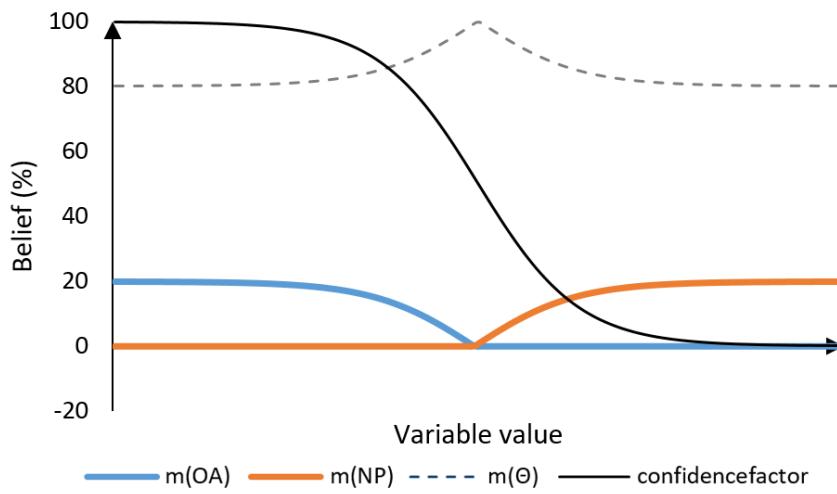


Figure 3.24 The conversion from confidence factor to a BOE when the upper and lower boundaries of uncertainty are 1 (100%) and 0.8 (20%) respectively, as used in the work of Jones (2004).

3.5.4 Evaluation of Classification Error

To assess the reliability of the classification, it is necessary to evaluate the error. It is widely known that estimations of errors of data classification techniques are biased if they are estimated from the same set of data that was used to define the classification parameters (Michie *et al.*, 1994). To test classification robustness, it is, therefore, necessary that the data used to test the classification accuracy was not part of the training body. Previous work has utilised the Leave-One-Out (LOO) cross-validation algorithm (see Figure 3.25) in which $(n-1)$ cases are used to train the classifier and hence define the control parameters, and these parameters are used to classify the remaining subject (Beynon *et al.*, 2006). This process is repeated n times until each subject has

been used as test data and classified. The classification accuracy is then determined by whether the greater belief value, $m(OA)$ or $m(NL)$, matches the actual class label.

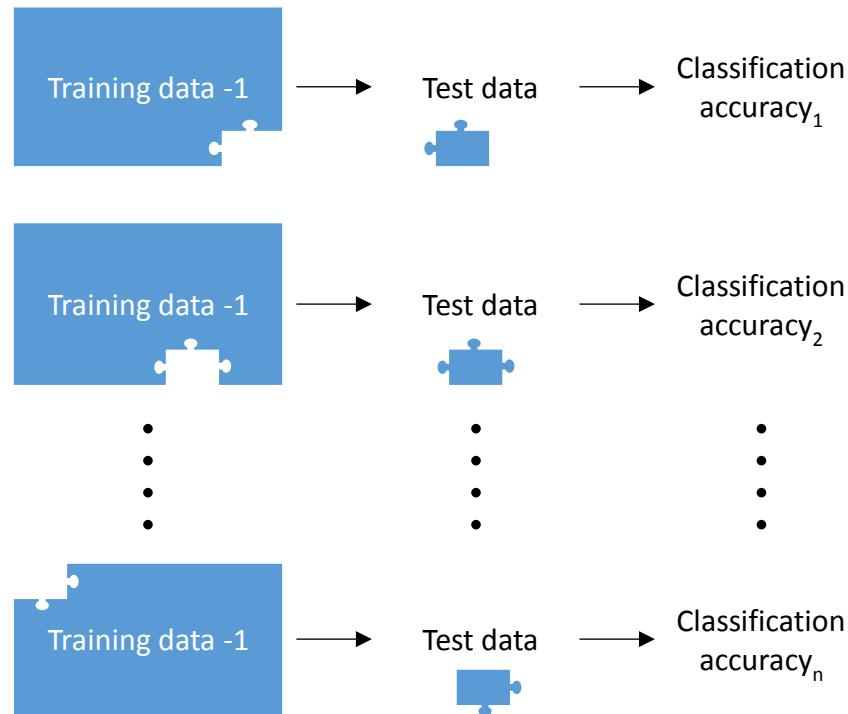


Figure 3.25 Illustration of the Leave-One-Out (LOO) cross-validation technique. A subject is removed from the training body, and this becomes the test data. The control parameters defined from the training body are used to calculate the belief values of the test data, and these belief values are compared to the class label to determine the accuracy. This is repeated n times until each subject has been used as test data.

A further test for classification robustness is to adopt a leave- p -out approach. This is methodologically the same as the LOO approach, except p subjects are retained as test data as opposed using new subjects. One of the challenges using this approach is that, for a dataset with n subjects, where p subjects are retained for the test set, there are:

$$\binom{n}{p} = \frac{n!}{p!(n-p)!} \quad (3.57)$$

Different combinations. I.e. if we had a dataset of 100 subjects, and we chose to do a leave-5-out, there are 75287520 combinations of which the training body can be comprised. It is therefore computationally expensive to classify using every single permutation. Instead, each sample will be randomly generated a large number of times.

3.6 Results and Discussions

3.6.1 Classifying Using the Same Variables and Principal Components as Jones (2004)

The new classification pipeline was utilised to increase the training body cohort to a total of 85 OA subjects and 38 NP subjects. This represents an additional 65 OA and 16 NP subjects. Firstly, the same principal components defined by Jones (Jones, 2004) were used to calculate the PC scores of the additional subjects. The data was classified using each definition of control variables k and θ . The results of the in and out-sample classification accuracy, alongside the in and out-sample objective function, are shown in Table 3.4.

As predicted, the original definition of θ , (θ_O), results in poor classification accuracy, and is significantly improved by the defining using θ_A . This is because of an uneven group size biasing the definition of the midpoint of the sigmoid curve. There is a slight improvement in accuracy when using θ_S as opposed to θ_A .

Table 3.4 Classification results of 85 OA and 38NP subjects using the variables and PC defined by Jones (2004), comparing definitions of control variables k and θ

K definition	Theta definition	In-sample accuracy (%)	LOO accuracy (%)
K_c	θ_O	84.6	84.6
	θ_A	94.3	93.5
	θ_S	94.9	94.3
K_s	θ_O	80	77.2
	θ_A	90.1	87
	θ_S	92.2	90.2
$K_{c/s}$	θ_O	86.1	86.2
	θ_A	94.3	93.5
	θ_S	94.9	94.3

Consistent with Jones (2004), k_s performed slightly worse than k_c and $k_{c/s}$. The performance of $k_{c/s}$ was very similar to that of k_c . The two best-performing techniques are plotted in Figure 3.26—using the STD adjusted midpoint θ_S , and using k_c or $k_{c/s}$ to define the steepness of the sigmoid. There was greater uncertainty when using $k_{c/s}$, with misclassified subjects also being closer to the vertex, representing uncertainty. Despite this, and considering similar classification accuracies between the two definitions, the definition of $K_{c/s}$ to define the steepness of the sigmoid curve is recommended, as it forgoes the aforementioned bias towards variables with a large order of magnitude.

Table 3.5 shows a comparison of the ranking of variables between this study and that of Jones *et al.* (2008) when using the θ_S and $K_{c/s}$ control parameters. The four highest ranked variables are the same but they have, however, changed order. Rankings for PCs of the internal/external rotation and external rotation angles have dropped considerably, and the knee measurements now fall within the top ten ranked variables.

Table 3.5 Comparison of the rankings of input variables using the 85OA and 38NL within this study to those reported for the 22OA and 20NL subjects within (Jones *et al.*, 2008). Only information on the top 10 ranked variables was available.

Rank	Input Variable	Classification accuracy (%)	Rank within Jones <i>et al.</i> (2008)
1	GRF Ant/posterior PC1	93.5	3
2	Knee Flexion Angle PC2	90.2	4
3	GRF Vertical PC2	90.2	1
4	Cadence	80.5	2
5	ML Knee Width	79.7	
6	GRF Vertical PC1	71.5	
7	Stance percent	70.7	7
8	Knee Depth	69.9	
9	Average Knee Girth	68.3	
10	GRF Vertical PC1	68.3	
11	BMI	68.3	9
12	GRF Ant/posterior PC2	65.9	
13	Knee Ad/Abduction Angle PC3	54.5	8
14	Knee Int/external Angle PC1	54.5	5
15	Knee Ad/Abduction Angle PC2	53.7	
16	GRF Ant/posterior PC3	50.4	10
17	Knee Ad/Abduction Angle PC1	49.6	6
18	Knee Flexion Angle PC1	49.6	

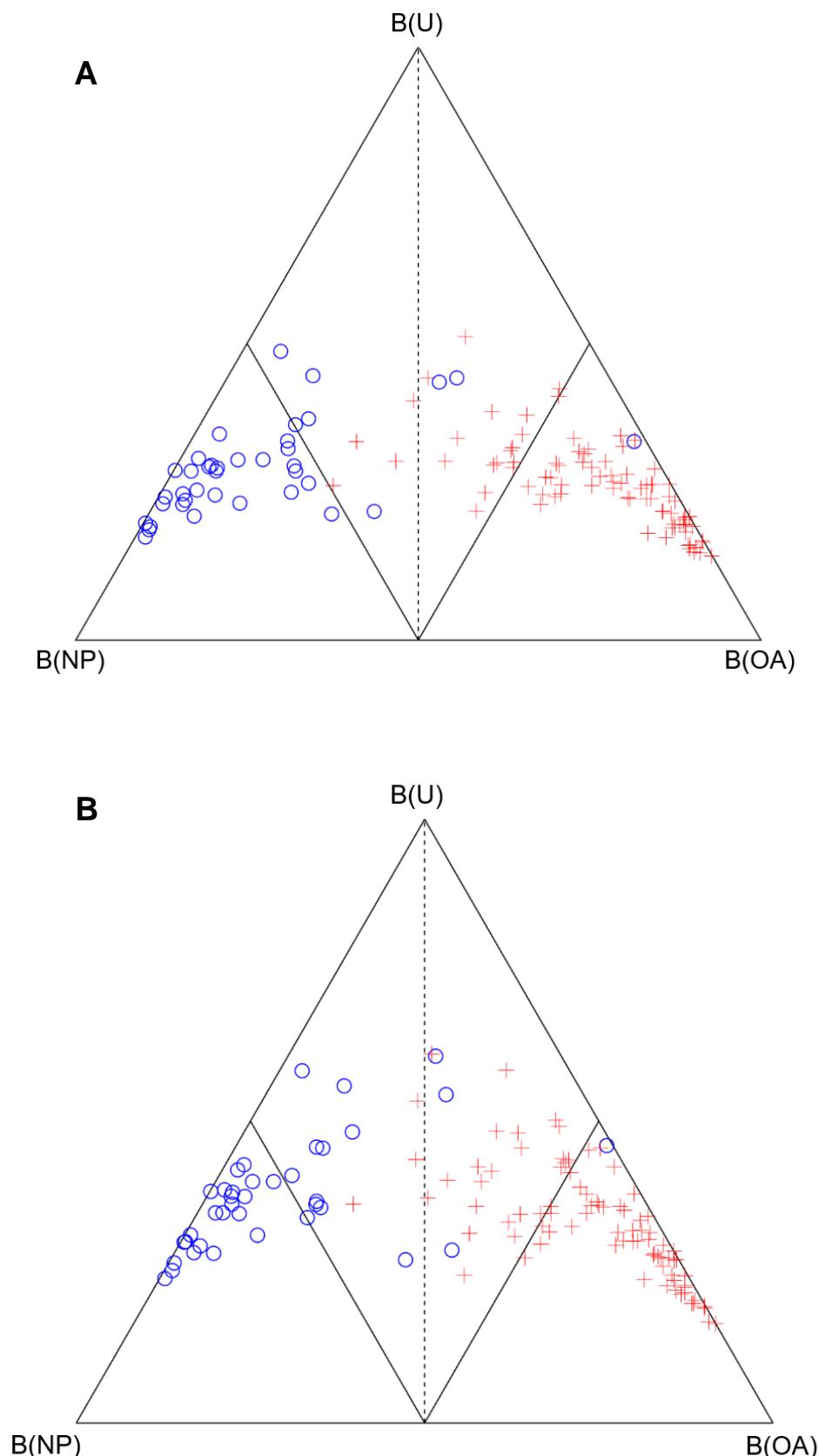


Figure 3.26 Simplex plots to illustrate the LOO classification of 85 OA (red cross), and 38 NP (blue circle), subjects. A) Using the correlation coefficient to define k , and the STD adjusted midpoint to define θ . B) Using the new adjusted correlation coefficient to define k , and the STD adjusted midpoint to define θ .

3.6.2 Updated Principal Components

As there is a significantly larger cohort of OA and NP subjects, the original PCs defined by Jones may not be the optimal components in describing the variance within the new dataset. Therefore, the total of 85 OA subjects and 38 NP subjects were used to define new principal components using the same methodology as Section 3.4. These updated principal components shall now be described:

Knee Flexion

The updated definition of the flexion/extension PC is shown in Figure 3.27. The first principal component still represents the difference changes in knee flexion during stance phase, where OA subjects on average appear to have less peak flexion, extension and hence ROM. Similarly, the second principal component represents an almost identical portion early to mid-swing phase. It appears OA subjects tend to adopt both a reduced and a delayed peak flexion angle. An additional PC has also been defined towards mid to terminal swing, where OA subjects appear to display a slower rate of knee extension. The explained variance of the three components was 58.4%, 13.9% and 13.4% respectively, resulting in a total representation of 85.6%.

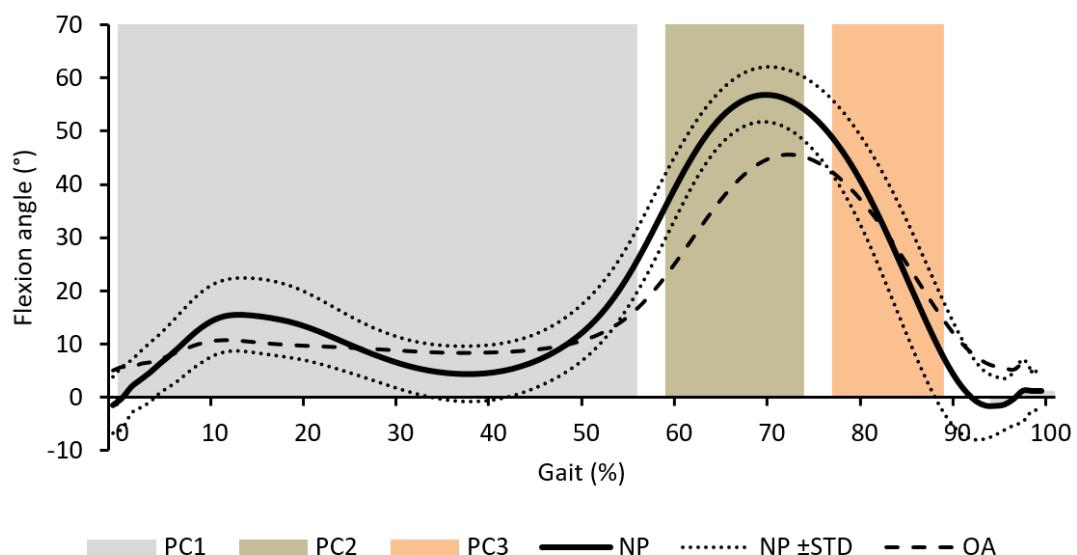


Figure 3.27 Knee flexion angle for NP (solid black, dotted standard deviation), and OA subjects (dashed), with regions of retained PC interpretation shaded.

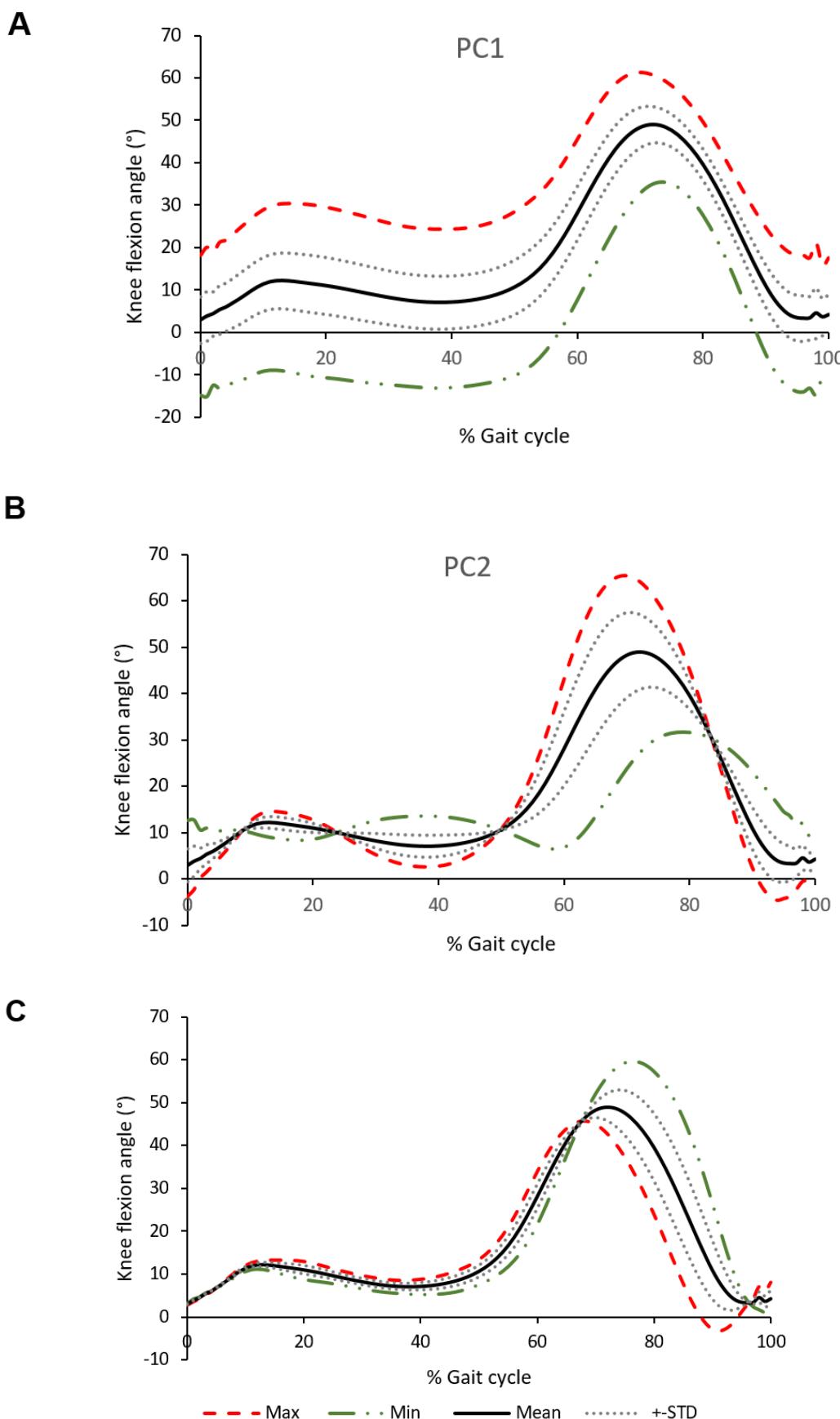


Figure 3.28 Reconstruction of the original flexion angle data using only the retained principal components, PC 1-3, individually. For each PC, the subject with the highest and lowest PC scores are reconstructed, alongside the mean and STD PC scores.

The reconstruction of the original data using the PC scores and the PC vectors significantly aids the interpretation of principal components. It can be seen for Figure 3.28A, that the first PC represents not only the increase/decrease knee flexion during the stance phase but also a great deal during swing phase. Before reconstructing this PC, we might have assumed that it represented the decreased ROM during stance phase of OA subjects; however, this is not the case. Surprisingly, it can be seen from Figure 3.28B that the difference of stance ROM is surprisingly better represented within the reconstruction of PC2. Again, the reconstruction of PC2 expands our understanding of what it functionally represents: it appears that subjects with less ROM during stance phase have not only reduced peak flexion angles, but also a delay of those peaks. This is very much in agreement with previous findings of the kinematic changes due to OA (Astaphen *et al.*, 2008a).

Finally, the third PC is a harder to interpret. It appears that there is a relationship between a delay in the peak flexion angles during the swing, and the magnitude of this peak. This is a very different relationship to that represented by PC2. A notable difference is that this relationship does not correlate to any changes during stance phase, and could, therefore, be a potential factor of variation within NP subjects, as opposed to a representative of changes due to OA.

Internal/External Rotation

The updated PC is used in representing variance within the knee internal/external rotation angles, as shown in Figure 3.29. On average, it does not appear that there is a detectable difference between OA and NP subjects. Variance within NP subjects also appears greatest during swing phase. The first component represents a similar region of variance to that defined by Jones; from load response to the end of initial swing. This is a period where we might expect progressive internal rotation, followed by the beginning of external rotation from toe-off (Lafortune *et al.*, 1992). A second PC was also defined during terminal swing, at which point we'd expect the knee to begin to internally rotate again into a slightly externally rotation position at heel strike (Lafortune *et al.*,

1992). The explained variance of the two components was 53.1% and 16.9% respectively, resulting in a total representation of 70% of the waveform.

The reconstruction of each PC is shown in Figure 3.30. The first PC has a similar interpretation to that of the first PC of the flexion/extension waveform: it represents an increase or decrease of knee external rotation throughout the entirety of stance phase. However, there is also an element of increased ROM during swing phase represented by the component. It seems that subjects whose knees are further from neutral transverse alignment during stance phase are more likely to have increased ROM during swing phase. Reconstructing using the second PC indicates that there appears to be a relationship between increased external rotation from weight acceptance to toe-off, and increased internal rotation during swing phase and heel strike.

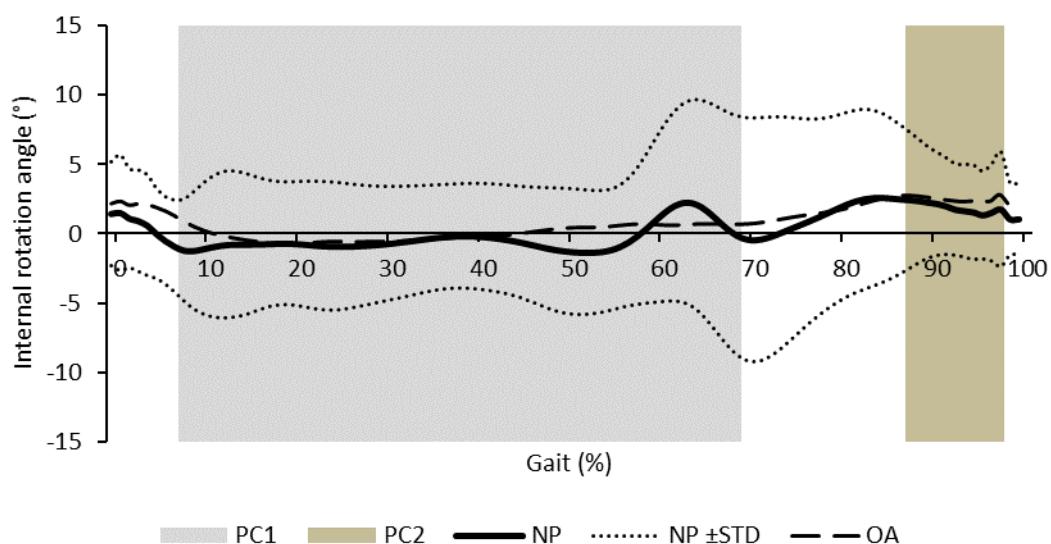


Figure 3.29 Knee internal/external rotation angle for NP (solid black, dotted standard deviation), and OA subjects (dashed), with regions of retained PC interpretation shaded.

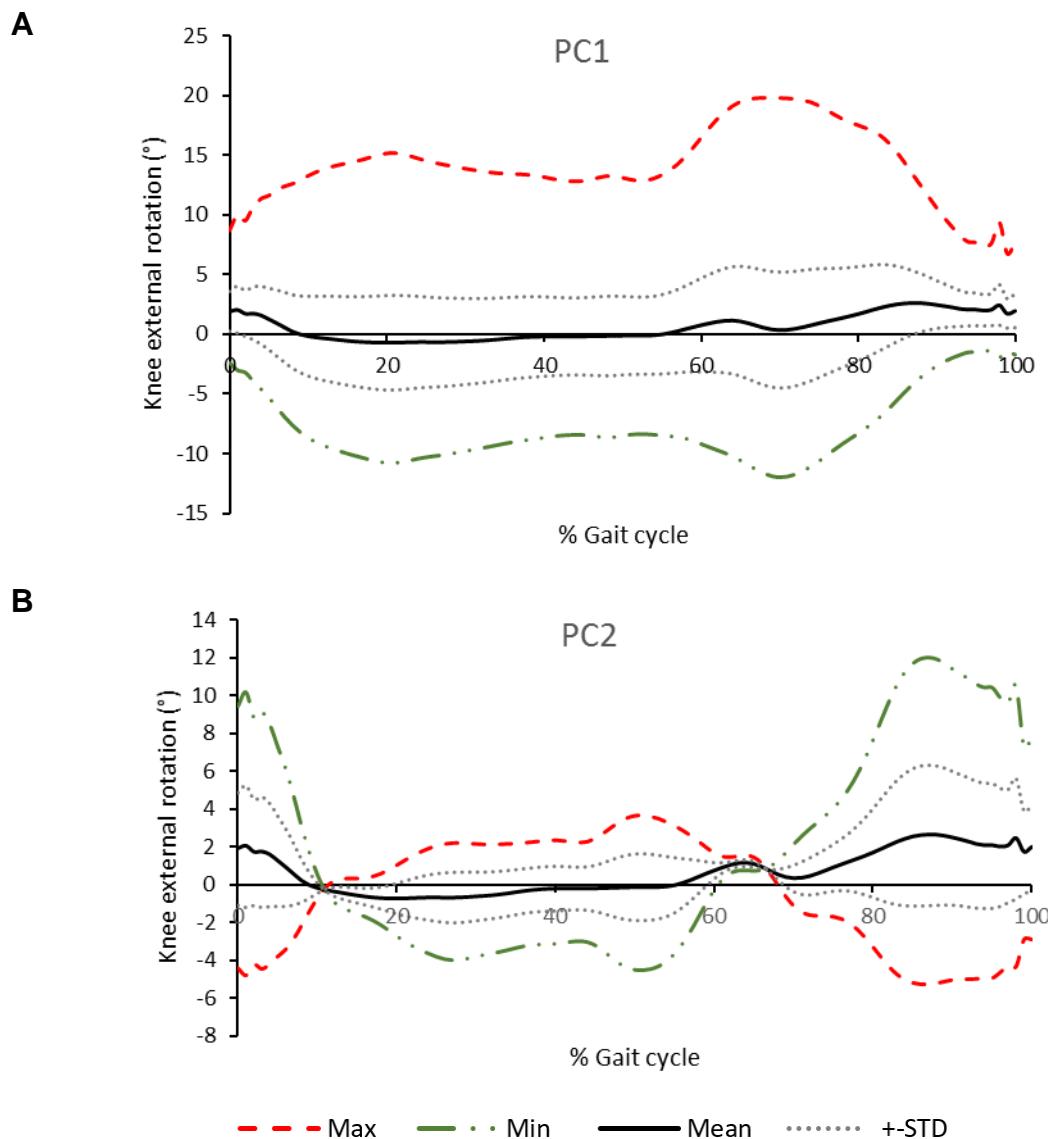


Figure 3.30 Reconstruction of the original ext/internal knee rotation angle data using only the retained principal components, PC 1-2, individually. For each PC, the subject with the highest and lowest PC scores are reconstructed, alongside the mean and STD PC scores.

Ab/Adduction

The updated PC is used in representing variance within the ab/adduction angles, as shown in Figure 3.31. On average, there does not appear to be any noteworthy differences between OA and NP subjects. The region of variance represented by the first PC remains the same as that described by Jones. We wouldn't expect a high amount of coronal plane ROM during stance phase in NP subjects (Lafortune *et al.*, 1992); however, this value could easily be affected by errors in the identification of anatomical landmark and the alignment of the lower limb. The second PC represented a similar

region – initial swing. This is generally the period in which the majority of abb/adduction occurs in NP subjects. Unlike Jones, a third PC was not identified. The two PCs explained a variance of 59.3% and 17.9% respectively.

Data reconstruction using the two PCs is shown in Figure 3.32. The reconstruction using the first PC reveals again a similar interpretation to that of the first PC of the external rotation angle: an increase/decrease in the angle throughout the duration of the gait cycle and an angle further from neutral during stance phase relating to an increased ROM during swing phase. The second PC definition also displays a very similar relationship as PC of the external rotation angle; a relationship between increased adduction from weight acceptance to toe-off, and increased abduction during swing phase and heel strike (and vice versa).

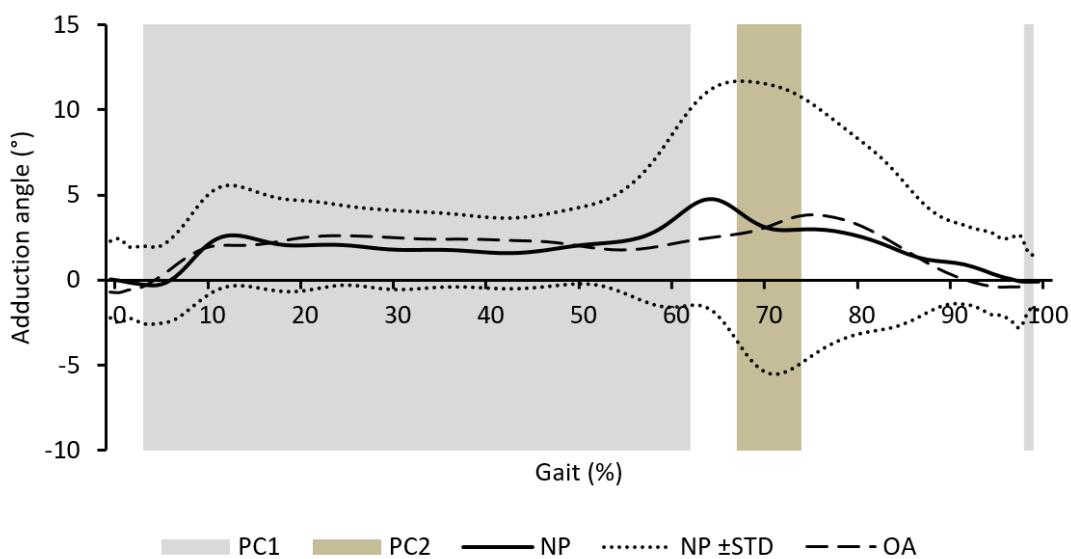


Figure 3.31 Knee ab/adduction angle for NP (solid black, dotted standard deviation), and OA subjects (dashed), with regions of retained PC interpretation shaded.

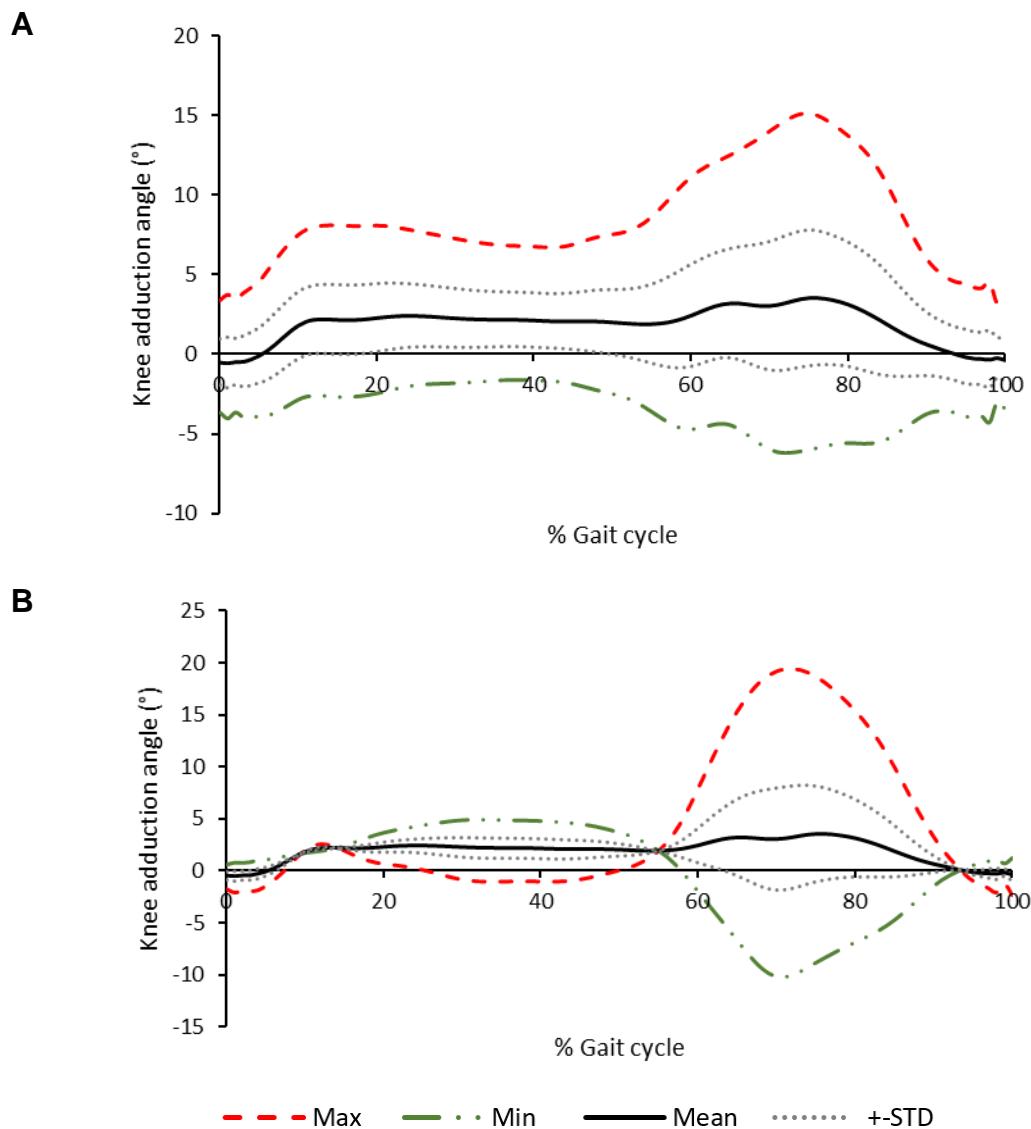


Figure 3.32 Reconstruction of the original abb/adduction angle waveforms using only the retained principal components, PC 1-2, individually. For each PC, the subject with the highest and lowest PC scores are reconstructed, alongside the mean and STD PC scores.

Vertical Ground Reaction Force

The region of variance represented by the first PC is different to that defined by Jones *et al.* (2008); both represent changes during midstance, however, the new definition also represents the region during second double limb support, as well as loading response between the first heel strike transient, and the first peak of the vertical GRF. The second PC appears to represent the region between heel off and opposite leg heel strike, towards the second peak of the vertical GRF. The third PC appears to represent the area towards toe-off, as the limb prepares for swing phase. The three PCs explained a

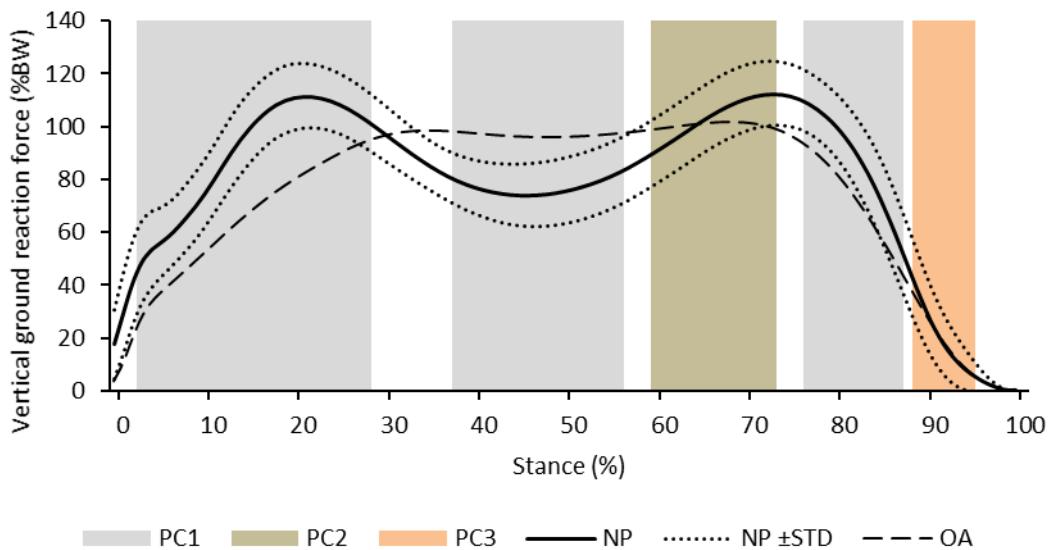


Figure 3.33 Vertical ground reaction force for NP (solid black, dotted standard deviation), and OA subjects (dashed), with regions of retained PC interpretation shaded.

variance of 50.2%, 23.3% and 11.8% respectively, resulting in a total representation of 85.3% of the waveform.

The reconstruction of the vertical GRFs using the three identified PCs is shown in Figure 3.34. The reconstruction of data using the first PC reveals that this component represents a reduction in the ‘double peak’ of the GRF. In NP subjects, the vertical GRF typically peaks above 100% of body weight during loading response, as the falling COM of the subject is decelerated by the limb, and again during the swing of the opposite leg as the COM is accelerated vertically. It appears that the OA subjects maintained a steadier vertical position of COM. This was also correlated with a much slower rate of weight acceptance, characterised by a longer period until the first GRF peak. These features all appear to be accounted for within the first PC.

The second PC appears to reconstruct a small amount of variance throughout the duration stance phase. If we consider that the first PC has already reconstructed how much of a “double” peak pattern exists, this second PC reconstructs variation in absolute magnitudes within these patterns. As these forces are normalised to body weight, the

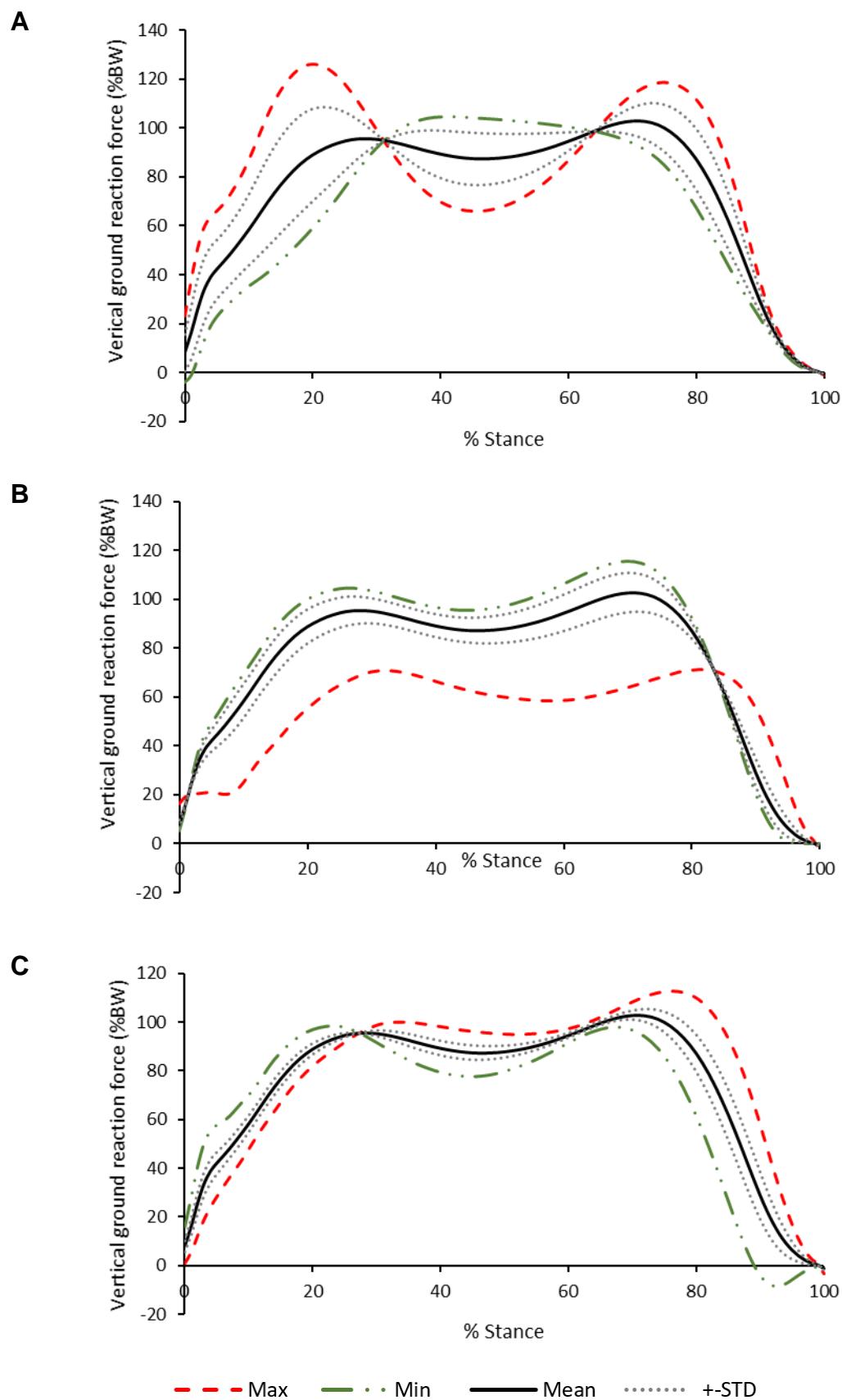


Figure 3.34 Reconstruction of the original vertical ground reaction force waveforms using only the retained principal components, PC 1-2, individually. For each PC, the subject with the highest and lowest PC scores are reconstructed, alongside the mean and STD PC scores.

accuracy of the measurement of participant weight would likely have a large influence on the PC scores of this component. The minimum PC score for this component appears to be very far from the standard deviation of all the PC scores. The original waveform of this subject was therefore checked, along with the bodyweight; however, the original waveform bears little resemblance to that which might be indicated by the second PC. This is a challenge when interpreting PCs through individual reconstructions, and is discussed further in Section 6.1.2. Finally, the third PC appears to reconstruct differences in timing of the initiation of push-off.

Anterior-Posterior Force

The region of variance represented by all three PCs of the anterior-posterior force match the regions described by Jones *et al.* (2008), and are plotted in Figure 3.35. The first PC represents the region of the anterior and posterior peaks which occur as the COM of the subject is decelerated and accelerated in the sagittal plane. The second PC appears to represent the region between midstance and midterminal stance. The third PC represents the region towards terminal pre-swing. The three PCs explained a variance of 50.0%, 28.3% and 8.8% respectively, resulting in a total representation of 87.1% of the waveform.

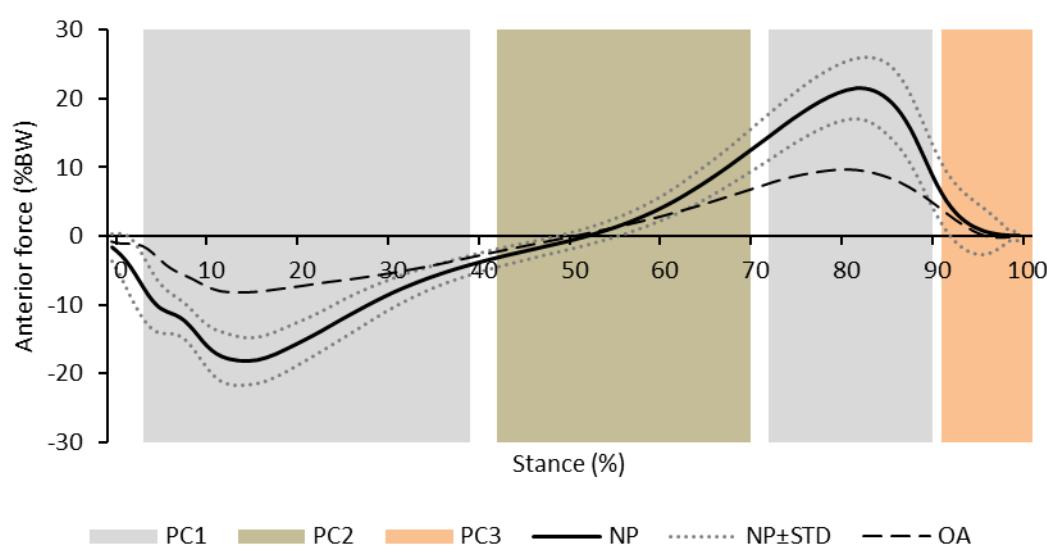


Figure 3.35 Anterior-posterior ground reaction force for NP (solid black, dotted standard deviation), and OA subjects (dashed), with regions of retained PC interpretation shaded.

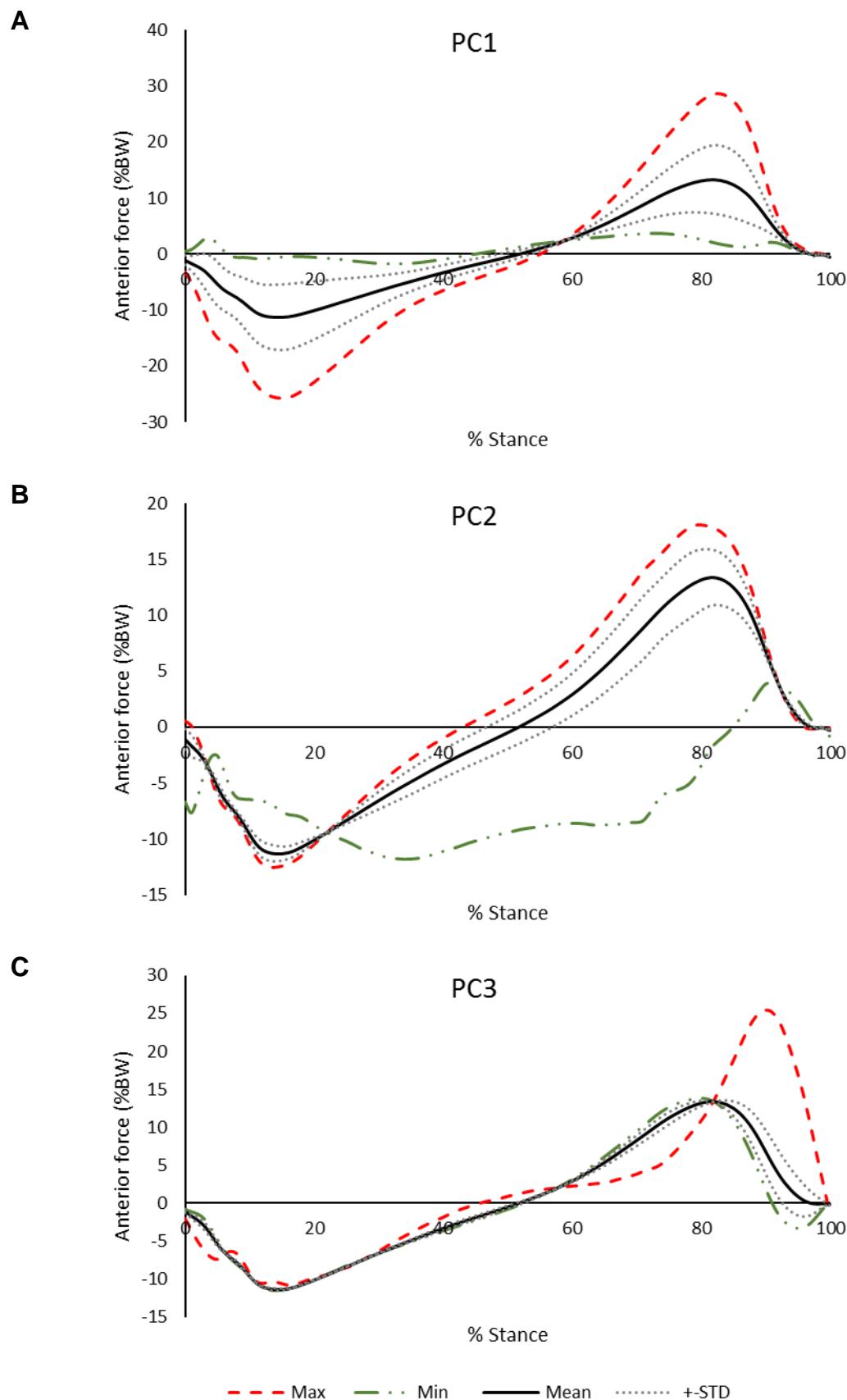


Figure 3.36 Reconstruction of the original anterior ground reaction force waveforms using only the retained principal components, PC 1-3, individually. For each PC, the subject with the highest and lowest PC scores are reconstructed, alongside the mean and STD PC scores.

The individual reconstruction of each PC is shown in Figure 3.36. The first PC reconstructs the relationship previously mentioned; indicating there is a strong correlation between the increase of the first peak of posterior force and the second peak of the anterior force. The second PC seems to reconstruct a relationship between the earlier timing of the initiation of an anterior force, and a greater magnitude of the anterior peak. Again, the minimum value for this PC falls very far from the standard deviation, and appears to show an avoidance of an anterior force, and hence a very prolonged posterior force. The third PC appears to reconstruct a relationship between a delayed peak of the anterior GRF, and the magnitude of that peak. As this component only represents a total of 8.8% of the variance and has a relatively small standard deviation, this relationship might only be present in a small number of the participants.

3.6.3 Classification Using Updated Principal Component Definitions

Following the redefinition of PCs using the increased cohort, the data was collated as a training body of NP and OA subjects. The previously included clinical and temporal-spatial parameters: BMI, stance percent, cadence, were again included in the classification, in order to make a direct comparison between the two PC definitions.

By comparing the results of the newly defined PCs, shown in Table 3.6, with those of the original components, shown in Table 3.4, it can be seen that there was a small improvement in the largest in-sample classification accuracy achieve (96.7% vs 94.9%). There was no increase in the maximal out-sample classification accuracy (94.3%). The control parameters which achieved the greatest in and out-sample classification accuracy was using $k_{c/s}$ as the weighting/ steepness of the activation function, and θ_s for the bias/midpoint.

Table 3.6 Classification results of 85 OA and 38NP subjects using newly calculated principal components and scores to those defined by Jones (2004), comparing definitions of control variables k and θ

K definition	Theta definition	In-sample accuracy (%)	LOO accuracy (%)
K_c	θ	86.9	86.9
	θ_A	93.4	93.4
	θ_S	93.8	93.4
K_s	θ	80.9	75.4
	θ_A	89.4	85.2
	θ_S	91.2	87.7
$K_{c/s}$	θ	86.9	86.9
	θ_A	93.4	92.6
	θ_S	96.7	94.3

3.6.4 Leave-P-Out Classification

To further assess the robustness of the classification, and evaluate the reliability of the LOO cross-validation, a k-fold classification was applied with an increasing number of hold-out subjects. The highest accuracy midpoint definition; θ_S , was used for each case. The three different definitions of k are compared. This is a non-exhaustive random sampling, where the average accuracy of 500 repeats was taken. The results are shown in Figure 3.37. For each value of p , an equal number of NP and osteoarthritic subjects is chosen for the validation set.

It can be seen from Figure 3.37 that the classification accuracy using each technique doesn't vary a great amount as the value of p increases. The classification accuracy when defining the steepness of the sigmoid transfer function as k_c or $k_{c/s}$ is consistently greater than when defined as k_s .

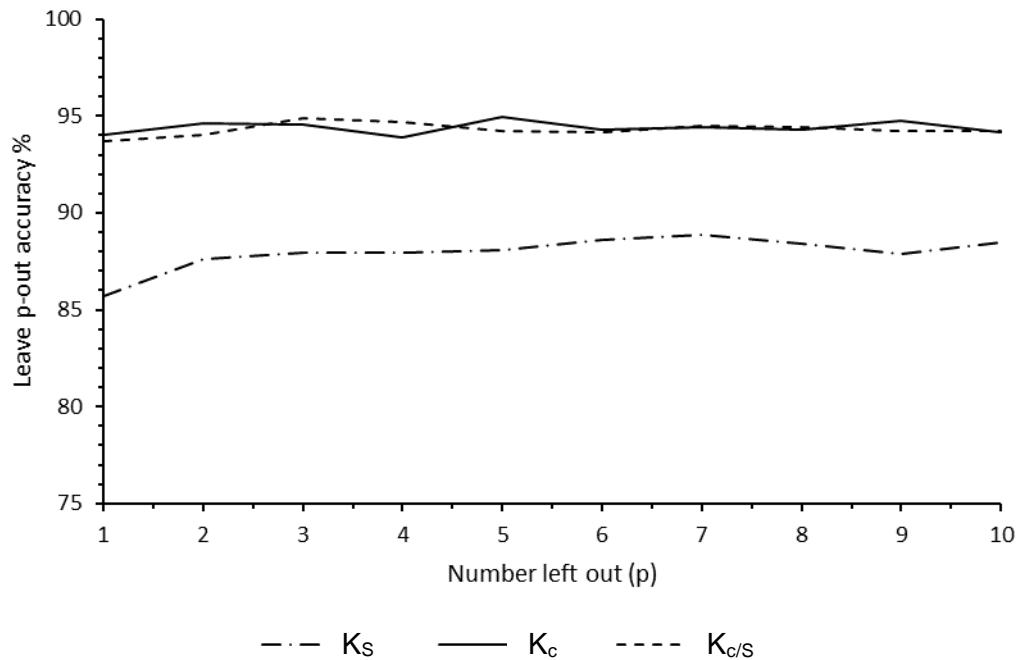


Figure 3.37 The non-exhaustive leave-p-out classification results for the training body of the 86 OA and 38NP subjects included in this study. The midpoint definition θ_S was used, and the three different definitions for k are compared. The points represent the average leave-p-out accuracy for 500 random resamples.

3.6.5 Increasing the Classification Cohort Size

To understand further understand how many subjects are required in order to accurate classify OA and NP subjects, the DST classifier was tested with increasing cohort sizes. For each classification cohort size, the subjects were randomly pooled from the dataset and used to classify the remaining data. This was repeated 500 times for each cohort size. The results are shown in Figure 3.38. Classification accuracy is surprisingly high even with small cohorts and reaches above 90% with only five subjects in each group. The classification accuracy seems to show signs of plateauing towards to maximum available of 38 subjects in each group. It is worth noting that due to imbalance in the total data pool (86 OA and 38 NP), as the training set size increases, the ratio of OA to NP subjects in the test set also increases, until eventually all 38NP subjects are within the training set and the test set comprises of the remaining 48 OA subjects.

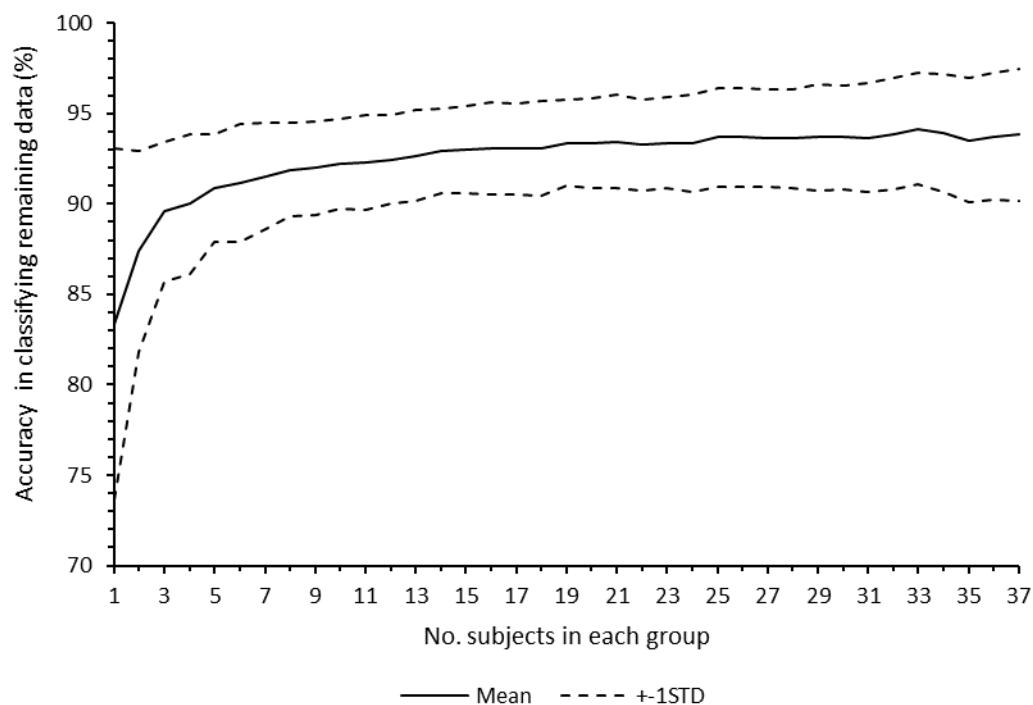


Figure 3.38 The effect of a change in the number of subjects in each group within the training body, and the resultant classification accuracy when used to classify the remaining test data. For each subject group size, subjects were randomly pooled from selected from the total 85OA and 38NP subjects. This was repeated 500 times for each training body size

3.6.6 Adding ML Forces and Moments to the Classification

Mediolateral Force

The principal components selected for the ML force and their corresponding regions of variance are represented in Figure 3.39. Much of the medial/lateral force is included in the first PC; including the lateral peak around the time of the heel strike transient, and from terminal loading response towards terminal stance. The second PC didn't represent more than 50% of the variance at any point of the gait cycle, however upon further inspection contained useful information. This is a drawback to the selection technique adopted by Jones (2004), which is discussed further in Section 4.2.8. The explained variance of the three components was 62.3%, 13.5%, 6.8%, respectively, resulting in 82.6% of the total variance being represented.

The reconstruction of the original data using the three PCs individually is shown in Figure 3.40. Much like many other PCs that have been defined, the first PC reconstructs a change in magnitude throughout the entirety of stance phase. The mediolateral GRF is related to the movement of the centre of mass during walking: the COM is decelerated within the coronal plane from heel strike to midstance and then accelerated

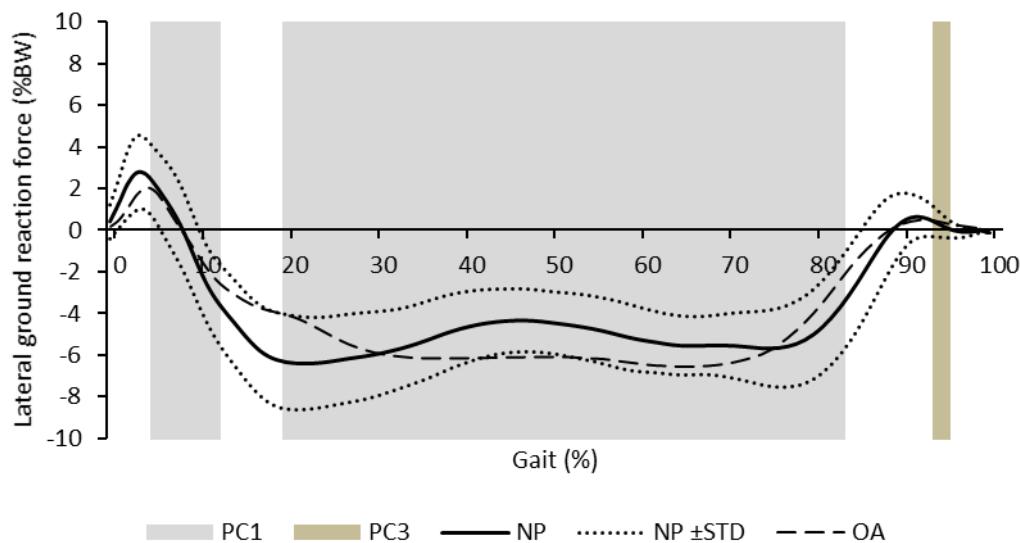


Figure 3.39 Mediolateral ground reaction force of NP (solid black, dotted standard deviation), and OA subjects (dashed), with regions of retained PC interpretation shaded.

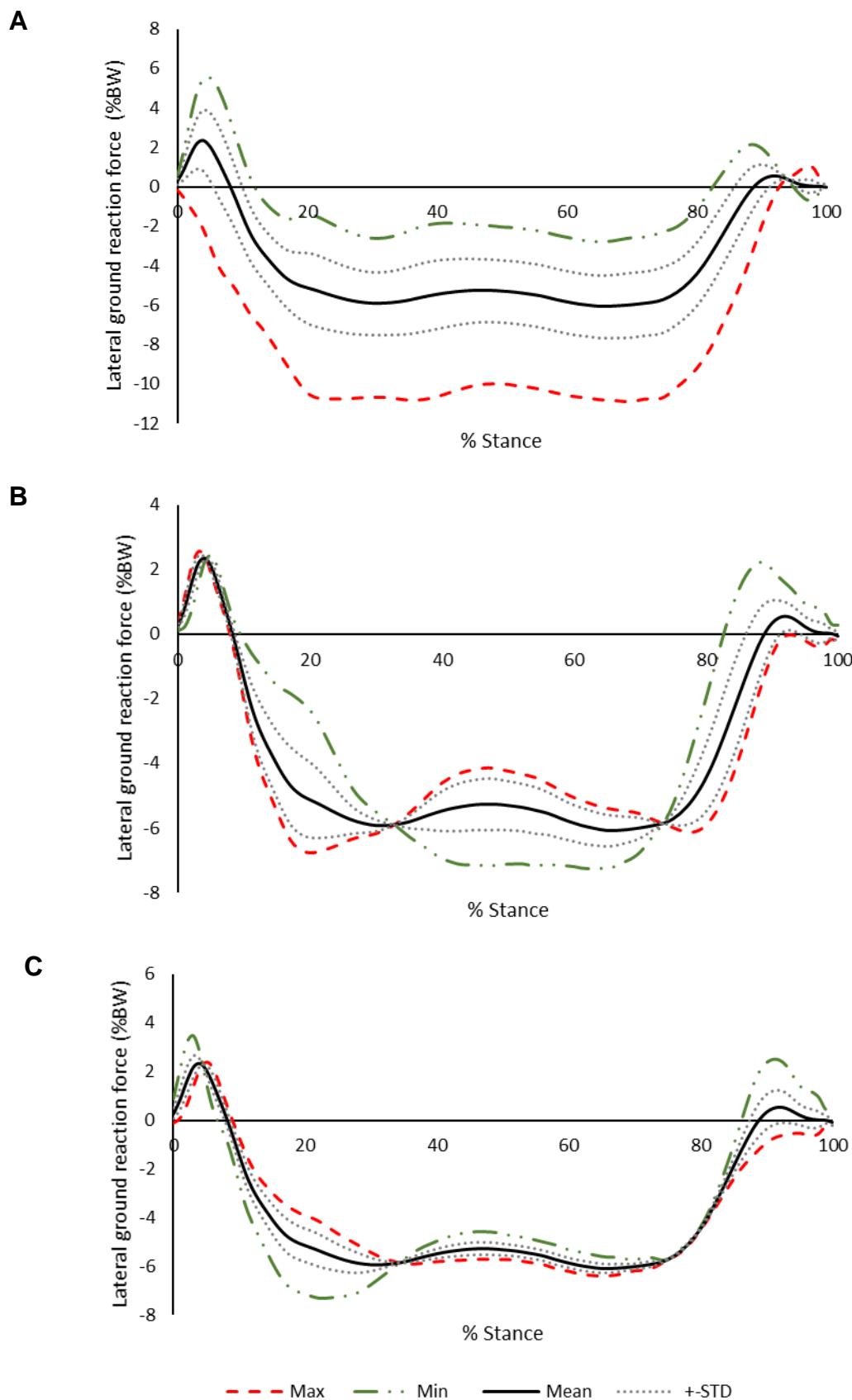


Figure 3.40 Reconstruction of the original mediolateral ground reaction force data using only the retained principal components, PC 1-3, individually. For each PC, the subject with the highest and lowest PC scores are reconstructed, alongside the mean and STD PC scores.

towards the other limb in the second half of stance. Both require a medial GRF, and deceleration and propulsion phases contribute to the presence of two peaks. The movement of the COM in the coronal plane is related to gait velocity and coronal plane stability; if the COM moves lateral to the supporting limb this can increase the risk of falls (Hof *et al.*, 2005).

Sagittal Moment

The principal components selected for the knee flexion moment and their corresponding regions of variance are represented in Figure 3.41. The first principal component represents most the waveform and includes the loading response, midstance, and terminal stance. Within these regions, it appears that, on average, OA subjects have a smaller magnitude of both flexion and extension moments. The second PC represents a small area before the first peak of the extension moment; where the COM has generally progressed over the supporting limb and the COP progresses towards the front of the foot. The explained variance of the two components was 54.5% and 16.8% respectively, resulting in a total representation of resulting in a total variance of 71%.

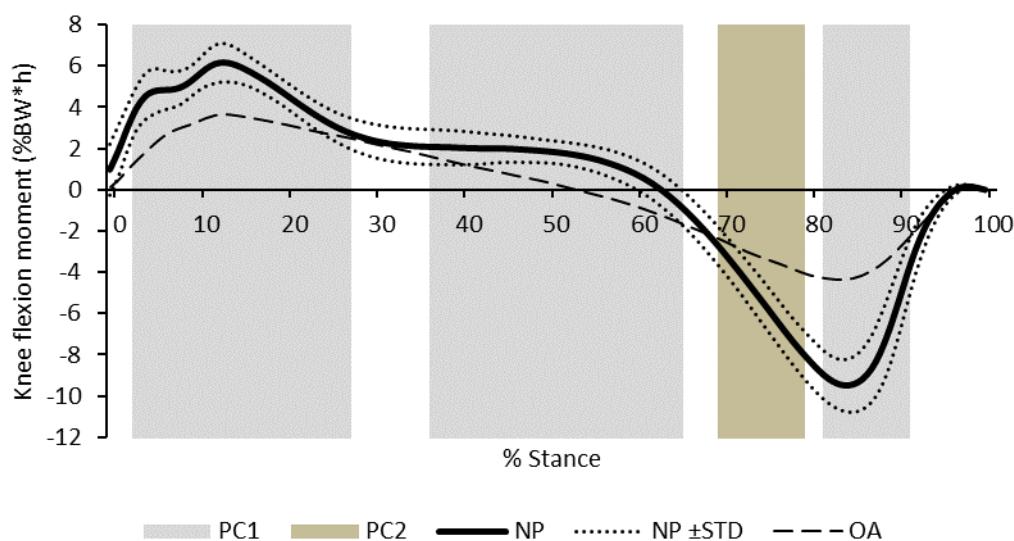


Figure 3.41 Knee flexion/extension moment of NP (solid black, dotted standard deviation), and OA subjects (dashed), with regions of retained PC interpretation shaded. Moments have been normalised and expressed as a percentage of bodyweight*height.

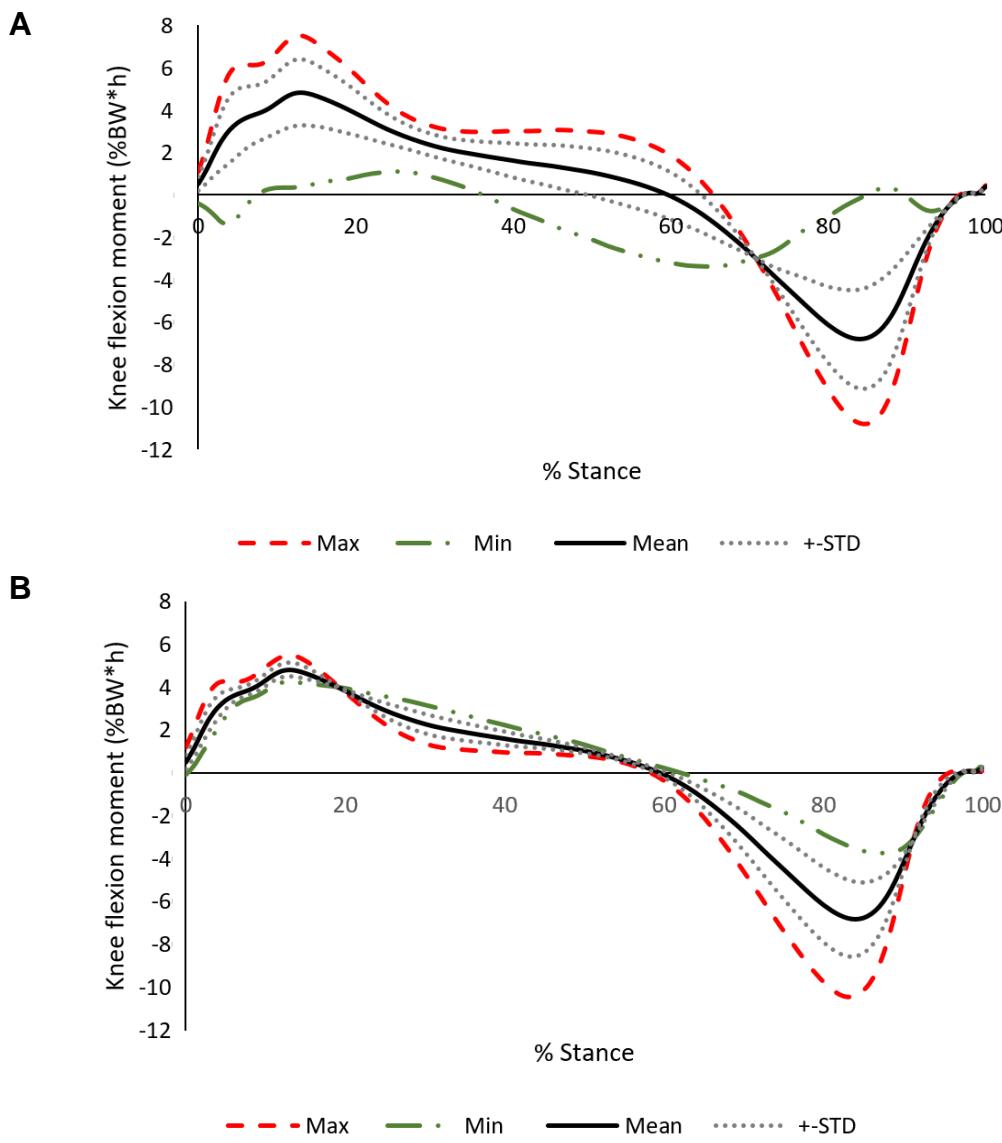


Figure 3.42 Reconstruction of the original knee flexion/extension moment waveform using only the retained principal components, PC 1-3, individually. For each PC, the subject with the highest and lowest PC scores are reconstructed, alongside the mean and STD PC scores. Moments have been normalised and expressed as a percentage of bodyweight*height.

The reconstruction of the individual PCs is shown in Figure 3.42. The first PC reconstructs an intuitive relationship between an increased flexion moment during loading response and midstance, and an increased extension moment during terminal stance. There also appears to be a relationship between a decreased peak of the flexion moment, and an early transition towards an extension moment. The second PC reconstructs mainly the extension moment peak towards terminal stance without affecting the flexion moment peaks, indicating that some subjects may avoid the extension peak without affecting the rest of the sagittal knee moment.

Coronal moment

The principal components selected for the knee flexion moment, and their corresponding regions of variance, are represented in Figure 3.43A. The first PC represents most of the variance throughout the duration of stance phase, during which it appears that, on average, OA subjects have an increased adduction moment. The second PC appears to represent only a small amount of variance towards the very end of stance phase. The explained variance of the two components was 80.3% and 6.8% respectively, resulting in a total representation of variance of 81%.

The reconstruction of the waveform using the two selected PCs is shown in Figure 3.44, B & C. The first PC reconstructs the change in magnitude of the adduction moment throughout the entirety of the stance phase. While subtle, it also appears that an increased magnitude of the adduction moment may be related to a small abduction moment towards toe-off. The reconstruction using the second PC accounts for very small changes in the amount of a dip between the two adduction moment peaks, as well as what appears to be differences in timing of these peaks.

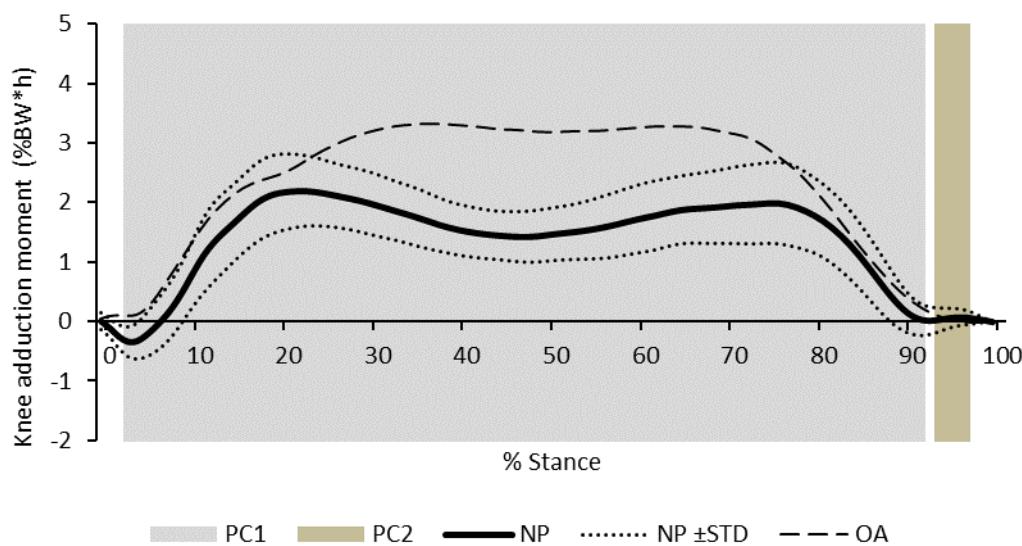


Figure 3.43 Knee adduction moment of NP (solid black, dotted standard deviation), and OA subjects (dashed), with regions of retained PC interpretation shaded. Moments have been normalised and expressed as a percentage of bodyweight*height.

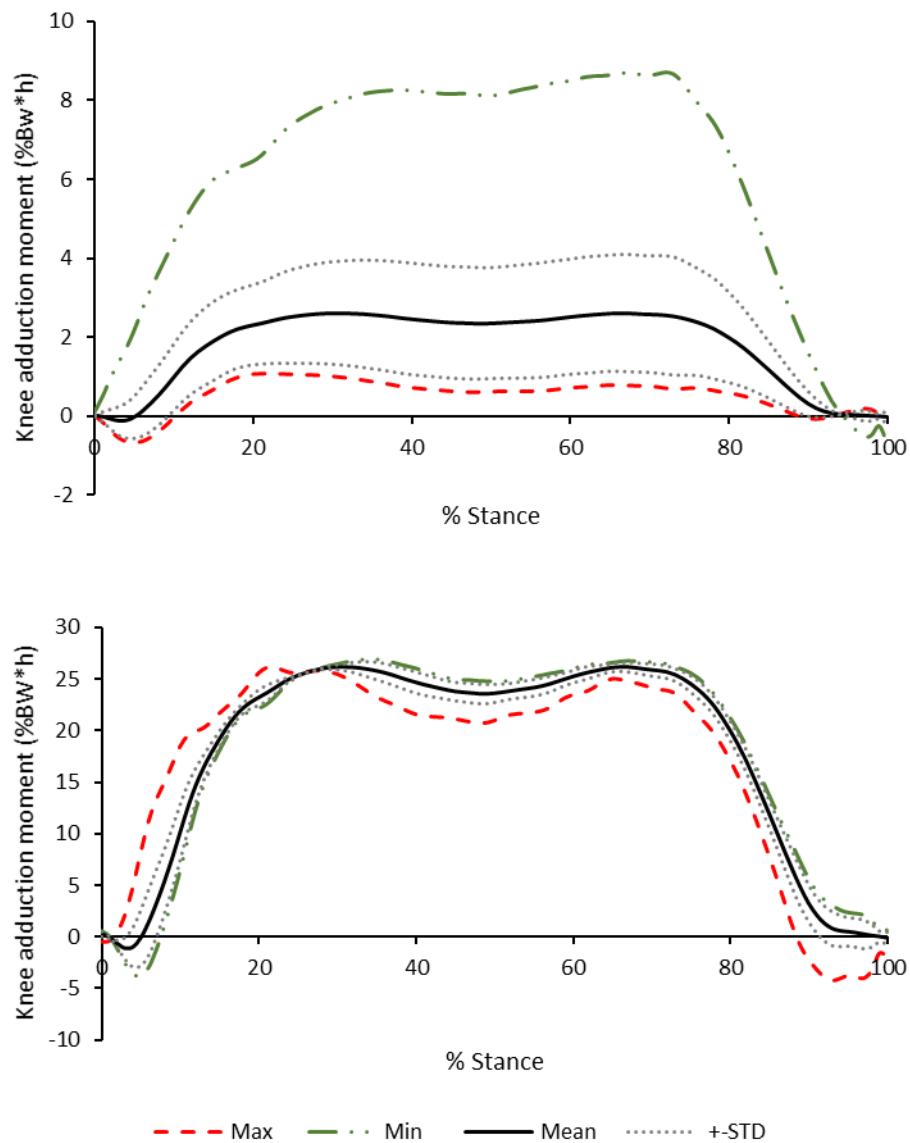


Figure 3.45 Reconstruction of the original knee adduction moment waveform using only the retained principal components, PC 1-3, individually. For each PC, the subject with the highest and lowest PC scores are reconstructed, alongside the mean and STD PC scores. Moments have been normalised and expressed as a percentage of bodyweight*height.

Transverse Moment

The principal components selected for internal knee moment and their corresponding regions of variance are represented in Figure 3.46. The first principal component represents the region of loading response and the second half of stance; where typically a peak external and internal moment would occur respectively (Brandon and Deluzio, 2011). The second PC appears to represent a region of variance towards the end of

loading response and the beginning of midstance. The explained variance of the two components was 60.9% and 25.8% respectively, resulting in a total variance of 87%.

The reconstruction of the waveform using the two selected PCs is shown in Figure 3.47. The reconstruction of the first PC shows an unexpected relationship between the first and second peaks of the transverse knee moment: if the first peak is an internal moment, the second peak is more likely to be an external, and vice versa. Other studies on transverse knee moments find NP subjects generally have a biphasic transverse moment with an external first peak and an internal second peak, which reduce in magnitude with severe OA (Brandon and Deluzio, 2011). One potential cause which could have resulted in this unexpected finding was that the positive definition of the knee moment might have been defined differently for each leg. The MATLAB code was double-checked, and this wasn't the case. Also, the transverse knee moments for the NP subjects, where the right knee was used, was compared to those of the left knee and no pattern of inverted sign convention was visible. The results for the transverse knee moments will therefore still be considered but will be treated with caution.

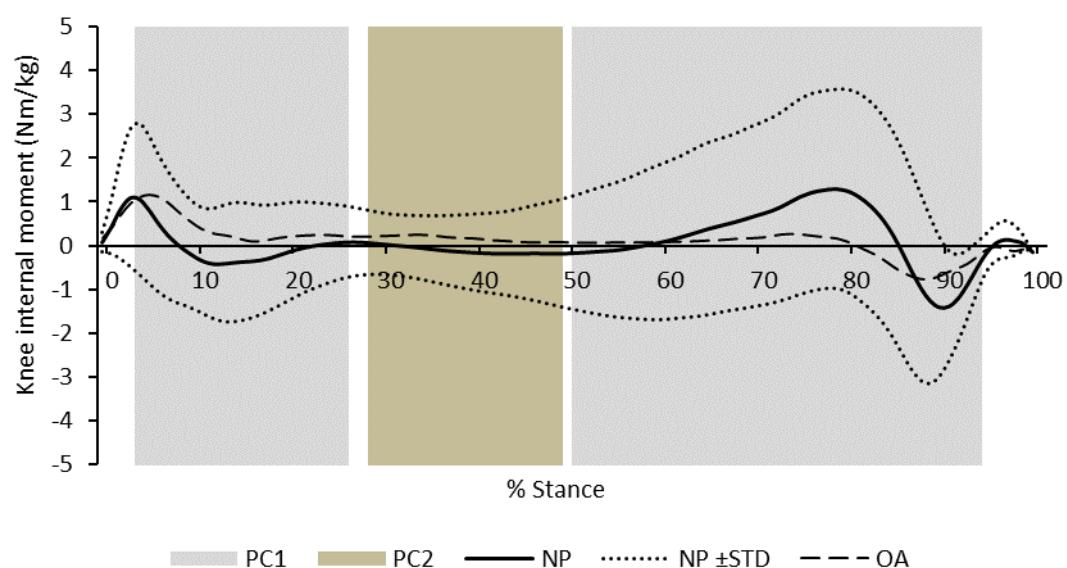


Figure 3.46 Knee internal moment of NP (solid black, dotted standard deviation), and OA subjects (dashed), with regions of retained PC interpretation shaded. Moments have been normalised and expressed as a percentage of bodyweight*height.

The second PC also appears to show a relationship between an internal moment throughout the beginning of stance, and an external moment during terminal stance. The second PC, however, reconstructs a prolonged internal or external moment throughout midstance, and hence a later transition towards the second peak.

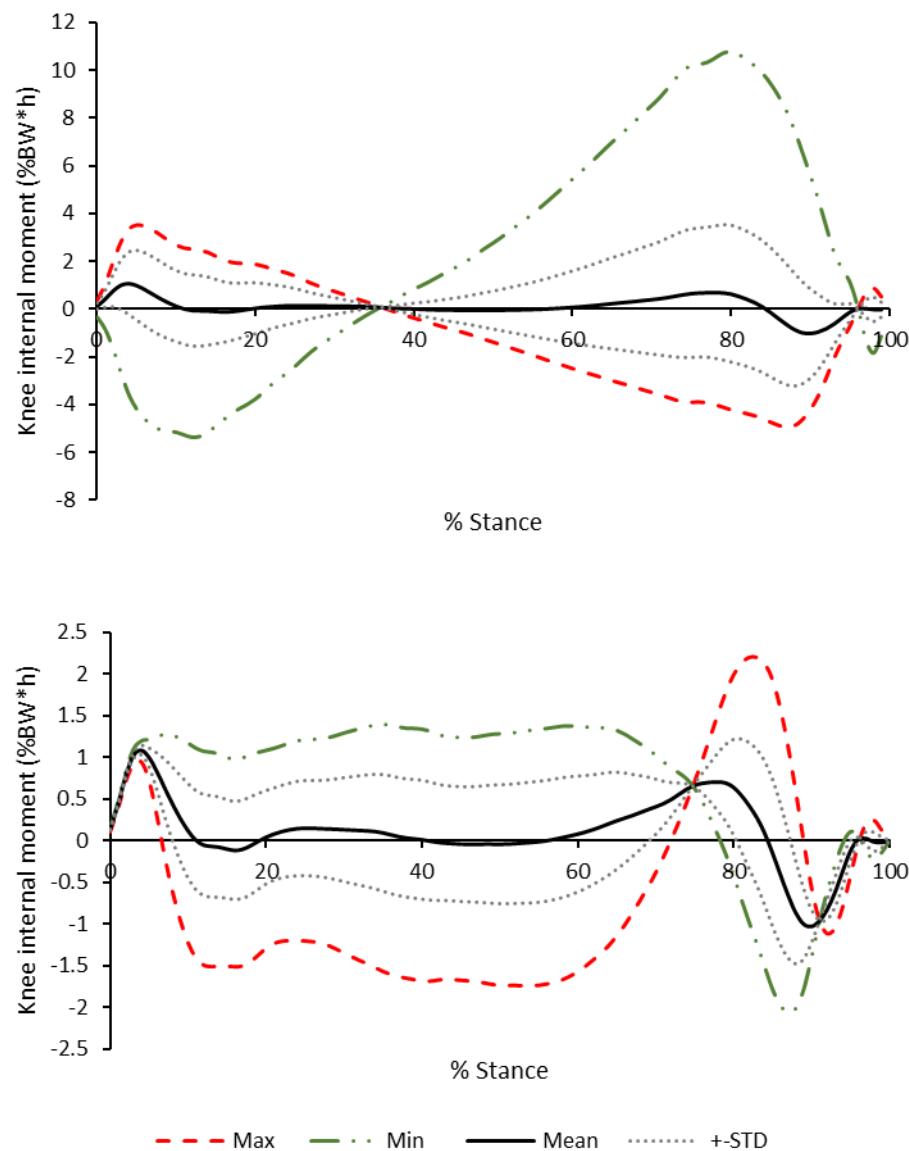


Figure 3.47 Knee internal moment of NP (solid black, dotted standard deviation), and OA subjects (dashed), with regions of retained PC interpretation shaded. Moments have been normalised and expressed as a percentage of bodyweight*height.

Classification Results

The classifier was trained on the retained PCs of the 18 NP and 20 OA subjects for whom joint moments and mediolateral forces could be calculated. The results of the LOO classification accuracy are shown in Table 3.7. Also shown, for a fair comparison, are the classification results when trained with the same subjects but without the additional joint moments and mediolateral force data added.

Similarly to the previous two datasets, the $K_{c/s}$ and K_c definition consistently achieved more favourable results than that of K_s . The definition using $K_{c/s}$ achieved slightly greater LOO classification accuracy than K_c when using the expanded variables. Overall, the $K_{c/s}$ definition also performed most accurately when using the original variables, despite LOO accuracy being lower for θ and θ_A definitions.

The original definition of the midpoint of the control function used by (Jones, 2004), θ_O , didn't perform as poorly as in the previous datasets. This is because the two groups were of similar size (18 vs 20) and the global average was therefore only slightly biased towards the OA group. As in the previous two datasets, the θ_S proved the best-performing theta definition, followed by θ_A .

Table 3.7 Classification results of the 20 OA and 18 NP subjects for which it was possible to calculate mediolateral GRF and joint moments, using the PCs defined within this study, comparing definitions of control variables k and θ .

		LOO classification accuracy (%)	
K definition	Theta definition	Moments added	Without moments
K_c	θ_O	97.4	94.7
	θ_A	100	94.7
	θ_S	100	97.4
K_s	θ_O	94.7	92.1
	θ_A	94.7	89.5
	θ_S	94.7	89.5
$K_{c/s}$	θ_O	100	92.1
	θ_A	100	92.1
	θ_S	100	100

Table 3.8 The ranking of classification input variables in for the 20 OA and 18 NP subjects for which it was possible to calculate mediolateral GRF and joint moments, using the PCs defined within this study. The additional PCs defined within this section are highlighted in bold.

Rank	Classification accuracy (%)	Variable
1	97.4	GRF Ant/posterior PC1
2	89.5	GRF Vertical PC1
3	81.6	Average Knee Girth
4	81.6	Knee Flexion Moment PC1
5	81.6	Knee Flexion Angle PC2
6	81.6	GRF Mediolateral PC2
7	76.3	AP Knee Depth
8	73.7	Knee Int/external Moment PC1
9	71.1	GRF Mediolateral PC3
10	68.4	ML Knee Width
11	68.4	Knee Ad/Abduction Moment PC1
12	68.4	Knee Flexion Moment PC2
13	63.2	Knee Int/external Mom PC2
14	60.5	GRF Vertical PC2
15	57.9	Knee Ad/Abduction Angle PC1
16	57.9	Knee Int/external Angle PC2
17	55.3	GRF Mediolateral PC1
18	55.3	Knee Ad/Abduction Angle PC2
19	52.6	Knee Flexion Angle PC1
20	52.6	Knee Ad/Abduction Moment PC2
21	52.6	GRF Ant/posterior PC2
22	50.0	Knee Flexion Angle PC3
23	50.0	GRF Ant/posterior PC3
24	47.4	Knee Int/external Moment PC2
25	47.4	GRF Vertical PC3

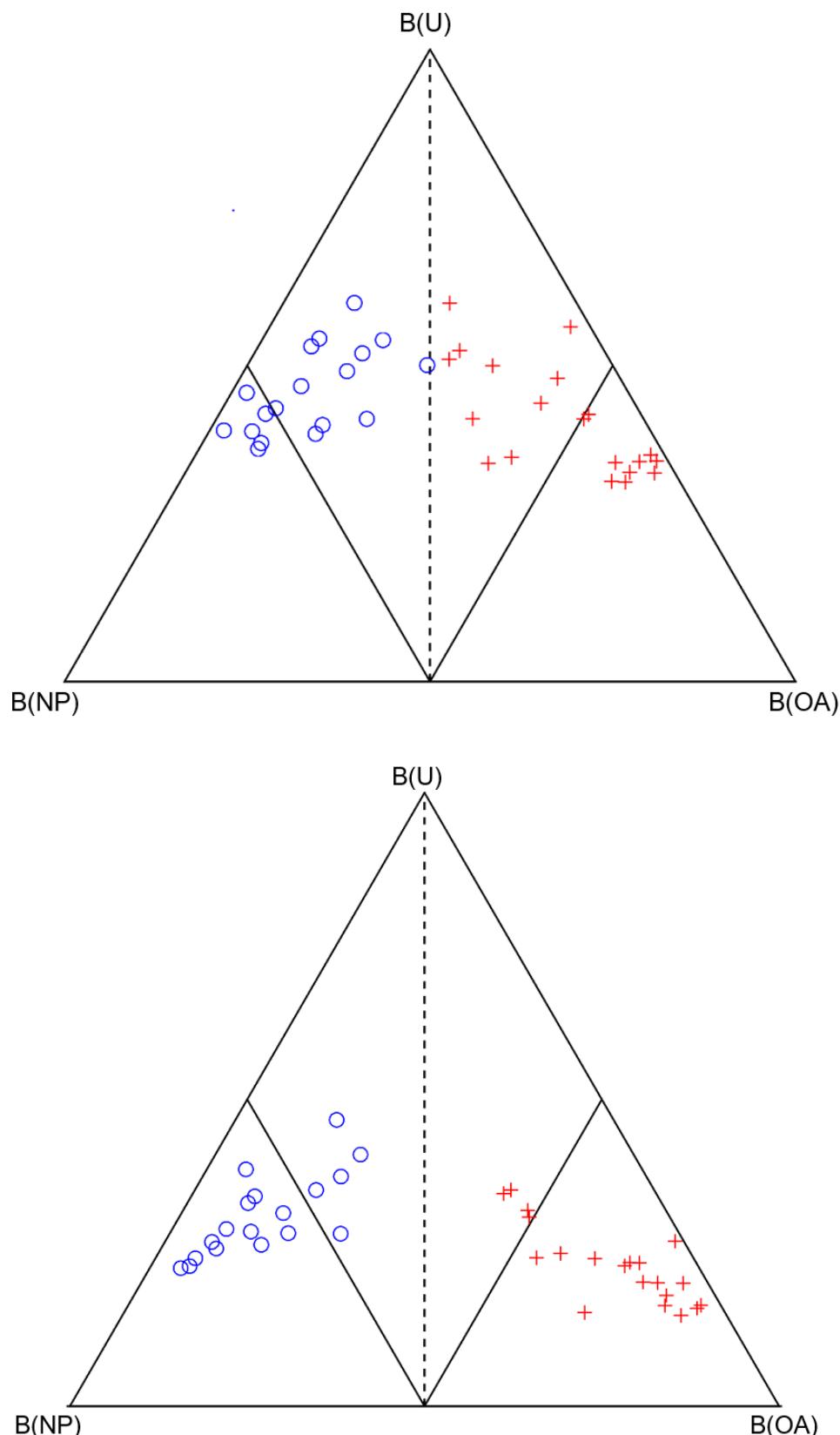


Figure 3.48 Simplex plots to illustrate the LOO classification of 20 OA (red cross), and 18 NP (blue circle), subjects, A) Using the original input variables selected by (Jones, 2004), and the updated PC definitions of Section 3.6.2 B) Using the addition of mediolateral GRF and knee joint moments. In both instances, the $K_{c/s}$ and the θ_S control parameter definitions have been used.

3.7 Conclusions

3.7.1 Exploring the Validity of the Classifier Control Variables

Within this study, previous MATLAB code has been edited and new bespoke code has been written in order to significantly automate and streamline the process of calculating tibiofemoral joint kinematics and kinetics, to reduce temporal waveforms using PCA, and to classify HMA data. This has facilitated the data processing of a huge cohort of subjects; some collected within the duration of this study, but most collected prior to the study commencing. This increased the cohort to 85 OA and 38 NP subjects.

Within this chapter, the control variables used in previous DST classifier studies have been investigated, and particular concerns have been expressed regarding the validity of the choice of two of the control variables; k – which defines the steepness of the sigmoid activation, and θ – which defines the midpoint at which the control function is 0.5. Alternative definitions have been reasoned and proposed for these activation functions.

It has been shown that one of the suggested definitions of k proposed by (Beynon *et al.*, 2000), as the inverse of the standard deviation (k_s), has the same effect as normalising the input variables to a unit variance. This is a common standardisation technique used in other classification techniques, such as in Neural Networks and the Gillette Gait Index, ensuring no bias is given towards datasets of larger magnitudes of scale. It has been shown that the use of only the correlation coefficient to define k (k_c), as proposed by (Beynon, 2002) and found most accurate within Jones (2004) produces a bias which could result in variables with smaller magnitudes of scale having limited contribution to the classification outcome. A new definition has been suggested, incorporates both the standardisation using the standard deviation and an adjusted weighting using the correlation coefficient ($K_{C/S}$). This new definition has consistently improved the performance of the classifier on the three datasets tested. The improvements have been more modest in comparison to the k_c definition; however, most importantly, the new

definition eliminates the bias while still maintaining equal or greater classification accuracy.

3.7.2 Exploring the Sample Size Required to Classify Osteoarthritic Subjects Accurately.

The increased cohort of 85 OA and 38 NP subjects were used to explore and estimate the required sample size to accurately distinguish between NP and OA subjects using a combination of clinical (knee joint measurements, BMI), temporal-spatial (cadence, stance/swing percent), and tibiofemoral kinematics. This was achieved by randomly subsampling subjects from the full dataset as the ‘training’ cohort, and then testing their accuracy at classifying the remaining subjects. This was repeated 500 times for each size of the training body.

The results indicate that, surprisingly, both the classification accuracy and the variability of that accuracy within the 500 repeats appears to begin to plateau quite early; at around ten subjects in each group. This indicates that to build a training body for the classification of pathologies with severe gait abnormalities, such as late-stage knee OA, a small initial training body may be sufficient.

One of the limitations of this study is that the variability of individual subject classifications was not analysed. There is a possibility that, while the classification accuracy might not vary a large amount between different repeats of the random subsamples, the individual BOE of each subject might alter significantly. For example, a subject could hypothetically be classified as 90% B(OA) when using one subsample and 55% when using another, and while the classification accuracies might be the same, the individual BOE has significantly changed. If the classifier is being used as an objective tool to quantify the overall change in knee biomechanics, it is important to know how dependent this objectively measured change might be on the specific ‘training body’ of OA.

3.7.3 Assess the Reliability of the LOO Cross-Validation Technique as an Estimate of Classification Accuracy.

Classification techniques are subject to over-fitting; which occurs when the classification rules model relationships which only exist by chance within the sampled training data. In this instance, the accuracy of a classifier was, for example, 95% accurate at classifying the ‘training’ data, might perform very poorly

The LOO cross-validation algorithm is a useful method for validating the robustness of the classification because it maximises the amount of training data used but has been criticised in certain applications as it can under-estimate classification errors. This is in part due to the self-similarity of the n-1 training cohort which is used to validate the remaining one subject.

A non-exhaustive leave-p-out classification was therefore performed with 500 repeats for each value of p. The classification accuracy was consistent between different leave-out groups, tested as a maximum of leave-20-out. This suggests that the LOO classification accuracy may be a reliable estimate of classification accuracy for this dataset. This could reflect the fact that, unlike other classification techniques, such as neural networks, there is only one iteration and over-fitting may, therefore, be less of a concern. Neural networks use a backpropagation algorithm in order to make small modifications to the classifier rules until an optimal classification is achieved. This may result in overly complex models which aren’t reflective the real-world data. By comparison, the resultant DST classification is achieved due to a combination of the *a priori* definition of classifier control parameters, alongside the application of imprecise probability using Dempster-Shafer Theory. This classification technique is still subject to over-fitting. One example in the use of Pearson’s correlation coefficient is used in the definition of the sigmoid function: a high correlation could be present in the dataset and hence be highly weighted within the classification. If this correlation was present by pure chance, it would be successful in classifying the training data but would perform poorly on real-world test data.

3.7.4 Does the Inclusion of Mediolateral GRF Force and Knee Joint Moments Have a Significant Impact the Ability to Classify Osteoarthritic Subjects?

The tibiofemoral joint knee kinetics during gait were calculated by modifying the previously written code which calculates knee kinematics using the Grood and Suntay JCS approach. The calculation of knee joint forces requires knowledge of the force platform relative to the lab GCS, allowing the centre of pressure to be transformed from the force platform LCS to the lab GCS. The eventual cohort was much smaller than the previous dataset, consisting of 20 OA and 18NP subjects.

The classification accuracy was improved significantly by the addition of knee kinetics and the mediolateral GRF, achieving the greatest LOO accuracy of 100%. Two PCs of the mediolateral force, one PC of the knee flexion moment, and one PC of the internal knee moment were all within the top 10 ranked variables of the classification. Surprisingly, the most influential PC of the knee adduction moment was ranked at only 11th.

3.8 Clinical Summary

Osteoarthritis of the knee can be the result of, and can result in, altered biomechanics during ADLs. Measuring and monitoring these biomechanical changes allows the quantification of a construct of physical function that isn't currently collected during self-reported assessments. There are obvious challenges to the adoption of HMA, notably the cost of equipment, level of expertise required, time and other resource costs for each assessment. There are also numerous analytical challenges once the data is collected.

Consider the following three challenges:

1. Data reduction: Kinetics and kinematics are temporal, and traditionally normalised to 100% of the gait cycle. There are too many values per variable for statistical analysis/comparisons to be made. The data must, therefore, be reduced in some way.
2. Data classification/summation: Joint biomechanics are, by their nature, highly influenced by changes in other joints. Hip, knee and ankle kinetics and kinematics are therefore very influenced by knee OA, and each only tells one part of the story. It is therefore in some instances useful to provide a summary measure which considered these factors.
3. Validation: Classification techniques are subject to over-fitting, or rather inferring relationships which incidentally exist in the example data, and don't exist in the whole population. Similarities can be drawn between over-fitting and type II errors. Tests of robustness are often challenging as datasets are typically small, however, are critical for the prevention of spurious findings.

3.8.1 Key methodological developments:

Data reduction - This chapter has explored the application of PCA, a data reduction technique, to first reduce a large biomechanical dataset before further classifying. This can be seen as an alternative technique to subjectively choosing discrete metrics from individual waveforms – e.g. 'peak knee flexion during swing', 'knee ROM during gait', etc. The research presented also reveals that previously adopted methods of interpreting

PCs by considering the ‘factor loadings’, which represent how well each PC reconstructs the original data at each time point, can actually result in misinterpretations (see Section 6.1.2). Data reconstruction is therefore recommended to aid clinical interpretations of the biomechanical features characterised by PCA.

Data classification: The framework of DST theory adopted to summate biomechanical information. Small methodological changes were made to the control parameters to counteract a bias which is introduced when classifying certain sets of data, particularly when data is of different scales of magnitude, data set sizes are uneven, and variance is not equal between the OA and NP groups.

Validation: Previously, a LOO cross-validation technique was applied to assess the classification performance on unseen data. It appears that less conservative techniques such as leave-10-out cross-validation results in similar classification accuracies. This increases the confidence in the validity of the results, however, further tests on a larger cohort of unseen data would be required to further validate the classification.

3.8.2 Key clinical findings

. When classifying 85OA and 38NP subjects using the same input variables defined by (Jones, 2004), the variables/features which were ranked highest in the classification of OA were:

1. The magnitude of the anteroposterior GRF peaks during weight acceptance and push-off (93.5% accuracy).
2. The ROM of the knee during stance phase, which is correlated with both a reduced and a later peak knee flexion (90.2% accuracy)
3. The reduction in the double peak, and hence trough of the vertical GRF. This was also correlated with a slower rate of weight acceptance (90.2% accuracy).
4. Cadence, i.e. steps per minute (80.5% accuracy)
5. Mediolateral knee width (79.7% accuracy).

The inclusion of mediolateral GRF and joint moments on a smaller cohort of 20 OA and 18NP subjects revealed that most discriminative knee kinetic feature was a reduction in external knee flexion and extension moments. This correlates with the previously observed reduction in anteroposterior GRFs. The next most discriminative feature was perhaps surprisingly the internal/external knee moment. Transverse knee moments are infrequently reported in the literature, perhaps due to the associated measurement errors. The reduction in the magnitude of transverse moments in OA subjects has previously been reported (Brandon and Deluzio, 2011), however on average the OA subjects were well within the normative range of NP subjects. The identified biomechanical feature appeared to represent a relatively more external knee moment during the first half of stance, and more internal during swing phase. This is perhaps a challenging feature to interpret – clinically if the knee moment is “relatively more external” than NP subjects, however, is still an internal joint moment, the joint moment would be termed as “reduced”. A reduction in gait velocity is strongly correlated with a reduction in the magnitude of knee joint kinetics, perhaps most easily explained by the decrease in accelerations and decelerations of the body.

Chapter 4 - Classification of Osteoarthritic Hip, Knee and Ankle Gait Biomechanics

4.1 Introduction

One of the overall aims of this thesis is to develop a robust classification of osteoarthritic biomechanical function during level gait. The previous chapter outlined limitations of the Cardiff Classifier control variables, and demonstrated appropriateness of LOO classification and classification cohort sizes. It also outlined the usefulness of PCA reconstruction in greater understanding and contextualising the contribution and clinical relevance of the PCs selected for further analysis.

One of the limitations of the previous chapter was the consideration of only the kinematics and kinetics of the affected knee. OA of the knee is a bilateral disease, which is known to affect the hip, knee and ankle kinematics and kinetics of both the affected and unaffected leg (Metcalfe *et al.*, 2013).

As previously mentioned and further described in Section 3.2, in 2006 additional markers were added to the TKR data collection SOP at Cardiff University which enables the calculation of hip and ankle kinematics and kinetics in addition to those of the knee. The ultimate aim is to quantify biomechanical recovery following TKR, and hence changes in biomechanical function of the hip and the ankle may be clinically important and relevant in assessing functional outcomes.

This chapter aims to explore and develop the techniques for reducing and classifying biomechanical level gait data, and follows on from the previous chapter as a continual development of methods. This chapter utilises a software package called Visual3D (C-motion). While the software could have been custom tailored to follow the joint segment definition and tracking methods of the previous chapter, the opportunity has instead been

taken to adopt more advanced segment definitions, which are more in-line with current research practices.

Some of the changes in methodology are likely to affect the resulting knee kinematics and kinetics. One of the changes in methodology is the definition of the anatomical axis of the femur. Previously, the line between the upper border of the greater trochanter to the midpoint of the medial and lateral epicondyles was used to define the proximal/distal axis of the femur. It isn't known to what extent these changes will alter resulting biomechanical variables. As PCA is sensitive to differences in shape as well as magnitude, it also isn't clear to what extent this would affect the PC scores used for classification. One of the aims of this chapter is therefore to assess the appropriateness of previously defined PCs in representing variance between subjects.

The study reported in this chapter has two aims and hypotheses as follows:

Aim 1: Assess the appropriateness of previously defined PC in representing variance between subjects collected with the updated methodology.

Hypothesis 1: The methodological changes will affect the appropriateness of the PCs of the knee kinematics and kinetics, but shouldn't affect those for the GRF.

One of the considerations when classifying data between two data groups is the consideration of differences in demographics. The matching of relevant subject demographics is generally recommended when attempting to identify statistical differences between groups. In the UK, the median age of patients undergoing TKR surgery is 70 years, with an interquartile range of 64-76 years old. Numerous studies have shown biomechanical changes in level gait associated with ageing in NP subjects. These changes are well-reviewed within Prince *et al.* (1997) and Moreland *et al.* (2004), which highlight physiological changes such as reduced muscle mass and strength, gait velocity, sagittal lower limb ROM, and sagittal plane moments. These may be a result of the vast amount of biological changes in the joint due to a natural ageing process, such

as sarcopenia; however, they could also be due to underlying comorbidities which the subject may not be aware of.

The ultimate aim of this thesis is to objectively define changes in biomechanical function following TKR surgery. The presence of age-related biomechanical changes in NP subjects raises an important consideration when classifying function. Therefore, should improvement in biomechanical function be defined as:

1. Restoration towards NP young/middle-aged gait biomechanics?
2. Restoration towards NP age-matched/elderly gait biomechanics?
3. Restoration towards an envelope of healthy gait biomechanics, which includes a range of ages of NP gait?

Previous research in classifying osteoarthritic level gait kinematics, presented within Watling (2014) also found statistically significant biomechanical changes relating to ageing. This motivated a decision to consider only young and middle-aged non-pathological subjects within the non-pathological training body, such that it represented a homogenous control cohort of NP biomechanics.

While some biomechanical characteristics of ageing are similar to those of OA, many features of OA are more specific to the mechanisms of the disease, and are not reported in the elderly subjects. Figure 4.1 shows a simplified illustration of the biomechanics of healthy (young/middle-aged), elderly healthy, and osteoarthritic function. Healthy subjects are symbolically shown as the smallest circle, as we would expect the greatest homogeneity within this cohort. As introduced in Section 2.1, OA is a very complex and multifactorial disease associated with several comorbidities and degradation of surrounding joints, often because of biomechanical adaptations. It is therefore symbolically shown as the largest circle.

The figure displays commonalities and differences between:

1. Healthy vs elderly gait
2. Healthy vs osteoarthritic gait

3. Elderly vs osteoarthritic gait

We can see that, if we combine elderly and healthy gait, and classify differences between this combined cohort and OA gait, the top-ranked discriminating variables should theoretically exclude shared characteristics between OA and either of the healthy cohorts, as shown in Figure 4.1. This study therefore proposes the use of a combined healthy adult cohort across a broader age range as that investigated by Watling (2014).

Aim 2: Assess the validity of a combined healthy and elderly cohort in classifying OA subjects.

Hypothesis 2: A combined elderly and healthy cohort will favour input variables which discriminate between OA and healthy, but aren't related to ageing. The belief of healthy, B(NP), will not, therefore, be significantly correlated to ageing.

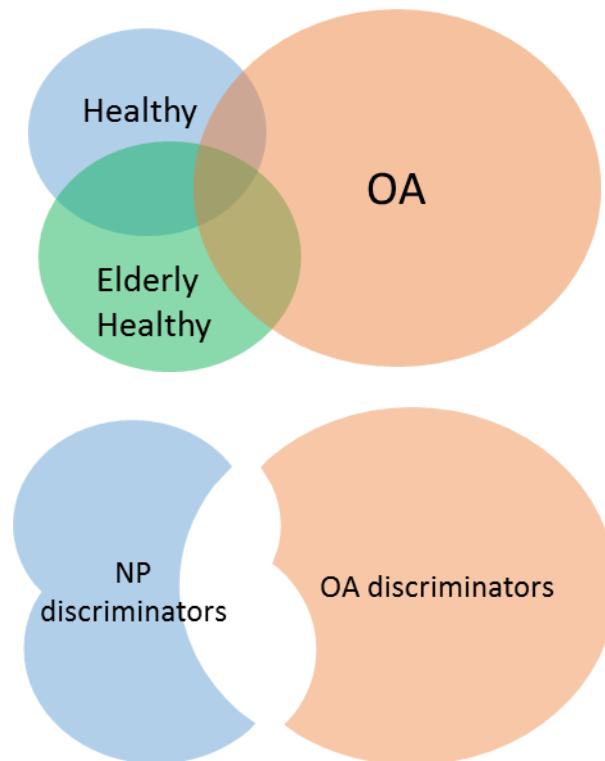


Figure 4.1 A) Simplified illustration of an overlapping in function between healthy, elderly and osteoarthritic gait. B) An illustration of what would theoretically be the discriminatory variables of function, when combining all age ranges of healthy subjects.

4.2 Methodology

4.2.1 Marker Placement

The marker set used within this study was based upon the CAST marker protocol (Benedetti *et al.*, 1998, Cappozzo *et al.*, 1995). The markers form an addition to that used within the previous chapter and are displayed within Figure 4.2. Two rigid marker clusters are placed bilaterally on the shank and thigh, as described previously. As opposed to defining anatomical axes using the pointer technique, anatomical landmarks are now defined using markers placed on anatomical landmarks and captured during a static calibration trial.

One of the benefits of using skin-based markers, as opposed to a pointer in determining the location of anatomical landmarks, is that it is possible to leave these markers on for the duration of the trial. Two common challenges in carrying out the assessment are the gradual movement of the rigid marker clusters relative to the thigh, and the tracking markers coming off the subject during a trial. If either of the two occur, it is necessary to re-calibrate the anatomical model and to create another static calibration file including all the markers. This can be very time-consuming, particularly as the model requires 16 anatomical landmarks.

4.2.2 Defining the Pelvis

Previous work within Cardiff University has defined the pelvis in three stages (Watling, 2014):

- Create a CODA pelvis using the anterior superior iliac spine (ASIS) and posterior superior iliac spine (PSIS) markers
 - This will create virtual hip joint centre landmarks
 - This pelvis is not used for joint kinematics, but is used for joint kinetics
- Create a virtual iliac crest marker which is directly vertical within the lab axes of the greater trochanter (GT) markers

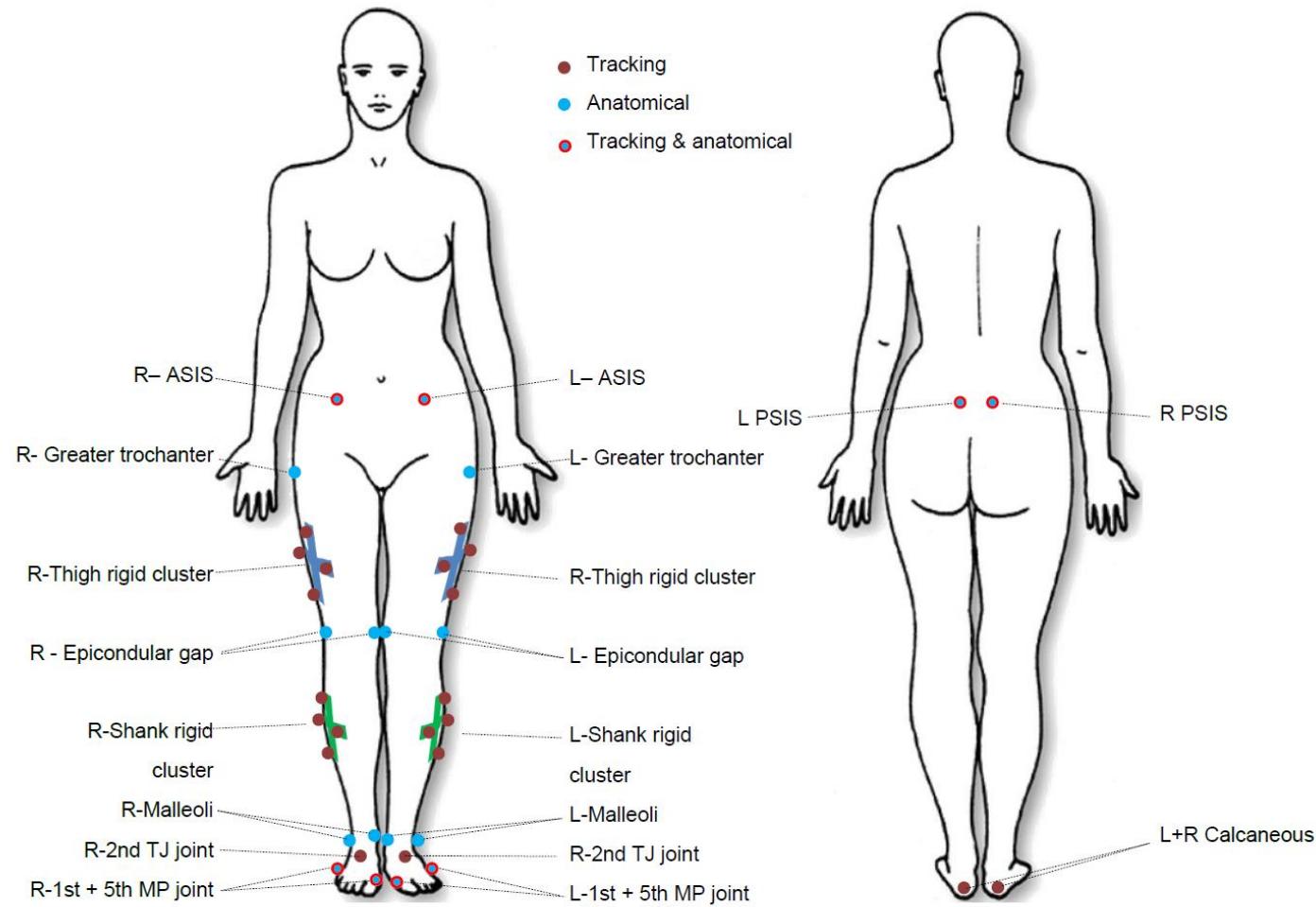


Figure 4.2 CAST protocol marker set used within this chapter, illustrating tracking (red), anatomical (blue), and joint anatomical and tracking (blue with red circle).

- Create a Visual3D pelvis using the GT and the iliac crest markers
 - This pelvis is used for the calculation of hip kinematics

By defining the pelvis segment using virtual iliac crest markers which are vertically aligned above the GTs, the pelvis tilt relative to the lab GCS is essentially set to zero during the static trial. The coronal plane angle, which shall be referred to as the pelvic obliquity, is also dependent on the relative height of the GTs.

A study by Della Croce *et al.* (1999), examining the intra- and inter-examiner precision of identifying the GT, found intra-examiner precision of 11.1m and inter-examiner precision of 9.8mm in the superior/inferior direction. Each assessment of precision was carried out six times by physical therapists, on two able bodies with BMIs of 24 and 19.8. This same study found much lower imprecision in this anatomical plane when identifying the LASIS, but much higher when identifying the RASIS. No information within the literature was found on the intra and inter-examiner precision of GT and ASIS markers on obese subjects.

From personal experience, the superior-inferior identification of the ASIS has been more challenging to locate than that of the GT on obese subjects. It was therefore decided that the relative height of the virtual iliac crest markers should be offset such that the relative height between each side matches that of the ASIS markers.

4.2.3 Hip Joint Centre Definition

One of the challenges of defining biomechanical models is the definition of the hip joint centre (HJC), which can be used to define the centre of rotation of the hip joint. Within traditional lower body gait models, the HJC defines the centre of the superior end of the femur. This therefore affects the angle of the thigh segment – affecting the hip and the knee angles kinetics and kinematics.

Methods for estimating the location of the HJC can be categorised into two groups; functional methods based on movement, and predictive methods based on regressive equations using anthropometric measurements. A recent systematic review by Kainz *et*

al. (2015) found that functional methods improved accuracy of HJC location on healthy subjects, and may reduce between tester (inter-operator) and between session (intra-operator) variations due to the removal of the reliance on the manual palpation of anatomical landmarks. This aforementioned study builds upon previous ISB guidelines (Wu *et al.*, 2002), and concludes that the current best-performing functional method only performs marginally better than the best-performing predictive method, and only in ideal conditions. It is therefore doubtful as to whether it is worthwhile to collect the additional functional calibration trials, particularly as subjects with late-stage knee OA will likely have difficulty performing the required movements.

There were no reports found on the effect of obesity on the accurate of HJC estimation using functional sphere-fitting methods or predictive regression equations. One small study found significant differences between functional and predictive methods, resulting in significant changes in hip kinematics and kinetics (Chohan *et al.*, 2013). Within this study, there was no gold standard of HJC location (such as radiographic measurements), therefore it isn't possible to conclude which method is preferable. It was noted that obese subjects often found it difficult to perform the movement required to calculate the functional HJC, and that the anatomical landmarks required for predictive HJC definitions were difficult to locate. Predictive methods also couldn't account for pelvis asymmetry.

This thesis builds upon the research of Watling (2014) whom used the Bell and Brand predictive regression equations (Bell *et al.*, 1989, Bell *et al.*, 1990), which are used by default within Visual3D. This method uses the inter-anterior superior iliac spine distance to estimate the HJC using the following regression:

$$HJC \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \pm 0.36 \times PW \\ -0.19 \times PW \\ -0.3 \times PW \end{pmatrix} \quad (4.1)$$

Where the x, y and z coordinates relate to the mediolateral, anteroposterior, and superoinferior coordinates relative to the origin of the pelvis segment, and PW refers to the distance between the left and right ASIS.

A number of studies have recently found improved predictive accuracy of the Harrington equations over other predictive techniques, such as the Bell and Brand regression (Sangeux *et al.*, 2014, Andersen *et al.*, 2013), and it has therefore been recommended as the best current predictive method in a systematic review by (Kainz *et al.*, 2015). The Harrington equations were defined from the MRI scans of eight adults (ages 23-40), 14 healthy children, and ten children with cerebral palsy (Harrington *et al.*, 2007).

The Harrington equation defines the HJC as:

$$HJC \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0.33PW + 0.0073 \\ -0.24PD - 0.0099 \\ -0.30PW - 0.0109 \end{pmatrix} \quad (4.2)$$

Where the x, y and z coordinates relate to the mediolateral, anteroposterior, and superoinferior coordinates relative to the origin of the pelvis segment, and PW refers to the distance between the left and right ASIS, the pelvic depth (PD) is the distance between the midpoints of the line segments connecting the two ASIS and the two PSIS (Harrington *et al.*, 2007).

4.2.4 Hip, Knee and Ankle Axis Definitions

Joint segments were defined following ISB recommendations for joint coordinate systems (Wu *et al.*, 2002), except for the following differences:

- The knee joint centre was defined using the midpoint of the epicondylar gap, as opposed to using the knee epicondyles themselves
- The tibia coordinate system was defined using the epicondylar gap instead of the tibial condyles
- The foot segment was normalised to the floor, i.e. was assumed to be parallel to the floor during the static trial. It was decided that this method of normalisation was suitable when comparing barefoot NP and OA subjects. The assumption that inversion/eversion was deemed to be valid for subjects without severe deformity.

4.2.5 Upsampling to the Analogue Capture Frequency

When calculating gait kinematics, it is necessary to locate the timing of heel strikes so that gait kinematics can be normalised as a percentage of heel strike. The most accurate method of defining this gait event is to use a threshold on the force platform data. A minimum threshold of 20N has previously been used to detect the initial heel strike and toe-off events. Until recently, there have only been two force platforms within the motion analysis laboratory at Cardiff University. In the absence of force platform data to define the following heel strike, the “pattern recognition” feature within Visual3D had previously been used to define the following heel strike.

This feature uses the technique of Stanhope *et al.* (1990). The pattern of the proximal end of the foot segment is taken at the first heel strike, and a pattern recognition algorithm is then used to detect the following HS. Within Visual3D, the analogue data is used to locate the first heel strike and the next closest kinematic frame is taken. The marker capture data within this study was sampled at 60hz, and the analogue data was sampled at 1080Hz. There are therefore 18 analogue samples for every one marker sample. The theoretical precision of the resultant HS identification is therefore:

$$\pm \frac{9}{1080} = \pm 0.0083 \text{ seconds} \quad (4.3)$$

The algorithm then attempts to match the kinematics at this point using pattern recognition to determine the second heel strike. Had the initial pattern been identified accurately, the precision for the identification of the second HS would again match that of Equation (4.3). However, as the pattern now used for matching might already be $\pm \frac{9}{1080}$ seconds out, the theoretical maximal precision for the identification of the second HS is:

$$\pm \frac{9}{1080} + \frac{9}{1080} = \pm 0.0167 \text{ seconds} \quad (4.4)$$

Remember, this is only considering one element of precision, and is not considering the accuracy of the 20N threshold, the variability of foot segment kinematics during walking, the robustness of the pattern-fitting algorithm and its associated detection threshold.

Considering only this aspect of precision, the precision of calculating the stride time can be seen as follows:

$$\pm 0.0083 \pm 0.0167 = \pm 0.025 \text{ seconds} \quad (4.5)$$

The resultant effect that these precisions might have on a single knee flexion waveform is modelled within Figure 4.3A. The different timestamps at which each HS might have been identified within the precision boundaries have been selected, and each waveform has then been normalised over 100% of the gait cycle. Not only the shape of the waveform is changed noticeably, but also the timings of gait events such as the peak knee flexion angle during swing phase.

Within this study, the marker data was upsampled to the analogue capture frequency, such that the correct timestamp can be assigned to the initial HS (identified from the vertical GRF), and hence a pattern more representative of the kinematics at this HS can be adopted to identify the second HS. This was achieved within Visual3D (C-motion, USA), using an in-build algorithm called “Upsample_Point_Data_To_Analog_Rate”, which adopts the cubic spline method. Marker data was low-pass filtered 6Hz before upsampling, and Fast Fourier Transforms were calculated before and after this process to ensure the frequency content of the signal was unchanged.

Following this change, the HS is identified at the point at which the HS passed the threshold, and hence theoretically precise to:

$$\pm \frac{0.5}{\text{Analogue sampling frequnecy}} = \pm \frac{1}{2160} = \pm 0.41 \times 10^{-3} \text{ seconds} \quad (4.6)$$

The second HS is still subject to the same marker imprecision of $\pm \frac{9}{1080}$ seconds, however as the HS profile used within the pattern matching is more precision, the precision would be:

$$\pm \frac{1}{2160} \pm \frac{9}{1080} = \pm \frac{19}{2160} = \pm 0.0088 \text{ seconds} \quad (4.7)$$

The resultant theoretical effect of this imprecision is modelled within Figure 4.3B for comparison.

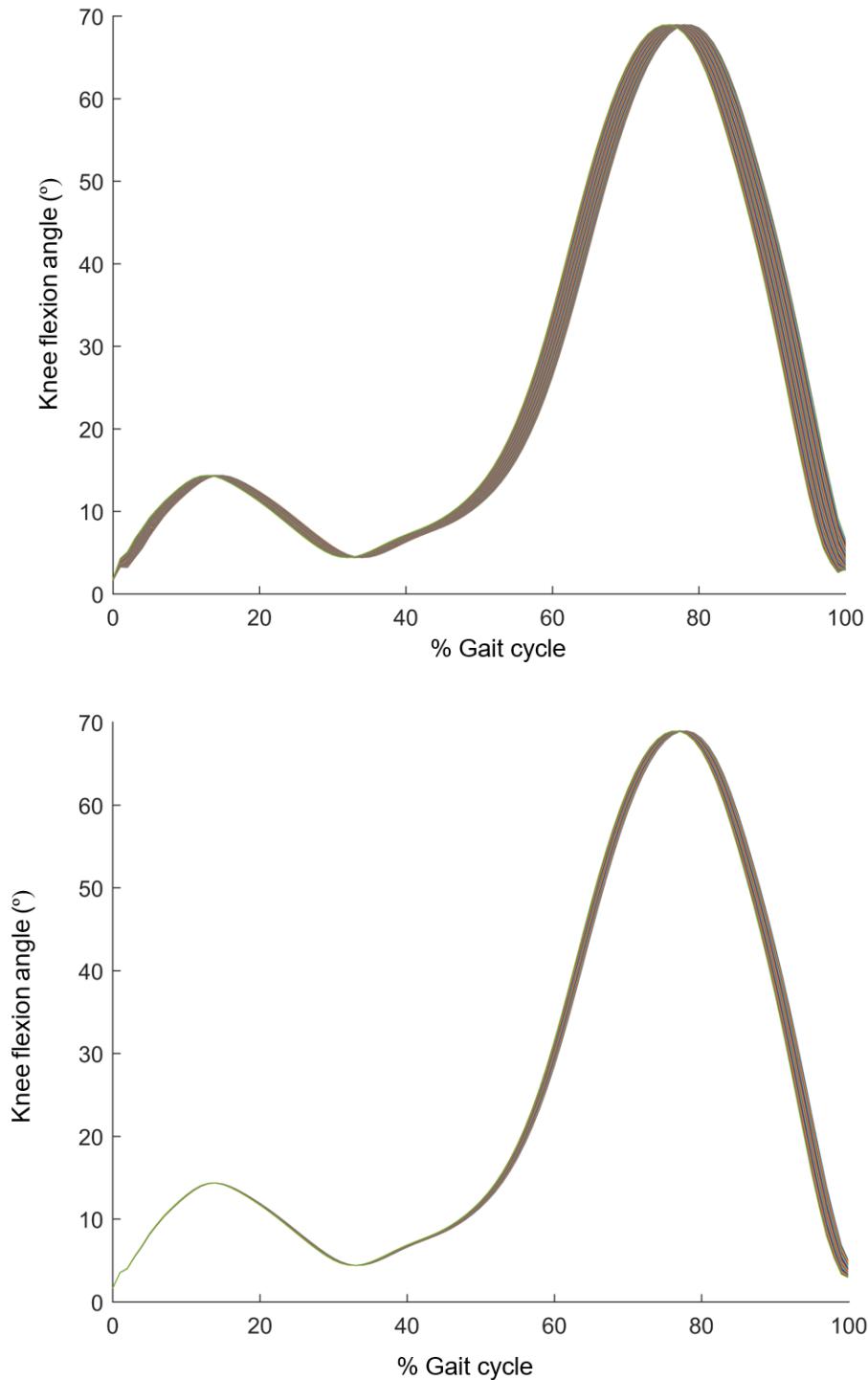


Figure 4.3 Example simulation of the effect imprecision in HS identification can have on the resultant kinematic waveform. A single knee flexion waveform has been cut and normalised to 101 points. Every permutation of the possible HS locations within the precision boundaries have been considered for:

- A) Before upsampling: $\pm \frac{9}{1080}$ seconds for the first HS, and $\pm \frac{18}{1080}$ seconds for the following HS, and
- B) Following upsampling: $\pm \frac{1}{2160}$ seconds for the first HS and $\pm \frac{19}{2160}$ seconds for the following HS.

4.2.6 Filtering Data

Electrical interference, instrumental errors in the digitization of retroreflective markers in 3D space, and STA can all result in errors within motion analysis data (Winter *et al.*, 1974). These errors are often referred to as “noise” within the signal, and attempts are made to improve the signal to noise ratio by “smoothing” or “filtering” the data. While a number of techniques have been applied to motion analysis data, such as cubic spline fitting (McLaughlin *et al.*, 1977), Fourier series (Hatze, 1981), and polynomial smoothing (Pezzack *et al.*, 1977), a large proportion of studies use a fourth-order zero-lag Butterworth filter (Sinclair *et al.*, 2013).

Digital filtering of marker kinematics relies on the principle that the actual kinematics of the bones relative to one another during locomotion are of a lower frequency than the signal noise. To minimise any detrimental effects on the true signal, it is necessary to define a filter cut-off frequency above that of the highest expected frequency component of the true kinematics. The frequency components of the true relative bone kinematics would depend on the velocity of the task being carried out. For example, it would be expected that running would have large amplitudes of higher frequency components than walking.

The frequency content of normal gait is generally considered to be around 4-6Hz (Angeloni *et al.*, 1994) (Winter *et al.*, 1974). Previously, within this research group, a cut-off frequency of 6Hz has been applied (Watling, 2014), which is commonplace in the context of gait analysis (Kirkwood *et al.*, 2011). The choice of a fixed cut-off frequency has previously been criticised (Chiari *et al.*, 2005), largely due to the fact that frequency content varies between markers, and also during activities. This may be less of an issue in the context of gait, as Winter *et al.* (1974) found markers to fall within a narrow boundary of frequency content.

One of the highest frequency components found by Winter *et al.* (1974) was that of the heel, at 6Hz. For demonstration purposes, Figure 4.2 displays the frequency content of

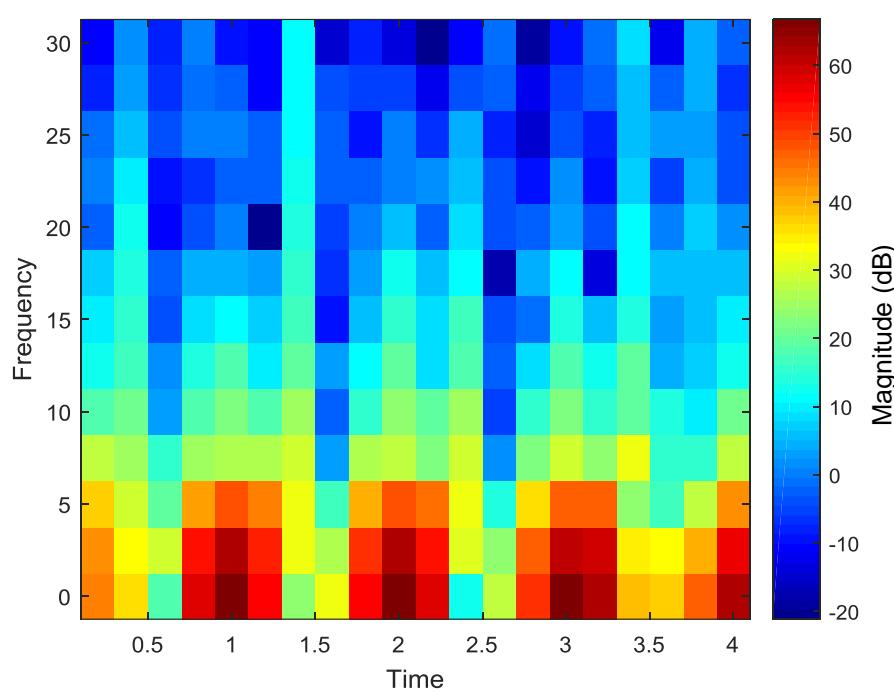
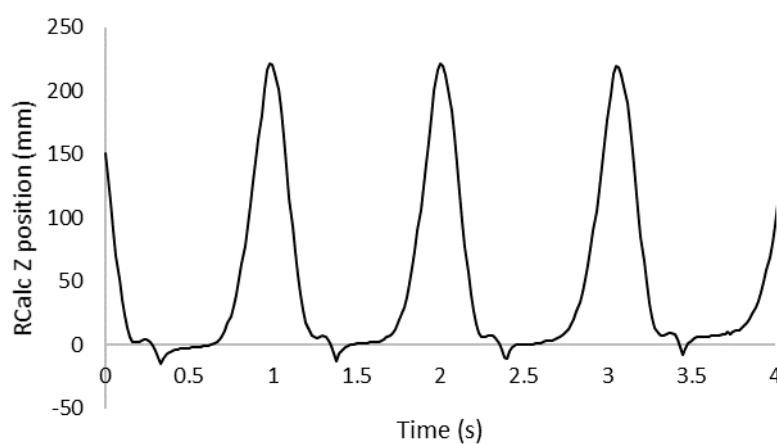
A**B**

Figure 4.4 Example of the frequency components of a vertical positions of the right calcaneus (RCalc) marker during a gait trial. A) A spectrogram of the change in frequency components over time. Red corresponds to higher amplitudes, and blue lower amplitudes. B) The positional data of the heel marker in the vertical axis over the gait trial. Multiple heel strikes occur during this trial.

a right calcaneus marker that has been plotted during a gait trial of a healthy adult. There are four heel strikes within this trial, occurring as the heel marker approaches zero in the z coordinates. A spectrogram of the frequency content is displayed in Figure 4.2A, in which frequencies are the highest towards the peak Z coordinate value. Within Figure 4.2B, towards heel strike at the minimum points of the curve, there appears to be some additional high-frequency content other than that expected, which is likely due to the STA and inertia of the marker at heel strike.

The effect of a fourth-order zero-lag low-pass Butterworth filter, with a cut-off frequency of 6hz is displayed in Figure 4.5. While subtle, changes can be seen at the peak of the vertical position, which is slightly reduced by the filtering. The digital filtering has, however, dampened the artefact identified towards the minimum curve around heel strike.

This study followed recommendations regarding the filtering of GRF data with the same filter parameters as that of marker kinematics, in order to reduce errors during inverse dynamic analysis (Kristianslund *et al.*, 2012, Van den Bogert and De Koning, 1996). Raw GRF was then separately filtered using a low-pass filter of 25Hz before analysis.

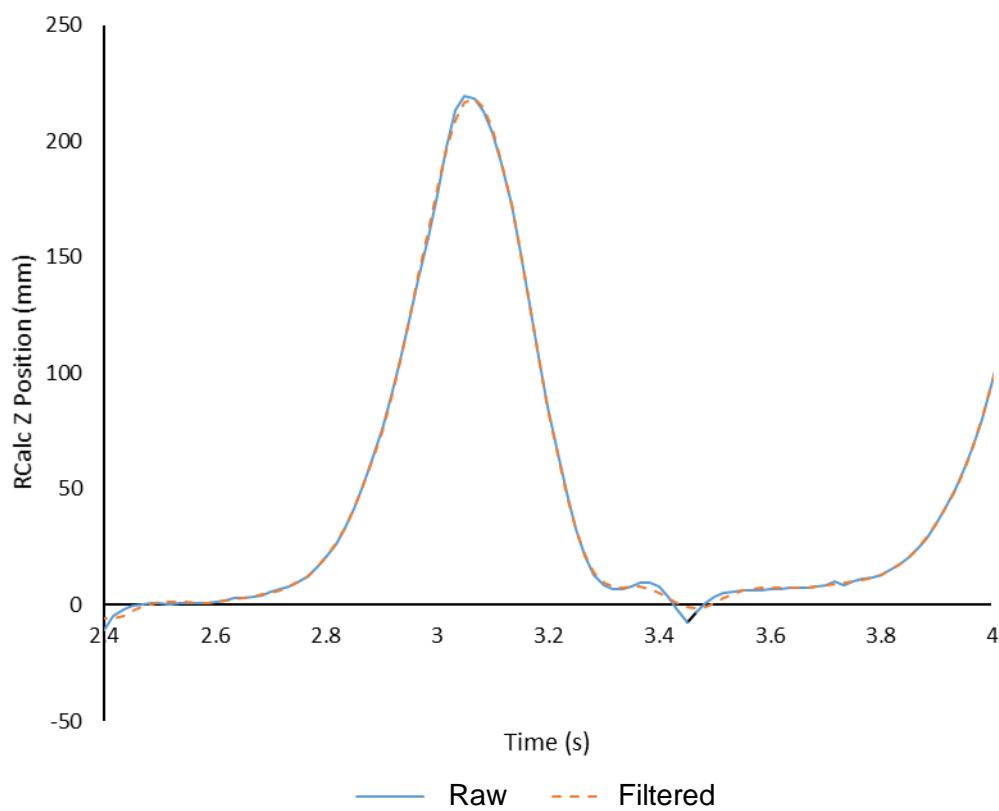


Figure 4.5 Demonstration of the effect of a fourth-order zero-lag low-pass Butterworth filter with a cut-off frequency of 6hz on right calcaneus (RCalc) marker during a gait trial.

4.2.7 Comparison of Previously Defined PCs

One of the aims of this chapter is to assess the validity of the PCs defined in the previous chapter, describing differences between subjects using the updated methodology. The pointer technique used in the previous chapter requires the presence of the captured pointer data files which, as previously mentioned, were not consistently captured for NP subjects at Cardiff University. The methods presented in this chapter require the presence of additional markers which define the pelvis (and hence HJC), and foot, which were added to the data collection in 2006. In total, nine NP and nine OA subjects who had been processed using both techniques were collated.

The data for knee kinematics, knee kinetics and GRF were collated, and the previously defined PCs and associated eigenvectors presented in Section 3.6.2 and Section 3.6.6 were used in order to calculate the PC scores for the same subjects processed using the two different techniques.

Statistical analysis was performed within SPSS (IBM Corp, USA). Each variable was tested for normality using a Shapiro-Wilk test, resulting in a total of four significant results; hence indicating the assumption of normal distribution was not valid for these variables. Data meeting the assumptions of parametric tests was analysed using a Paired-Samples T-Test, while the remaining four were tested using a Wilcoxon paired signed rank test.

4.2.8 Initial PCA Selection

Within the previous chapter, PCs were retained using the factor loading retention rule suggested by Comrey and Lee (2013) and adopted by Jones (2004). The results within the Section 3.6.2 and Section 3.6.6 showed that in some instances, the first PC can represent large magnitude differences throughout the duration of a waveform. While these may reconstruct a large amount of the total variance in the data, this may not always be the most clinically significant element of the variance. For example, had the rule been strictly followed, only the first and third PCs of the mediolateral GRF would have been selected. The first PC represented a large difference in variance throughout the duration of stance phase, whereas the second PC represented the level of 'double

'peak' within the waveform, alongside the gradient of the waveform during loading response and terminal stance. Because both PCs were influential within the same area, they are essentially competing for selection. The first PC consistently represented more than 50% of the total variance at each of these points of the gait cycle, therefore the second PC. The results shown in Table 3.8, however, suggested that the second PC is far more accurate in the classification of OA function (81.6%), as opposed to the first PC (55.3%).

This highlights the potential flaws of making the following assumption: the component which represents a larger amount of variance between all subjects is likely to be more clinically significant in the classification of pathological function.

Within this chapter, a new rule will be investigated: the first three PCs will always be selected for each waveform, and then further components will be selected only if they meet the previously mentioned factor loading threshold.

4.2.9 Further PC Retention Using Classification Ranking

PCA of ankle, knee and hip kinematic and kinetic waveforms can result in a large number of variables. Previously, other authors have used the DST classifier itself in order to further reduce the number of selected PCs (Warner *et al.*, 2015, Metcalfe, 2014). While this can be a very useful method of further reduction and PC retention, it also poses a potential bias in the final classification accuracy.

To demonstrate this, Figure 4.6A shows the classification results of a randomly generated dataset of 50 variables, and 25 subjects in each class, resulting in a LOO accuracy of 52%. The variables were then ordered based on the ranking of each variable's accuracy within the original classification, and split into datasets of different sizes based on the top-ranked variables. The classification results of the new datasets of top-ranked variables are compared within Figure 4.6B.

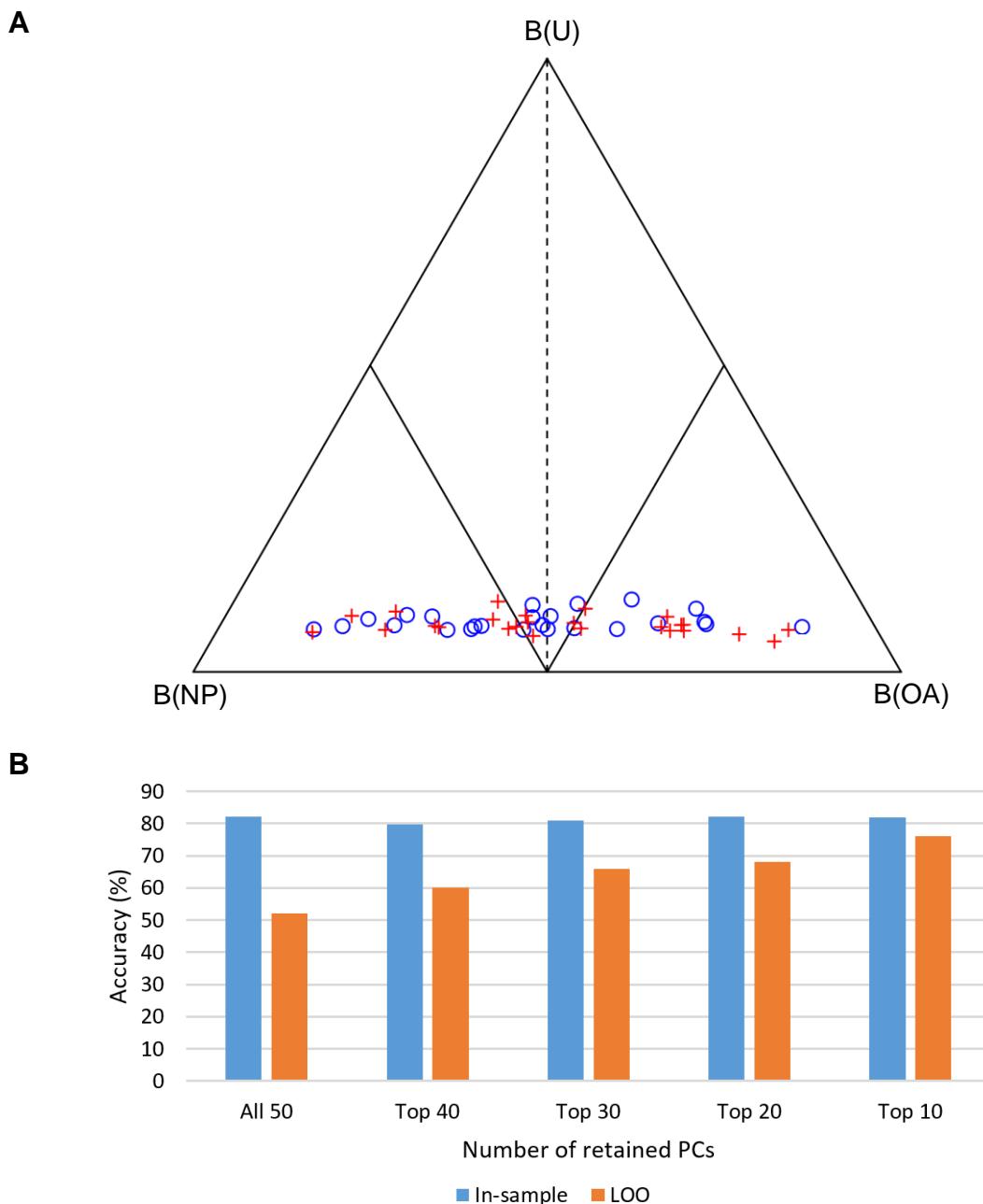


Figure 4.6 Comparison of in-sample and leave-one-out (LOO) classification accuracy of a dataset of randomly generated variables with 25 subjects in each group. A) The original classification simplex plot using all 50 variables. B) Comparison of in-sample and LOO classification accuracy when using the original classification variable ranking to reduce the number of selected PCs.

While this is anecdotal evidence of one randomly generated dataset, it matches what we would theoretically expect. The initial in-sample classification accuracy is biased, and therefore, as discussed in Section 3.5.4, it is necessary to perform cross-validation such as LOO. As expected, having found positive results using the LOO cross-validation within Section 3.6.4, the LOO accuracy is close to 50%. Within any relatively small, randomly generated dataset, some variables will be correlated with the marker label and others

will not. In other words, there are relationships within the data which wouldn't be present within an infinitely large dataset.

If only the variables which negatively influence classifier performance are eliminated, the classification accuracy can be increased through the inclusion of randomly correlated variables, and the exclusion of uncorrelated variables. While accuracy will appear to increase, this is likely to be much less indicative of performance on actual test data. To account for this, a different will be used:

1. Data will be randomised and split into two equal sets.
2. Both sets will be classified separately and the top-ranked variables will be calculated and compared.
3. The average “classification accuracy” of the individual variable, between the two datasets, will be taken. Variables will be ranked based on this average between two separate datasets.

It is hoped that this will reduce, as opposed to eliminate, the bias of this technique of variable retention. Ideally, if there were a larger number of subjects within this dataset, the retained PCs would be used on a completely separate data pool in order to train the classifier and evaluate classifier robustness.

4.3 Results and Discussion

4.3.1 Subject Demographics

A total of 31 NP and 41 OA subjects were included in the study. The subject demographics are displayed in Table 4.1. The NP control subjects in this study are, on average, younger, with a lower body mass and BMI.

Matching of relevant subject demographics is an important consideration for any research study. Numerous studies have demonstrated effects of ageing on gait biomechanics, as well as those of body mass and BMI. GRF data is normalised and described as either a percentage of body mass (%BW), and joint kinetics as a percentage of the product of body mass and height (%BW*h). This can't fully, however, account for the variations such as, kinematic and kinetic changes associated with obesity, increases in STA resulting in kinematic errors during the motion trial, and increased error in the identification of bony anatomical landmarks and identification of anatomic reference frames.

The ultimate aim of this study is to quantify the change in pathological function by quantifying a restoration towards healthy biomechanics. It is therefore important that the classification training body of NP function is a true representation of healthy level gait biomechanics. Both age and obesity are large risk factors of OA, alongside a huge number of other comorbidities which can affect locomotion. It is therefore a challenging process to collect age and BMI-matched subjects which meet the inclusion criteria listed in Section 3.2.1.

Table 4.1 Subject demographics for NP volunteer and patient groups used within this chapter.

	Number	BMI	Height/m	Mass/kg	Age/years	Gender
NP	31	24.6 (4.0)	1.69 (0.09)	70.32 (14.5)	40.7 (17.9)	12M 19F
OA	41	32.5 (6.4)	1.67 (0.10)	91.28 (20.3)	68.4 (8.6)	19M 22F

4.3.2 Assessing the Appropriateness of Previously Defined PC in Representing Variance Between Subjects Collected with the Updated Methodology

The results of the comparison in PC scores for the nine OA and nine NP subjects, whom had been processed using both techniques, are shown in Table 4.2. It was hypothesised that the changes in anatomical axis definitions would result in changes to knee kinematics and kinetics. It was not expected that the change in methods would affect the GRF data. The results from Table 4.2 support the hypothesis that the changes in methods result in significant differences between knee kinematics and kinetics. The methodology of this chapter defines the centre proximal end of the femur as the HJC, as opposed to the GT. The HJC is of course more medial to the GT, however, is often in a similar position in the sagittal and transverse plane. In fact, some studies have estimated the location of the HJC as one quarter of the distance between the ipsilateral to the contralateral GT (Weinhandl and O'Connor, 2010). The large changes in coronal plane alignment, and perhaps smaller changes in sagittal and transverse plane, result in a change in the local coordinate system (LCS) of the femur. Knee joint kinematics are calculated by defining the transformation between the tibial LCS and the femoral LCS. This might explain why significant differences in ab/adduction (coronal) angle PCs, were seen, alongside poor correlation between the two techniques. While two of the three PCs of the flexion (sagittal plane) angle were strongly correlated between techniques, there did appear to be a difference in the means of the two methods.

In the previous chapter, joint kinetics were calculated by resolving moments about the predicted centre of mass of the tibia. This assumes the moment caused by the GRF acts directly about the tibia, as well as negating the inertial effects of the surrounding limbs. This chapter uses the inverse dynamic calculations of Visual3D (C-Motion, USA) which uses an iterative algorithm to calculate proximal joint forces of the foot, the tibia, and the femur. Moments are then resolved about the COM of the segments to estimate the net joint moment.

As hypothesised the differences were smallest within the GRF data. Six of the eight GRF PCs were significantly correlated to one another ($r=0.78 - 0.97$, $p<0.05$). Only two resulted in significant differences between the means of the PCs calculated using the two methods, both being within the vertical (transverse) GRF.

Table 4.2 Table to compare PC scores of the selected biomechanical knee kinematics and kinetics, which have been calculated using both the methodology of Chapter 1 and that of Chapter 2. Variables which were significantly correlated to one another between the cohorts are shown in green. Significant differences between the PCs are shown in red.

Plane	Angle/ moment/ GRF	PC	Correlation	Sig.	Difference in Means	Sig. (2- tailed)
Coronal plane	Angle	PC1	0.54	0.02	6.50	0.01
			0.07	0.77	2.65	0.01
	Moment	PC1	0.47	0.05	3.15	0.14
		PC2	0.39	0.11	1.22	0.15
	GRF	PC1	0.80	0.00	-1.57	0.23
		PC2	0.80	0.00	-0.44	0.34
		PC3	0.78	0.00	-0.03	0.95
	Angle	PC1	0.82	0.00	4.95	0.00
		PC2	0.88	0.00	-2.72	0.00
		PC3	0.45	0.06	0.34	0.72
	Moment	PC1	0.69	0.00	-7.58	0.01
		PC2	0.04	0.88	-3.24	0.02
	GRF	PC1	0.97	0.00	0.42	0.37
		PC2	0.24	0.34	-0.69	0.55
Transverse	Angle	PC1	0.29	0.24	9.65	0.00
		PC2	0.68	0.00	-0.18	0.78
	Moment	PC1	0.46	0.05	-1.61	0.20
		PC2	-0.57	0.01	-5.75	0.03
	GRF	PC1	0.95	0.00	-0.16	0.77
		PC2	0.82	0.00	0.38	0.03
		PC3	0.09	0.71	3.14	0.00

4.3.3 Principal Component Analysis and Retention

PCs were defined for the 23 variables: ankle, hip and knee angles and moments, GRFs, and centre of pressure relative to the foot. This resulted in an initial total of 70 PCs being defined. The data was randomly split into two halves, one with 15 NP and 21 OA subjects, and the other with 16 NP and 20 OA subjects. Each dataset was classified and the LOO cross-validation accuracy and objective function were used to rank the input variables. The top n variables were selected for different values of n between one and 30. For each value of n , the number of total differences in the retained PCs (disregarding ranking) were calculated. Following this, the difference in rank of those retained PCs between the two datasets was calculated. The results are shown in Figure 4.7. If the top 18 variables are selected, there are few discrepancies between datasets. A selection of 18 variables also results in a variable to dataset size ratio of 1/3.

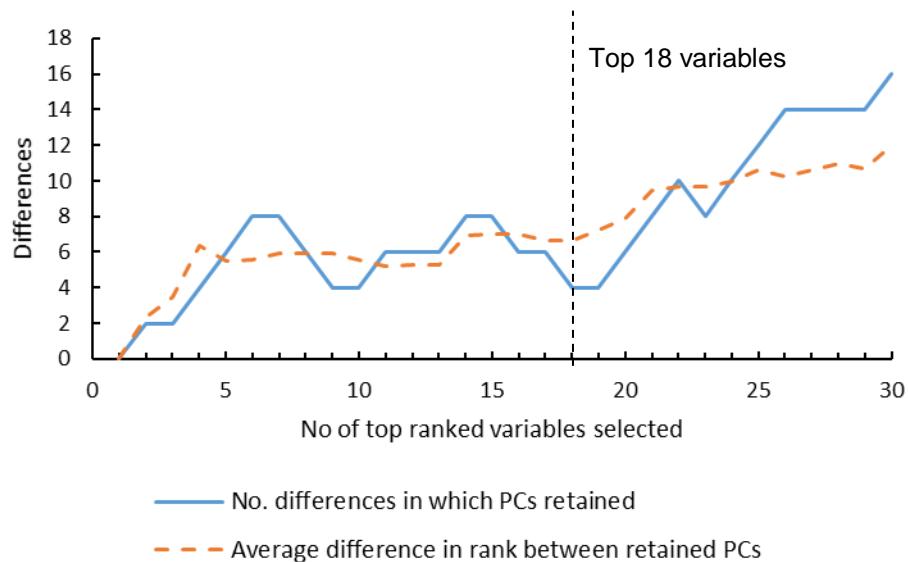


Figure 4.7 Illustration on differences between picking the top-ranked variables between the two datasets. The number of differences on variables selected when picking the top “n” variables is shown as a solid blue line. The average actual difference in ranking of those retained PCs between datasets is shown as a dashed orange line.

4.3.4 Classification Using Top 18 Ranked Variables

The LOO classification results of the final classification using the chosen 18 PCs is shown in Figure 4.8. The control parameters used to define the sigmoid function were as defined previously; $k_{C/S}$ and θ_{STD} . In the absence of an objective way of defining the uncertainty boundaries, the upper and lower boundaries were defined as 1 and 0.8 respectively. The DST classifier could discriminate between OA and NP gait biomechanics in all but 1 case, resulting in a LOO classification accuracy of 98.6%. Of the 41 subjects, 36 were classified as “dominant OA”, where the $B(OA)$ is larger than the sum of $B(NP)$ and $B(U)$. Of the 31 NP subjects, 26 were classified as dominant NP function.

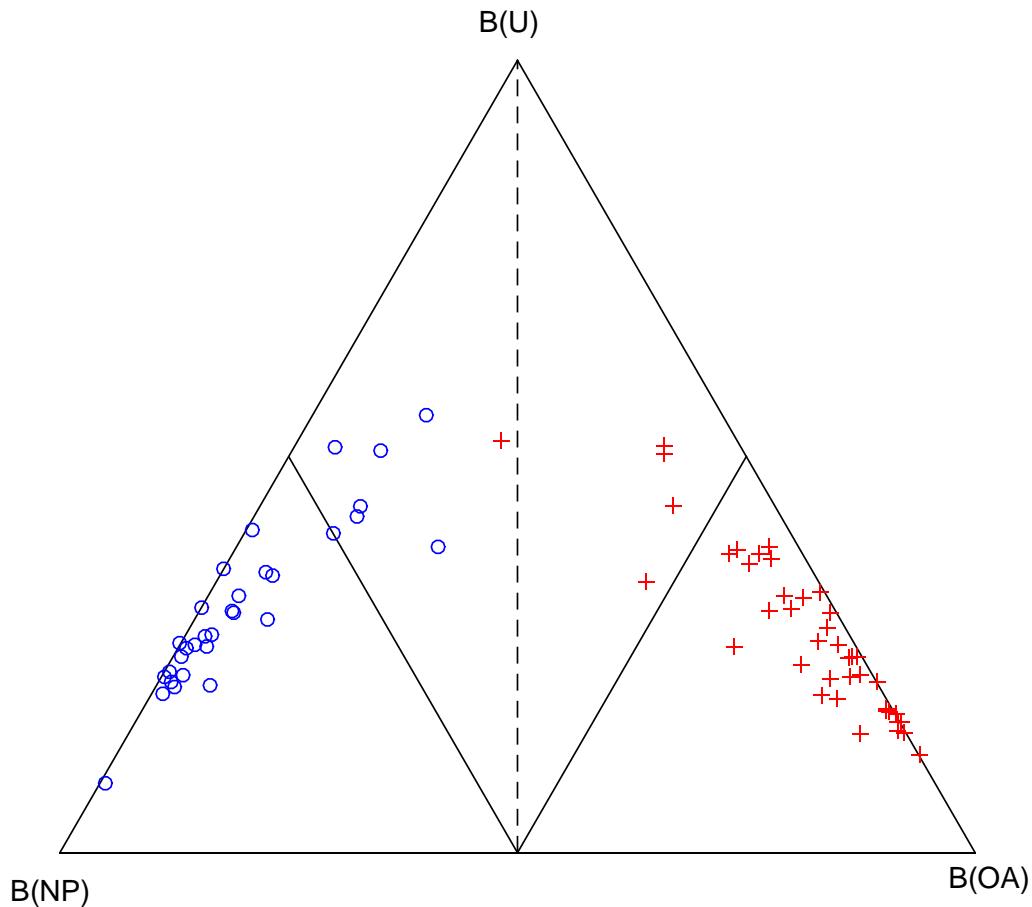


Figure 4.8 Simplex plot of the final classification, using the 18 chosen PCs, and the control definitions $k_{C/S}$, θ_{STD} , and upper and lower uncertainty boundaries of 1 and 0.8 respectively. The 41 OA subjects are plotted as a red cross, the 31 NP subjects are plotted with a blue circle.

Each input variable was ranked as to how well it could discriminate between OA and NP gait biomechanics. The results are shown in Table 4.3. Features extracted from the vertical and AP GRF waveforms appear to be the most accurate in discriminating OA. This is interesting considering this data is by far the least challenging to both collect and process. It can also be noted that several PCs marked with an asterisk would not have been retained for analysis had the factor loadings PCA retention rule been applied, however appear to be good discriminators of OA gait.

Table 4.3 The rankings of how accurately each variable was able to discriminate between OA and NP gait. Input variables marked with an asterisk * would not have been selected for further analysis had the factor loading retention rule been applied.

Ranking	Input variable	Classification accuracy (%)
1	GRF Vertical PC1	98.6
2	GRF Ant/posterior PC1	97.2
3	Knee Ad/Abduction Moment PC2	93.1
4	Hip Ad/Abduction Moment PC2	87.5
5	Hip Int/external Moment PC1	86.1
6	Hip Ad/Abduction Angle PC2*	86.1
7	Ankle Plantarflexion Moment P _c 2	84.7
8	Knee Flexion Moment PC2	84.7
9	Knee Flexion Angle PC2	84.7
10	Knee Int/external Mom PC1	83.3
11	GRF Medial PC2*	83.3
12	Hip Flex/extension Ang PC1	80.6
13	Ankle Inv/eversion Mom PC1	79.2
14	Hip Flex/extension Mom PC2	77.8
15	Knee Flex/extension Mom PC1	76.4
16	Hip Flex/extension Ang PC2*	76.4
17	Ankle Plant/dorsiflexion Ang PC2*	76.4
18	Hip Int/external Ang PC2*	76.4

Ranking 1: The first PC of the GRF was overwhelmingly the top-ranked variable in classifying between OA and NP subjects. The data reconstruction using this PC is shown in Figure 4.9. The PC is very similar to that defined in Section 3.6.2 however, the classification accuracy was higher at 98.6% vs 89.5%. The PC represented 67% of the total variable between subjects. The clinical description of this PC remains the same, therefore please refer to Section 3.6.2 for a full description. OA subjects appear to accept weight slower (relative to the duration of stance phase) than NP subjects, and they also unload slower. They also appear to have a reduced “double peak” characteristic, implying less vertical movement of COM.

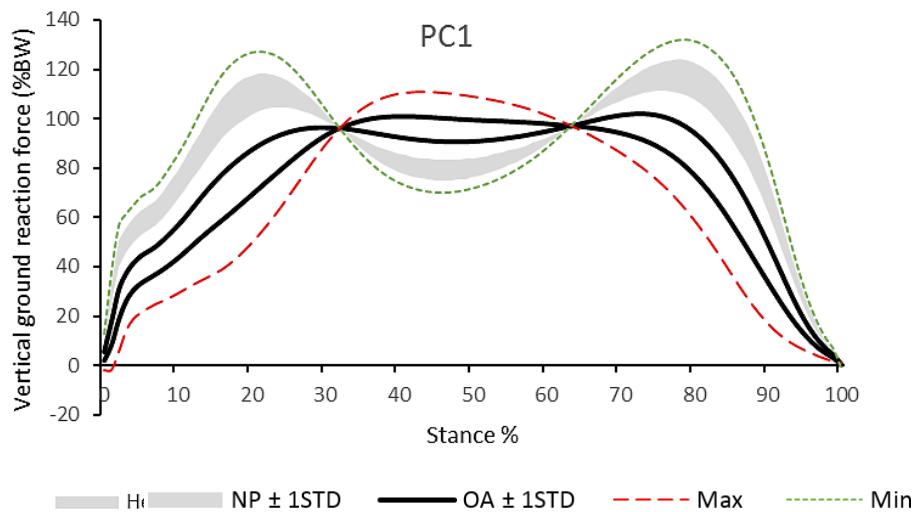


Figure 4.9 Reconstruction of the vertical ground reaction force using PC1. The $\pm 1\text{STD}$ confidence interval of NP subjects is shown as a grey shaded area for NP subjects, and as solid lines for OA subjects. Reconstruction using the maximum PC score is shown as a thin red long-dashed line, and the minimum as a thin green long-dashed line.

Ranking 2: The second ranked variable was the first PC of the anterior GRF, which represented 58% of the variable variance. The data reconstruction using this PC is shown in Figure 4.10. The PC is very similar to that defined in Section 3.6.2 and therefore the clinical description remains the same. As seen previously, OA subjects appear to have reduced sagittal plane GRFs throughout stance phase.

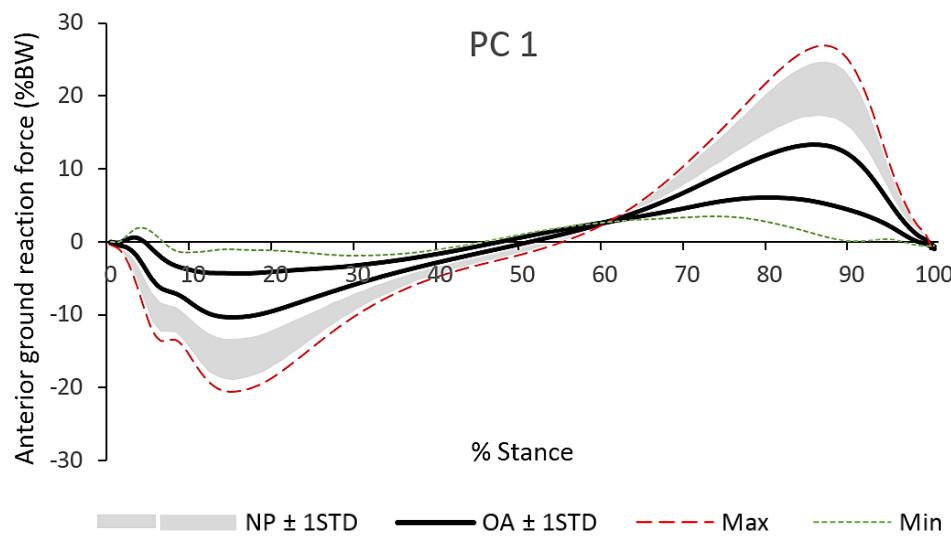


Figure 4.10 Reconstruction of anterior ground reaction force using PC1. The $\pm 1\text{STD}$ confidence interval of NP subjects is shown as a grey shaded area for NP subjects, and as solid lines for OA subjects. Reconstruction using the maximum PC score is shown as a thin red long-dashed line, and the minimum as a thin green long-dashed line.

Ranking 3: The third ranked variable was the second PC of the knee adduction moment, which represented just 13% of the variable variance. The first PC was also included in the classification, and was ranked 15th (see Figure 4.23). The data reconstruction using this PC is shown in Figure 4.11. The clinical description of this PC is slightly different to that described previously. Subjects with a high PC score have a much-prolonged larger adduction moment in midstance, a less of a bi-phasic loading pattern. Unlike within the

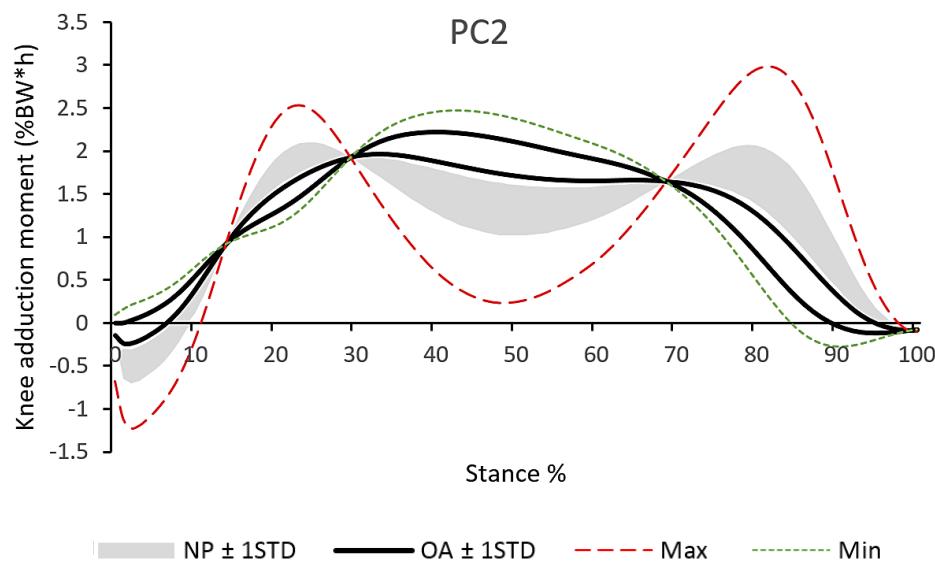


Figure 4.11 Reconstruction of knee adduction moment using PC2. The $\pm 1\text{STD}$ confidence interval of NP subjects is shown as a grey shaded area for NP subjects, and as solid lines for OA subjects. Reconstruction using the maximum PC score is shown as a thin red long-dashed line, and the minimum as a thin green long-dashed line.

previous definition, this PC doesn't reconstruct large magnitudes in the peak of adduction moment.

Ranking 4: The fourth ranked variable was the second PC of the hip adduction moment, which represented just 22.7% of the parameter variance. The data reconstruction using this PC is shown in Figure 4.12. The first PC wasn't included in the classification, as it was initially ranked 37th, and is therefore also graphed for reference. The clinical description of this PC is similar to that of PC2 of the knee adduction moment. Subjects with a high PC score have a much-prolonged larger adduction moment in midstance, and less of a biphasic loading pattern. Some of the biphasic/ "double peak" feature was also reconstructed within the first PC, which described large magnitude variations between subjects. It may be that these magnitude differences either weren't clinically significant when diagnosing OA, or perhaps they are due to a systematic error such as the coronal plane position of the HJC. It is known that this is more prone to error in overweight subjects due to soft tissue on the pelvis landmarks, as discussed in Section 4.2.3. This highlights an advantage PCA may have had in comparison to defining discrete parameters, such as the maximum magnitude at peaks, where the differences highlighted in PC2 would have been concealed within large magnitude differences.

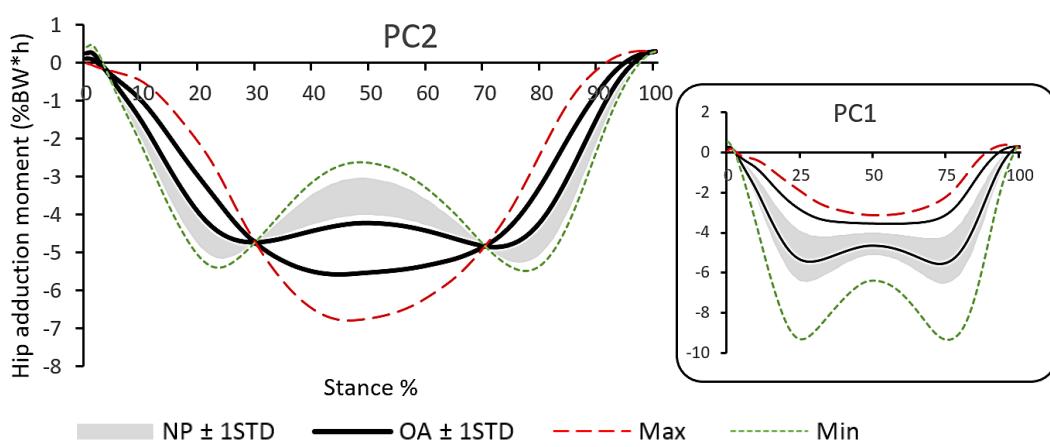


Figure 4.12 Reconstruction of hip adduction moment using PC2. The $\pm 1\text{STD}$ confidence interval of NP subjects is shown as a grey shaded area for NP subjects, and as solid lines for OA subjects. Reconstruction using the maximum PC score is shown as a thin red long-dashed line, and the minimum as a thin green long-dashed line. As the first PC wasn't retained, it is plotted in the square box for reference.

Ranking 5: The fifth ranked variable was the first PC of the internal hip moment, which represented 54.9% of the parameter variance. The data reconstruction using this PC is shown in Figure 4.13. It appears that a reduced magnitude of the internal hip moment throughout stance phase may be a good discriminator of osteoarthritic function. Subjects with OA commonly report pain when pivoting/twisting on the knee. It could be that OA subjects avoid transverse moments through decreased walking speeds or other gait adaptations. Reduced peak of the internal hip moment during stance has previously been found to be progressive with OA severity (Astephen *et al.*, 2008a); however, in this study, cadence was also different between the three groups of OA severity. Transverse plane moments appear to be less frequently reported in the literature; however, they have been found by multiple studies to be consistent between able-bodied subjects (Schache *et al.*, 2007). The choice of reference plane can have a significant effect on the profile of the transverse moment; NP moment profiles appear to be consistent with those reported by Schache *et al.* (2007) when resolving in the lab reference frame.

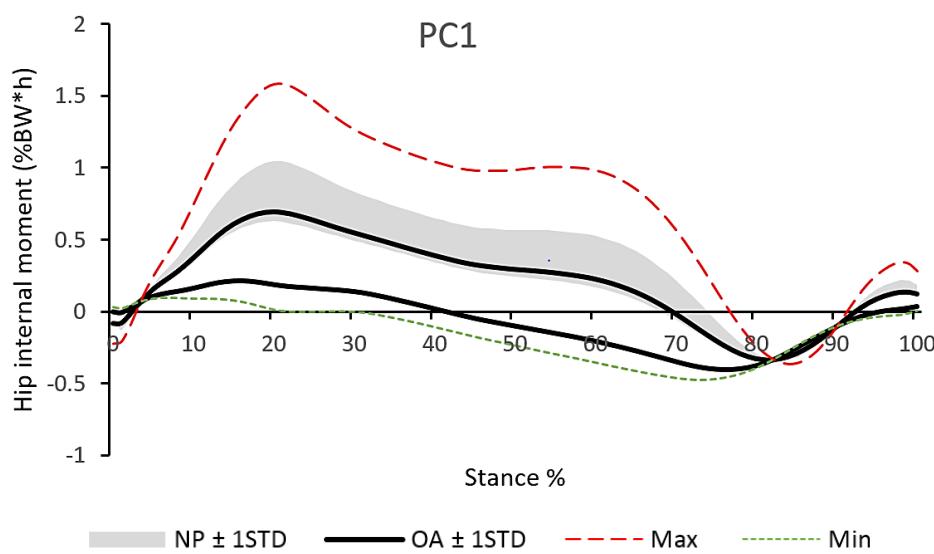


Figure 4.13 Reconstruction of the hip internal moment using PC1. The $\pm 1\text{STD}$ confidence interval of NP subjects is shown as a grey shaded area for NP subjects, and as solid lines for OA subjects. Reconstruction using the maximum PC score is shown as a thin red long-dashed line, and the minimum as a thin green long-dashed line.

Ranking 6: The sixth ranked variable was the second PC of the hip adduction angle, which represented just 11% of the hip adduction angle variance. The data reconstruction using this PC is shown in Figure 4.14. The first PC wasn't included in the classification, as it was initially ranked 45th, and is therefore also graphed for reference. Interestingly, had the previously used PC retention rule suggested by Comrey and Lee (2013), and adopted by Jones (2004), been used, only the first PC would have been selected. This highlights an important benefit from the adoption of a new “minimum of three” initial retention rule. The second PC of the hip adduction angle appears to represent changes in the coronal plane ROM of the hip throughout the gait cycle; however, towards the minimum value of the PC it appears to reconstruct a relative abduction during swing phase, as opposed to a relative adduction which would normally be seen.

PC appears to show higher and lower magnitudes of hip adduction angle throughout the gait cycle in OA subjects. If this was the case, and not due to errors in measurement, this would not result in a good classification accuracy using this variable, due to the use of a single sigmoid activation function per variable. These large magnitude offsets throughout stance phase could however be because of errors in the coronal plane location of the HJC or the KJC, particularly as the OA cohort had a much higher average BMI.

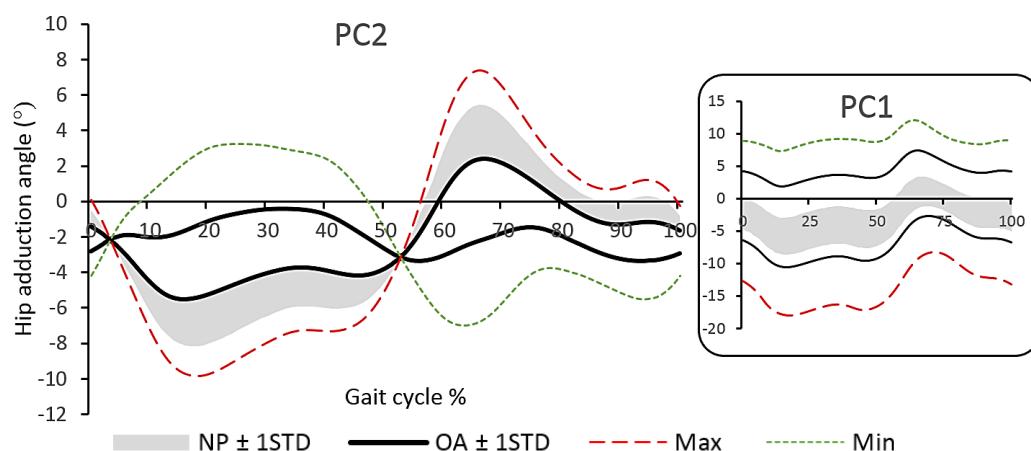


Figure 4.14 Reconstruction of hip adduction angle using PC2. The $\pm 1\text{STD}$ confidence interval of NP subjects is shown as a grey shaded area for NP subjects, and as solid lines for OA subjects. Reconstruction using the maximum PC score is shown as a thin red long-dashed line,

and the minimum as a thin green long-dashed line. As the first PC wasn't retained, it is plotted in the square box for reference.

Ranking 7: The seventh top-ranked variable was the second PC of the plantar-flexor moment, which accounted for 28.5% of the parameter variance. The data reconstruction using this PC is shown in Figure 4.15. The first PC wasn't retained for further analysis due to having an initial ranking of 31st, and is therefore illustrated alongside for reference. Both PCs appear to reconstruct magnitude difference in the plantarflexion moment which occurs during push-off; however, in the second PC, these differences appear to be related to a reduced dorsiflexion moment after heel strike, and an increased plantarflexion moment in the first half of stance. It might be that PC1 is more related to natural changes in gait velocity, and PC2 reconstructs a more clinically significant dorsiflexion moment avoidance, or perhaps a more anterior position of the COP or COM relative to the ankle during the first half of stance.

Anecdotally, the second PC wasn't correlated to gait velocity in NP subjects ($R=0.04$); however, the first PC was moderately correlated ($R=0.49$). This supports the suggestion that PC1 might represent more of a change in gait velocity than a presence of pathology; however, this would need to be analysed in a more controlled experiment.

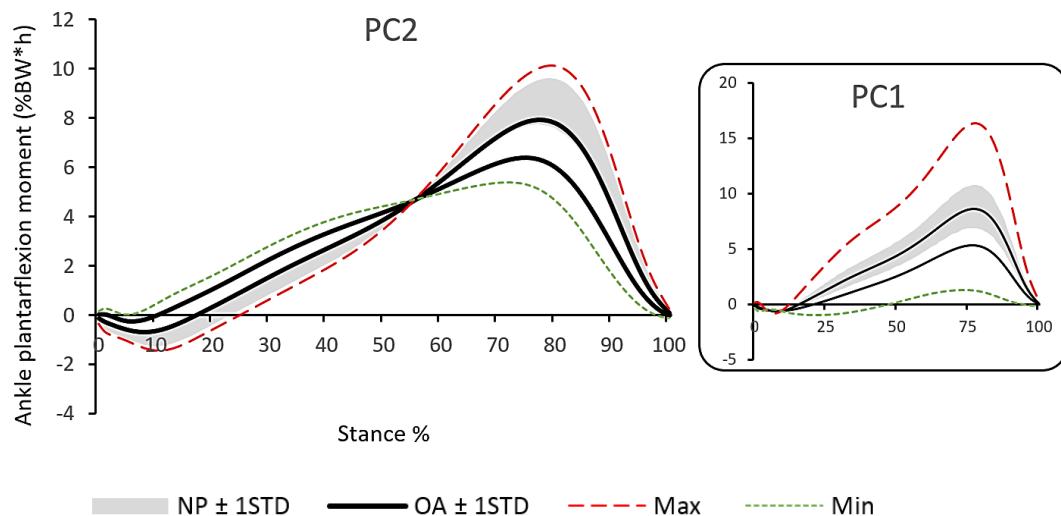


Figure 4.15 Reconstruction of the ankle plantar-flexor moment using PC2. The $\pm 1\text{STD}$ confidence interval of NP subjects is shown as a grey shaded area for NP subjects, and as solid lines for OA subjects. Reconstruction using the maximum PC score is shown as a thin red

long-dashed line, and the minimum as a thin green long-dashed line. As the first PC wasn't retained for analysis, it is plotted in the square box for reference.

Ranking 8: The eighth ranked variable is PC2 of the knee flexion moment, representing 30% of the parameter variance. The data reconstruction using this PC is shown in Figure 4.16. The first PC was ranked 15th, and is shown in Figure 4.23. While the first PC reconstructs large differences in the knee extension moment in the second half of stance, or at the extremes a complete avoidance of extension moment, the second PC reconstructs a reduction in both the knee flexion moment and, to a lesser extent, the knee extension moment during the second half of stance. Numerous studies have shown a strong relationship between sagittal moment peaks and gait velocity (Lelas *et al.*, 2003). While the high ranking of this PC could be a result of the differences in gait velocity between groups, it could be that the reduction in knee flexion and extension moment is a pain/instability avoidance strategy, as opposed to merely a consequence of reduced walking speed.

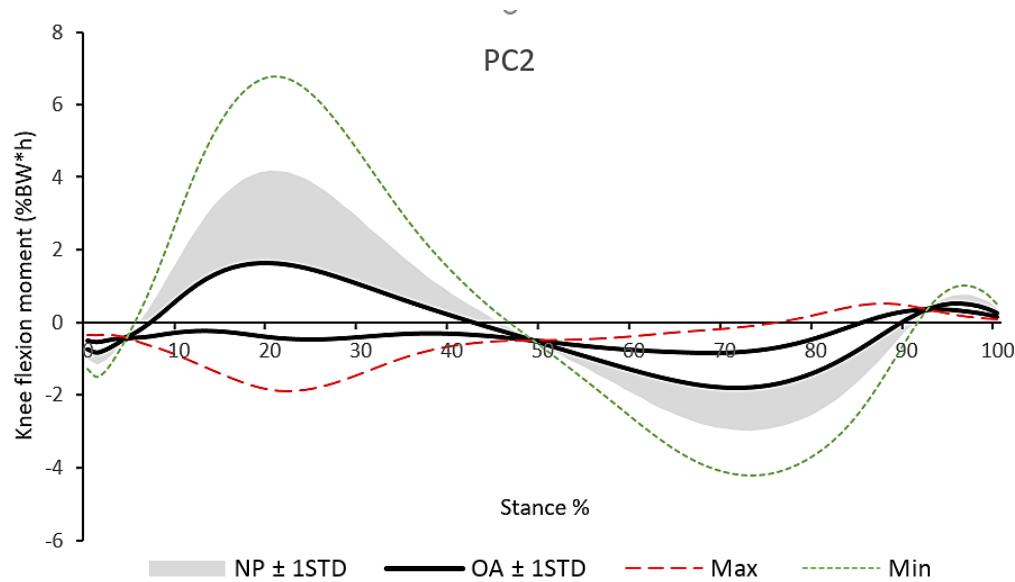


Figure 4.16 Reconstruction of the knee flexion moment using PC2. The $\pm 1\text{STD}$ confidence interval of NP subjects is shown as a grey shaded area for NP subjects, and as solid lines for OA subjects. Reconstruction using the maximum PC score is shown as a thin red long-dashed line, and the minimum as a thin green long-dashed line.

Ranking 9: The ninth ranked variable was the second PC of the knee flexion angle, which represented 24% of the parameter variance. The data reconstruction using this PC is shown in Figure 4.17. The first PC wasn't retained for further analysis due to having an initial ranking of 29th, and is therefore illustrated alongside for reference. The interpretation of the first two PCs is very similar to those defined for the flexion angle in Section 3.6.2, except that PC1 doesn't reconstruct changes towards peak swing.

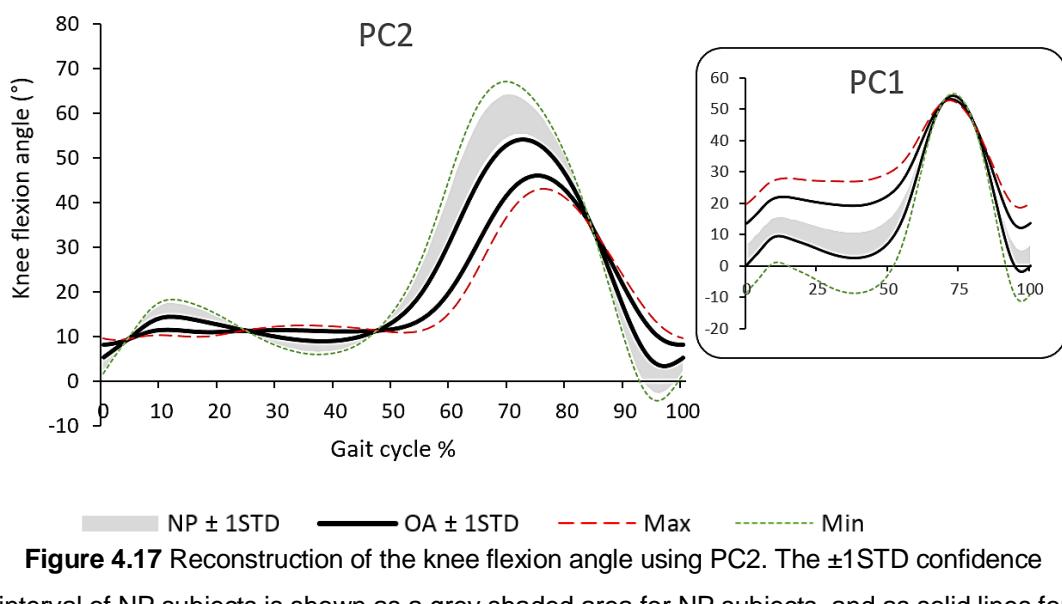


Figure 4.17 Reconstruction of the knee flexion angle using PC2. The $\pm 1\text{STD}$ confidence interval of NP subjects is shown as a grey shaded area for NP subjects, and as solid lines for OA subjects. Reconstruction using the maximum PC score is shown as a thin red long-dashed line, and the minimum as a thin green long-dashed line. As the first PC wasn't retained for analysis, it is plotted in the square box for reference.

Ranking 10: The 10th ranked variable was the first PC of the knee flexion angle, which represented 54% of the parameter variance. The data reconstruction using this PC is shown in Figure 4.18. The PC appears to largely reconstruct differences in magnitude of a peak internal or external moment during the second half of stance phase. The reconstruction of this PC suggests OA subjects have a reduced internal moment during the second half of stance, which has previously been reported in the literature (Ast Stephen *et al.*, 2008b). Changes in the internal knee joint moment in OA subjects has received little attention in the literature, and the few studies which do report changes acknowledge

the difficulty in relating these changes to the pathological mechanisms of the disease (Landry *et al.*, 2007, Astephen *et al.*, 2008b). It has however been suggested that changes in transverse plane biomechanics might change contact locations and stress distributions within the knee joint, and hence initiate degenerative changes through altered loading of articular cartilage (Andriacchi and Mündermann, 2006). As with many kinetic parameters, increased gait speed has previously been found to correlate with increased internal knee moments in both OA and NP subjects (Landry *et al.*, 2007).

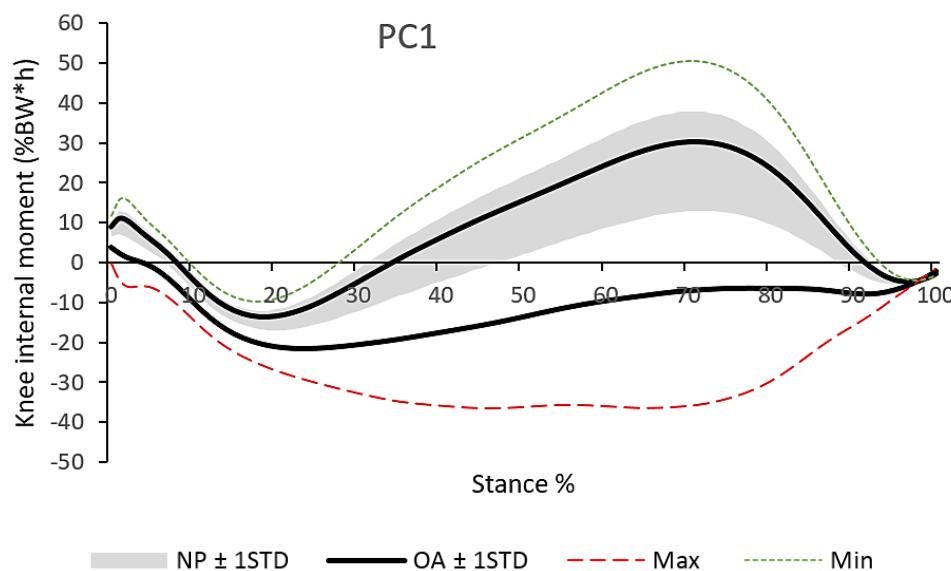


Figure 4.18 Reconstruction of the knee internal moment using PC1. The $\pm 1\text{STD}$ confidence interval of NP subjects is shown as a grey shaded area for NP subjects, and as solid lines for OA subjects. Reconstruction using the maximum PC score is shown as a thin red long-dashed line, and the minimum as a thin green long-dashed line.

Ranking 11: The 11th ranked variable was the second PC of the mediolateral GRF, which represented just 11% of the parameter variance. The data reconstruction using this PC is shown in Figure 4.19. The second PC wasn't retained for further analysis due to being initially ranked 19th, and is therefore also plotted for reference. The second PC appears to reconstruct the amount of "double peak" of the mediolateral force. The results suggest this characteristic of the ML force is more useful in distinguishing OA gait than the overall magnitude of the ML force. This supports the findings of Section 3.6.6 in which the

second PC representing the biphasic nature of the waveform was again found to be a much more valuable in discerning OA gait than the first PC.

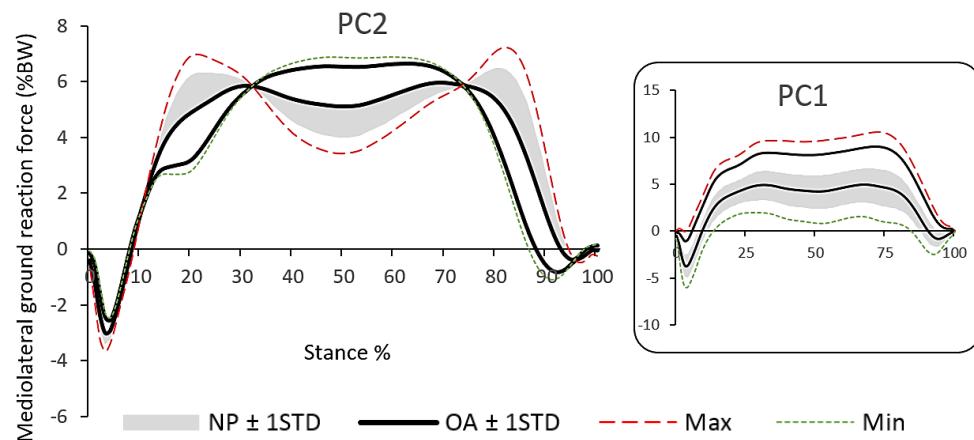


Figure 4.19 Reconstruction of the mediolateral GRF using PC2. The $\pm 1\text{STD}$ confidence interval of NP subjects is shown as a grey shaded area for NP subjects, and as solid lines for OA subjects. Reconstruction using the maximum PC score is shown as a thin red long-dashed line, and the minimum as a thin green long-dashed line. As the first PC wasn't retained for analysis, it is plotted in the square box for reference.

Ranking 12: The 12th ranked variable was the first PC of the hip flexion angle, which represented 90% of the parameter variance. The PC appears to reconstruct changes in magnitude of hip flexion during gait, as well as reduced ROM as hip flexion decreases. The standard deviation of hip flexion across OA subjects is much larger than the narrow band associated with NP subjects.

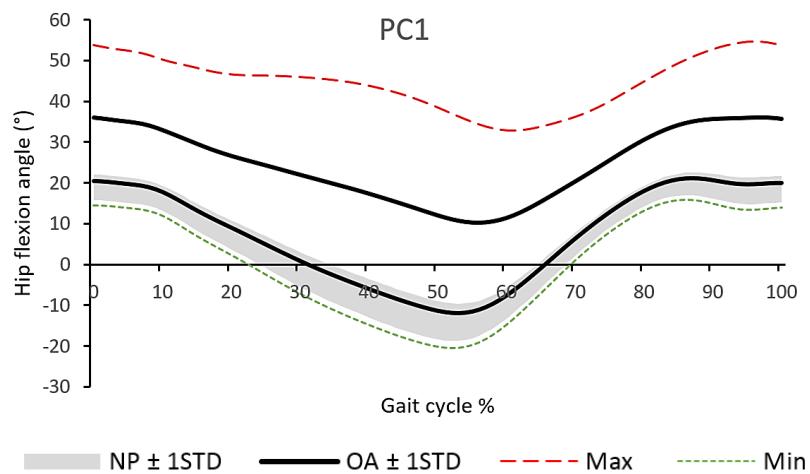


Figure 4.20 Reconstruction of the hip flexion angle using PC1. The $\pm 1\text{STD}$ confidence interval of NP subjects is shown as a grey shaded area for NP subjects, and as solid lines for OA

subjects. Reconstruction using the maximum PC score is shown as a thin red long-dashed line, and the minimum as a thin green long-dashed line.

Ranking 13: The 13th ranked variable was the first PC of the ankle internal moment, which represented 62% of the parameter variance. The data reconstruction using this PC is shown in Figure 4.21, and reconstructs changes in magnitude of the internal moment magnitude throughout gait. Much like at the knee, the internal ankle moment during the second half of stance also appears to be reduced in OA subjects. Again, it is difficult to relate these differences to the pathology. These differences might be related to changes in gait velocity, or perhaps changes in the foot progression angle as a learnt mechanism of reducing the knee adduction moment (Chang *et al.*, 2007).

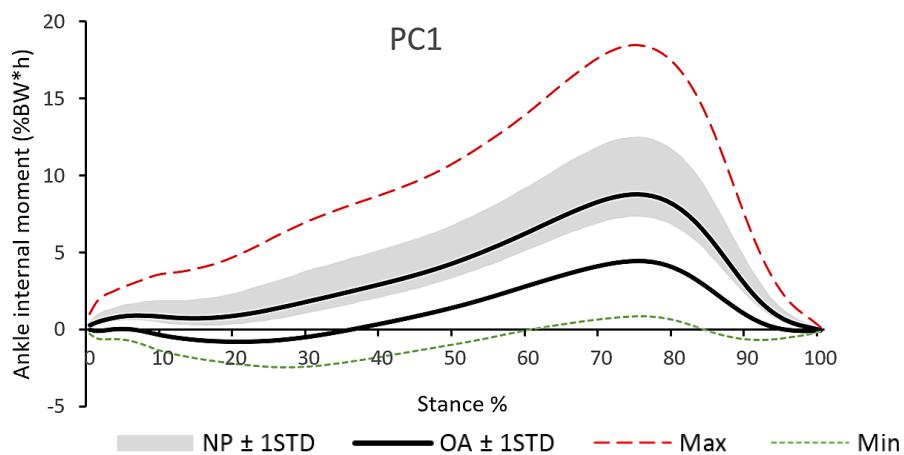


Figure 4.21 Reconstruction of the ankle internal moment using PC1. The $\pm 1\text{STD}$ confidence interval of NP subjects is shown as a grey shaded area for NP subjects, and as solid lines for OA subjects. Reconstruction using the maximum PC score is shown as a thin red long-dashed line, and the minimum as a thin green long-dashed line.

Ranking 14: The 14th ranked variable was the second PC of hip flexion moment, which represented 22% of the parameter variance. The data reconstruction using this PC is shown in Figure 4.22, alongside the first PC which wasn't included in the analysis due to an initial ranking of 32nd. The second PC appears to only reconstruct differences in the hip flexion moment during loading response and terminal stance. It appears that the

reduced hip flexion moments during these points are a better discriminator of OA function than the large magnitude differences throughout midstance reconstructed using PC1.

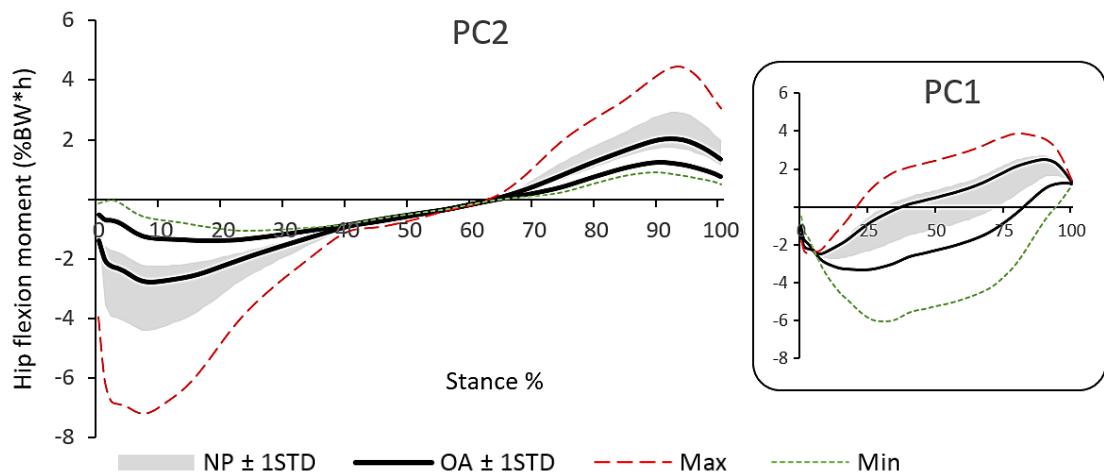


Figure 4.22 Reconstruction of the hip flexion moment using PC2. The $\pm 1\text{STD}$ confidence interval of NP subjects is shown as a grey shaded area for NP subjects, and as solid lines for OA subjects. Reconstruction using the maximum PC score is shown as a thin red long-dashed line, and the minimum as a thin green long-dashed line. As the first PC wasn't retained for analysis, it is plotted in the square box for reference.

Ranking 15: The 15th ranked variable was the first PC of the knee flexion moment, which represented 54% of the parameter variance. The data reconstruction using this PC is shown in Figure 4.23. The second PC of the knee flexion moment was ranked eighth and is shown in Figure 4.16. The first PC mainly reconstructs large differences in magnitude of the knee flexion moment throughout the second half of stance. It appears some subjects avoid an external extension moment altogether. This is likely related to fixed knee flexion during stance phase – which places the KJC more anteriorly relative to the COP of the GRF.

Ranking 16: The 16th ranked variable was the second PC of the hip flexion angle, which represented just 6% of the parameter variance. The data reconstruction using this PC is shown in Figure 4.24. The first PC of the hip flexion angle was ranked eighth and is shown in Figure 4.20. The first PC mainly reconstructs large differences in magnitude of the knee flexion moment second PC appears to reconstruct a reduced and delayed flexion angle towards toe-off.

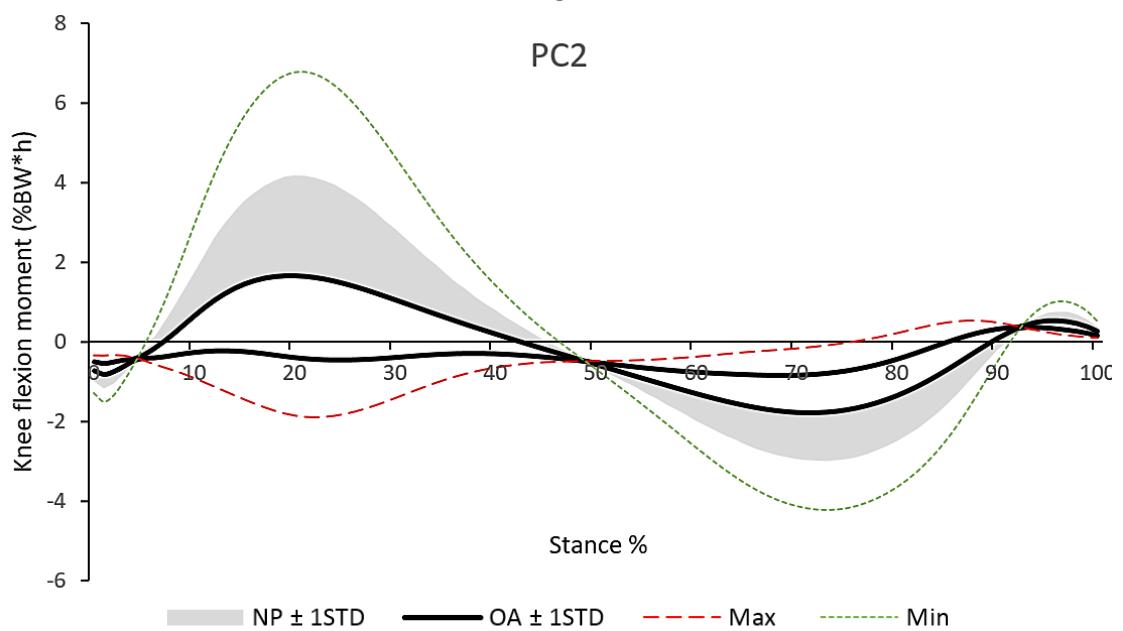


Figure 4.23 Reconstruction of the knee flexion moment using PC2. The $\pm 1\text{STD}$ confidence interval of NP subjects is shown as a grey shaded area for NP subjects, and as solid lines for OA subjects. Reconstruction using the maximum PC score is shown as a thin red long-dashed line, and the minimum as a thin green long-dashed line.

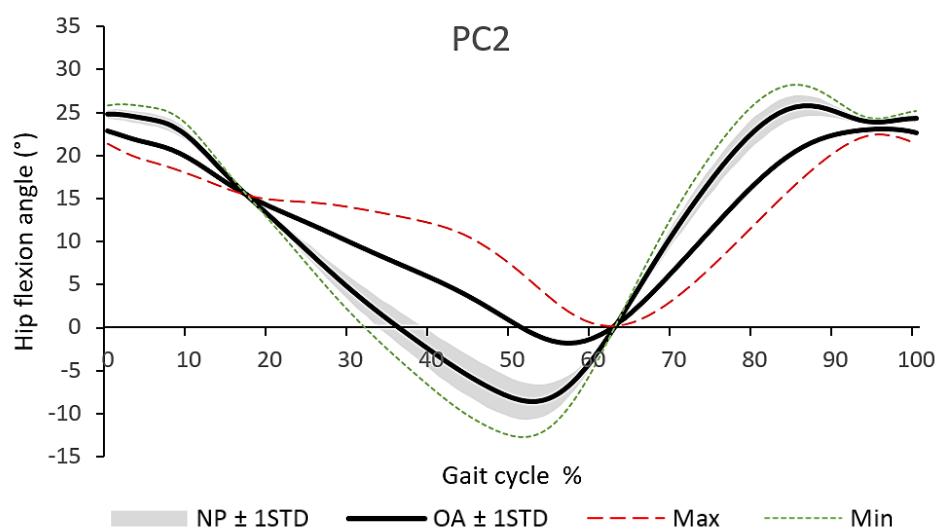


Figure 4.24 Reconstruction of the hip flexion angle using PC2. The $\pm 1\text{STD}$ confidence interval of NP subjects is shown as a grey shaded area for NP subjects, and as solid lines for OA subjects. Reconstruction using the maximum PC score is shown as a thin red long-dashed line, and the minimum as a thin green long-dashed line.

Ranking 17: The 17th ranked variable was the second PC of the ankle dorsiflexion angle, which represented just 13% of the parameter variance. The data reconstruction using this PC is shown in Figure 4.25. The first PC wasn't retained for further analysis due to having an initial ranking of 24th. The second PC reconstructs a delayed ankle dorsiflexion during stance phase, and a delayed and reduced plantarflexion during push-off and swing phase of gait.

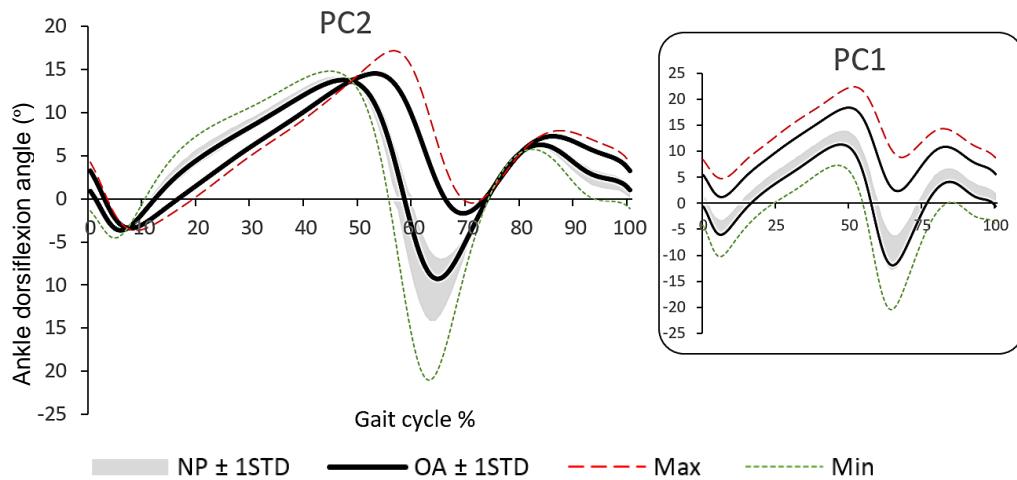


Figure 4.25 Reconstruction of the ankle plantar/dorsiflexion angle using PC2. The $\pm 1\text{STD}$ confidence interval of NP subjects is shown as a grey shaded area for NP subjects, and as solid lines for OA subjects. Reconstruction using the maximum PC score is shown as a thin red long-dashed line, and the minimum as a thin green long-dashed line. As the first PC wasn't retained for analysis, it is plotted in the square box for reference.

Ranking 18: The 18th ranked variable was the second PC of the hip internal rotation angle, which represented just 5% of the parameter variance. The data reconstruction using this PC is shown in Figure 4.26. The first PC wasn't retained for further analysis due to having an initial ranking of 29th. The second PC reconstructs a reduced transverse plane ROM in OA subjects, particularly a reduced tendency to internally rotate at heel strike and toe-off. This appears to be a better discriminator of OA gait than the large magnitude differences throughout the gait cycle reconstructed by PC1 – which may well be a result of errors in joint axis definition.

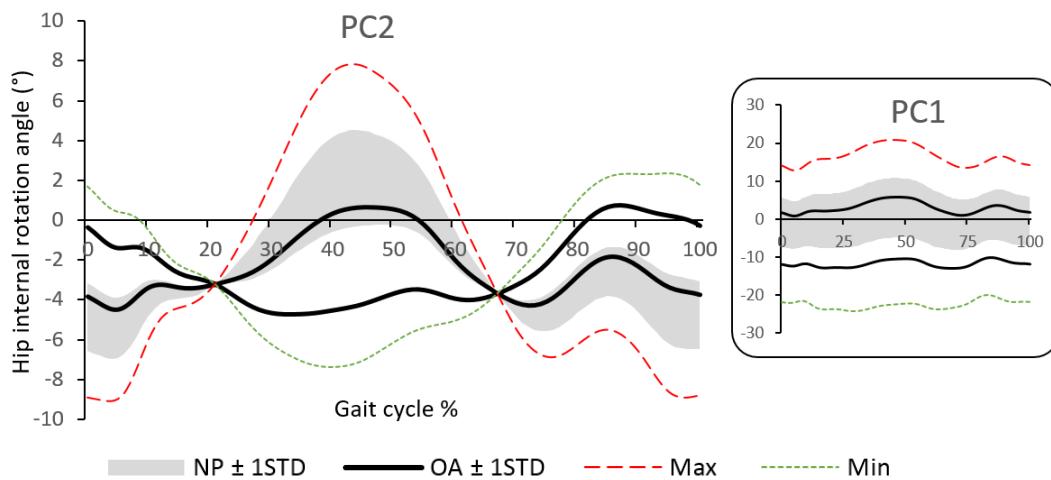


Figure 4.26 Reconstruction of the hip internal rotation angle using PC2. The $\pm 1\text{STD}$ confidence interval of NP subjects is shown as a grey shaded area for NP subjects, and as solid lines for OA subjects. Reconstruction using the maximum PC score is shown as a thin red long-dashed line, and the minimum as a thin green long-dashed line. As the first PC wasn't retained for analysis, it is plotted in the square box for reference.

4.3.5 About the Misclassified Subject

The misclassified subject had a TKR on the other leg three years prior to the motion analysis trial. The subject had no history of knee injuries, but has always actively participated in various sports. The participant didn't have any functional problems climbing normal household stairs, but did mention that it can cause pain. The trigger of pain was reported as a "twisting motion", or when standing or walking for long periods of time.

Within the Knee Outcome Survey, the participant commented that pain in the leg of interest affects their activity severely, and that grinding and stiffness affects their activity moderately. However, the participant also agrees with statements that the leg never feels weak, and never buckles or slips. Within the Oxford Knee Score, the subject also comments that they can easily walk down a flight of stairs, easily do household shopping on their own, and that pain in their knee only moderately affects their usual work (including housework). They also note that they can walk for more than 30 minutes without pain. They have a BMI of 34.5, which classifies them as obese. The subject notes they have trouble playing golf due to the twisting/pivoting. The subject had an

overall Oxford Knee Score of 70% for the left leg, and 96% for the right leg; where 100% represents no symptoms. Had the subject been classified using all the variables, they would have been classified as 59% OA, 32% NP and 10% uncertainty – the closest towards NP of all the OA subjects.

Unfortunately, it wasn't possible to retrieve the radiographic results of this subject, and the KL grade is therefore not known. Upon visual inspection of the motion analysis files, it was difficult to detect differences in gait to that of NP subjects.

The variables that most contributed to a NP classification were the PC2 of the knee adduction moment, second PC of the knee flexion moment, and the second PC of the anterior-posterior GRF. Variables that most contributed to the OA classification were overwhelmingly the 1st PC of the knee transverse moment, and the first PC of the ankle transverse moment. The decrease in transverse moments despite no decrease in sagittal moments might suggest a general avoidance of large internal moments, which have been reported by the subject to result in knee pain.

4.3.6 NP Subject with Lowest Belief in NP Function

Upon further inspection, the NP subject with the lowest belief of NP gait didn't meet the criteria as a non-pathological control subject stated in Section 3.2.1. The subject reported slight general knee pain, particularly during unexpected twists. The subject also reported stiffness in both knees, and pain in one leg while standing. Interestingly, this subject also had the highest level of classification uncertainty of the NP subjects ($B(U)= U 55.3\%$). Unfortunately, this was only discovered after the study, and re-running the analysis with this subject excluded would have been time-consuming. Considering the number of subjects within the training body and strong performance of the DST classifier it is unlikely that the re-analysis would have a significant impact upon results.

4.3.7 Assessing the Validity of a Combined Healthy and Elderly Cohort in Classifying OA Subjects

A comparison of the relationship between participant age and the belief values B(NP) and B(OA) are plotted within Figure 4.27A and B respectively. There is no obvious relationship between either belief value and the age of the NP participant, however statistical analysis was continued for completeness. The B(OA), B(NP) and the age of participants were all tested for normality, and the age of participants was not normally distributed. This was, perhaps not surprising – many of the younger subjects recruited were University students and hence tend to be aged 18-26. A Spearman's rank correlation test was performed. No significant correlation was observed found between age and the belief of OA ($r=0.046$, $p=0.811$) or NP ($r=-0.146$, $p=0.44$) for the non-

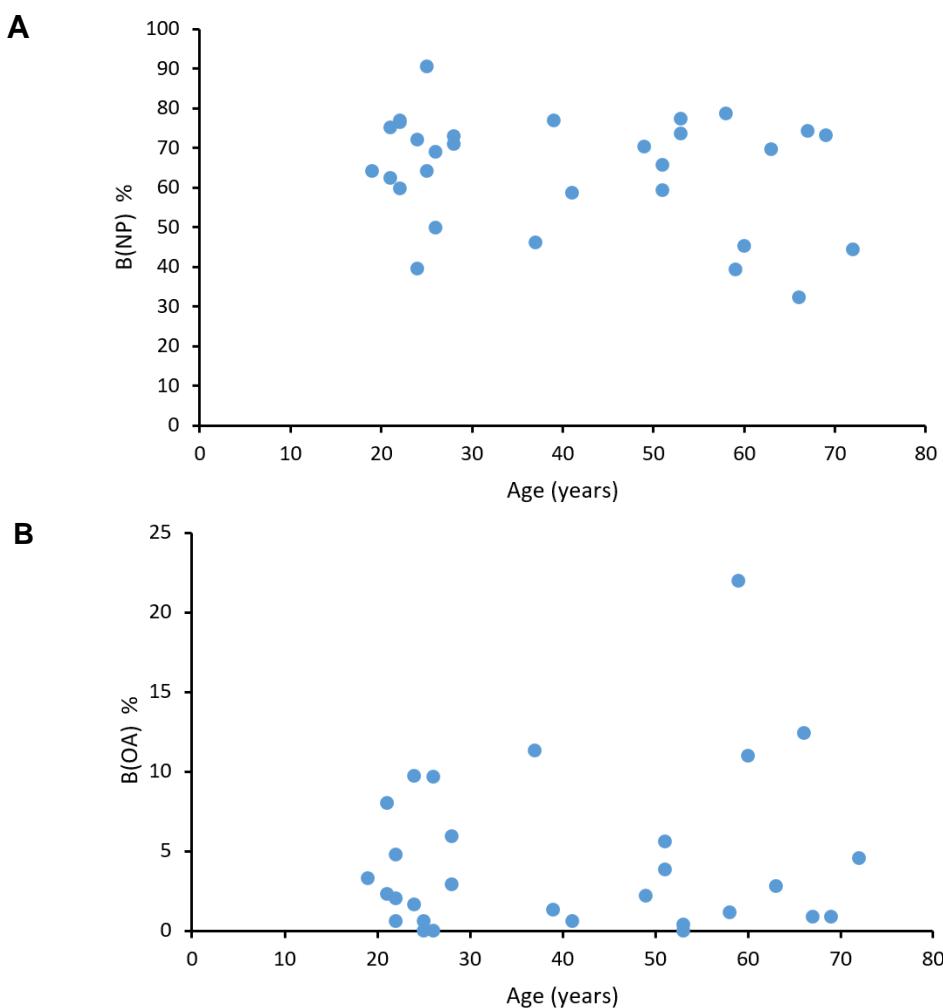


Figure 4.27 The relationship between participant age, in years, and

- A) % belief of OA, B(OA),
- B) % belief in NP, B(NP).

pathological control cohort. It could be argued that the effects of ageing might only result in significant changes past a threshold age, therefore reducing the strength of any correlations when the group is treated as a whole. Data shown within Figure 4.27 seems to suggest that this is not the case, however, this would need to be explored with a larger dataset with equal subject numbers of subjects within specified age groups so that this could be further explored.

It is also worth noting that the oldest NP control subject within this study was 72 years. Of the 41 OA subjects, 14 were older than 72, and the oldest was 84 years. It therefore isn't possible to deduce whether the classification might have been sensitive to age-related changes, had age-matched controls been used. In practice, it would be very difficult to recruit NP subjects within this age category who match the inclusion criteria within the study.

4.4 Conclusions

Aim 1: Assess the appropriateness of previously defined PC in representing variance between subjects collected with the updated methodology.

It was hypothesised that the changes to the methodology of calculating knee joint kinematics and kinetics would reduce the appropriateness of the previously defined PCs in representing biomechanical features. An analysis of 9OA and 9NP subjects calculated showed significant differences in several kinematic and kinetic measures by considering the linear correlation between PC scores calculated using the different methods, alongside the differences in means. It was hypothesised that there wouldn't be any changes in the resultant PC scores from the GRF data, however there were still some differences within the AP and ML ground reaction forces. It was therefore decided to define and contextualise new PC vectors within this chapter.

Aim 2: Assess the validity of a combined young, middle-aged and elderly cohort in classifying OA subjects.

To assess biomechanical changes relating to OA it is necessary to compare OA biomechanics to that of NP subjects. The validity of using a combined young, middle-aged and elderly cohort to define NP function. It was hoped that having a heterogeneous NP cohort would train the classifier to be less sensitive to biomechanical changes which might be related to ageing as opposed to specifically OA. This was confirmed by testing for linear correlations between B(OA), B(NP) and age within the NP cohort: no relationship appeared to exist between belief of osteoarthritic function and age of the NP subject.

The final classification of OA function within this chapter accurately distinguishes between OA gait biomechanics using only 18 biomechanical variables. These top-ranked variables were chosen in a way which was shown to reduce a positive bias in the resultant classification accuracy. One OA subject was misclassified, however upon

further review it appeared this person seemed to have a surprisingly high level of function for a TKR subject. It is possible that this person is a high-performing outlier within the dataset. This highlights difficulty of classifying the function of such a heterogenous cohort.

The objective classification of osteoarthritic gait biomechanics presented in this chapter forms the foundations in which pre and post-operative function shall be objectively quantified within the following chapter.

4.5 Clinical Summary

The application of PCA has again proved accurate at objectively describing differences in biomechanical gait parameters. When performing PCA, an ‘eigenvector’ is calculated which describes a prevalent feature of variance within the data. For each subject, a principal component value is then calculated for each subject. This method has been shown sensitive to changes in methodology – which was quantified by significant changes in PC values for subjects processed using two different techniques. PCs defined in Chapter 4 might therefore no longer be valid for objectively discerning between subjects within this chapter.

There are a whole host of challenges in comparing biomechanical information from subjects collected within different laboratories or processed using different approaches. Significant differences are likely to exist between datasets due to a plethora of factors, including biomechanical model definitions, hardware, and expertise in palpating anatomical landmarks. The same considerations should be considered when adopting a data reduction technique which has been defined or modelled on data which may have been collected or processed differently. Another example of this issue in literature is the findings of McMulin and MacWilliams (2008), who found the Gillette Gait Index (introduced in Section 2.5.1) varied as much as 20% between multiple sites using the index.

The biomechanical features of OA presented within this chapter define the parameters for quantifying biomechanical function before and after TKR surgery. Interpretation of the biomechanical features can be found within Section 4.3.4 and are further discussed in Section 6.2. Some key findings are:

1. **The magnitude of the double peak of the adduction moment discriminates severe knee OA better than the overall magnitude of the waveform.** Many studies consider the magnitude of the EKAM peaks, however this discrete measure would include the effect of large magnitude differences in adduction moment peaks

throughout the gait cycle. A more discrete measure might be the value of the peaks of the adduction moment, as a percentage of the trough between the two peaks. This discrete metric might be a useful addition or alternative to solely considering EKAM peaks.

2. **Hip ad/abduction can discriminate severe OA gait, but may be difficult to discriminate using traditional analytical techniques** - The hip adduction angle was able to discriminate OA gait with 86.1% accuracy (Table 4.3), and the feature detected is shown later in Figure 5.7. In short, the feature appears to show 'hip hiking'; perhaps a compensation strategy increase ground clearance and account for decreased knee flexion during swing phase. This relevant biomechanical feature, however, only represented 11% of the total variance between subjects. The vast majority of variance was accounted for in a feature which reconstructed large magnitude offsets throughout the gait cycle (see "PC1" within Figure 4.14). This is likely due to the difficulty in defining the coronal plane axis of both the hip and the pelvis. It may be of interest to calculate hip ad/abduction relative to the position at heel strike. Previous studies have, however, reported increased hip abduction throughout stance in severe OA subjects (Hunt et al., 2010), suggesting this feature is still detectable using traditional methods. The potentially enhanced ability of PCA to detect this feature over traditional methods warrants further investigation.
3. **Transverse moments were again shown to discriminate severe OA gait, despite receiving less attention in the literature.** It was discussed in the clinical summary of the previous chapter that the transverse (internal/external) knee moment was surprisingly useful in distinguishing the gait of severe OA subjects. The transverse moment of the hip, knee and ankle were all as highly accurate in classifying OA gait.

Chapter 5 - Quantifying Functional Changes Following Total Knee Replacement Surgery

5.1 Introduction

Total Knee Replacement surgery is a common procedure to treat moderate to late-stage OA of the knee. Alongside the increased incidence of TKR surgery, introduced in Section 0, there is also a rise in the treatment given to a relatively younger patient population, made possible by advancements in the durability of the endoprosthesis (Mizner *et al.*, 2005). The 11th report of the NJR analysed the risk of revision up to 10 years following surgery for a number of age groups (Figure 5.1), which highlighted the increased incidence of revision within younger cohorts (UK-NJR, 2014). The 12th report comments on the primary reason for knee revision surgery. It can be calculated, using the data presented within the report, that the primary reasons for revision in a cohort of 47,829 subjects were aseptic loosening (34%), infection (22.3%), pain (16.1%), instability (14.7%), implant wear (12.4%), lysis (10.1%), and malalignment (6.81%).

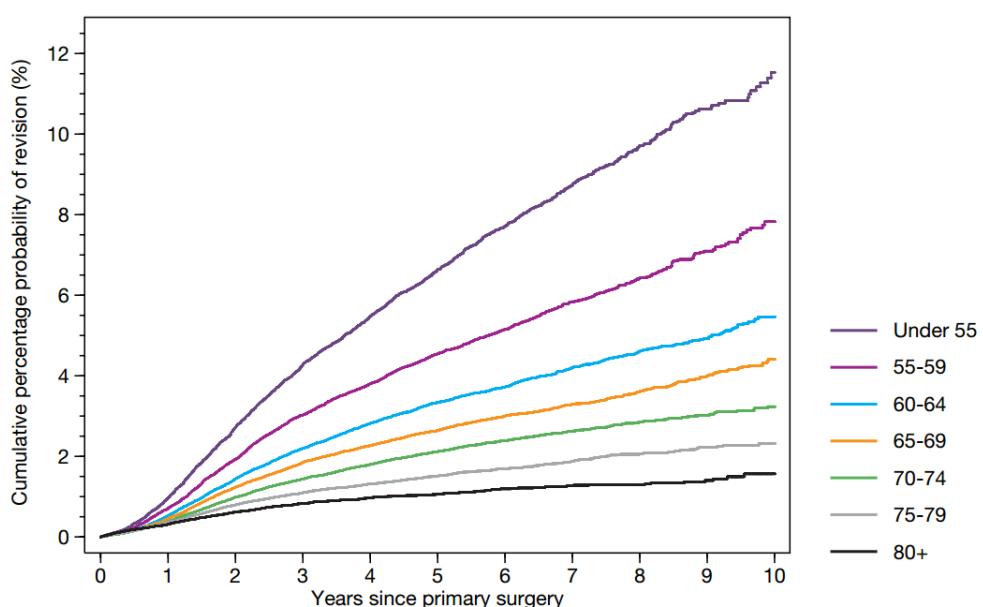


Figure 5.1 A comparison of the cumulative percentage probability of revision following primary TKR surgery for different age groups. Reprinted from UK-NJR (2014)

Despite improvements of outcomes following TKR surgery, patient satisfaction remains relatively low at around 80%, in comparison to 92% for hip replacement surgery (Baker *et al.*, 2007) (Anakwe *et al.*, 2011). This has led to increased scrutiny of healthcare providers to objectively assess outcomes of TKR surgery (Bourne, 2008). Patient-Reported Outcome Measures (PROMs) were introduced by the NHS in England in 2009, and became a first of its kind internationally in allowing patient perspectives to influence treatment (Partridge *et al.*, 2016). The EQ-5D-3L and the OKS were selected as the PROMs for monitoring outcomes of TKR surgery. The EQ-5D-3L is a widely used PROM for measuring general health status and is used to monitor health status in a wide range of populations. The OKS was designed specifically to assess the patient's perspective of pre-operative knee joint status and outcome following TKR surgery. It was initially designed to help eliminate surgeon bias in the selection of TKR surgery, and has since evolved to be the gold standard in measuring TKR outcome (Partridge *et al.*, 2016).

The overarching purpose of TKR surgery is to improve both current and future quality of life for the patient. This is expected to be achieved through two primary routes – a reduction in pain and an improvement in function. Functional improvement is by its very nature complex with multi-dimensional elements, and it is difficult to generalise which elements of physical functional are most important to any individual patient. PROMs therefore aim to assess physical function during a broad range of ADLs, as well as specific elements of physical function which might affect ability to perform a wide range of sports e.g. squatting, pivoting, jogging.

A challenging element of measuring physical function is that they measure the subject's perception of their own physical function. There are several influencing psychological factors which might affect how much a patient might over or under-estimate their own physical function, such as depression, pre-surgery expectations, dissatisfaction following surgery, the response-shift phenomenon (Razmjou *et al.*, 2009). There is growing evidence that performance-based measures capture a different element of physical

function of OA subjects than that of perception-based PROMs (Mizner *et al.*, 2011, Stratford and Kennedy, 2006).

The methods described in this thesis thus far represent an objective performance-based measure of functional changes in OA subjects during level gait. While the measure has yet to undergo the same reliability and repeatability assessment that would be required of an outcome measure in order to assess eligibility for TKR surgery and outcome following surgery, it may have the ability to provide a great deal of insight into changes in physical function following surgery. By comparing the change in B(OA) pre and post-surgery, using the trained classifier which reliably discriminates between OA and NP gait biomechanics, it is possible to measure how biomechanical alterations specific to OA are modified by TKR surgery. This methodology has previously been adopted within the PhD research of Jones (2004), Whatling (2009) and (Metcalfe, 2014).

These studies have used the same DST classification techniques to quantify TKR outcome; however, the data inputted into the classifier has varied depending on study methodology. The methods described thus far in this thesis result in a novel classification of osteoarthritic gait biomechanics, and are the resultant of proposed changes to the control variables of the DST classifier (Section 3.5), changes to the methodology of the calculation of kinematics and kinetics (Sections 4.2.2, 4.2.3 & 4.2.5) and changes to the retention and interpretation of PCs (Sections 4.2.8 & 4.2.9. The quantification of the net effect all these changes have on the sensitivity and specificity of the biomechanical classification of OA subjects would be very interesting, but is beyond the scope of this thesis.

Instead, one of the aims of this chapter is to analyse how the objective classification of gait biomechanics compared to PROMs before and after TKR surgery.

Aim 1: Do pre-, post-, and the relative change in subjective outcome measures correlate with changes in biomechanical gait classification?.

Hypothesis 1: Subjective outcome measures will correlate with the functional abnormality before surgery, as biomechanical function declines with pain. Several studies have found that the patient's own perception of functional improvement is often greater than objectively measured functional changes (Mizner *et al.*, 2011, Worsley, 2011) It is therefore hypothesised that PROMs will correlate less strongly with objectively measured function post-operatively.

Predicting Post-operative Functional Outcomes

As a result of the aforementioned adoption of OKS and EQ-5D questionnaires within England and Wales, Baker *et al.* (2012) were able to assess surgical factors on early PROMs in 22,691 primary TKRs. The study used stepwise multiple linear regression to identify factors most predictive of improvement in OKS and EQ-5D score. One of the findings was that subjects with a higher (less pathological) pre-operative OKS score are likely to gain benefit from surgery, as measured by the increase in OKS score. This is perhaps not surprising; the worse the condition of the pre-operative knee, the larger the room for improvement. Previous researchers using the objective classification of biomechanical function have, however, identified that subjects with particularly pathological gait biomechanics before surgery appear to have less improved biomechanical function following surgery (Metcalfe *et al.*, 2013) (Watling, 2014) (Worsley, 2011).

Scott *et al.* (2010) assessed predictors of dissatisfaction six months following TKR in a cohort of 1217 patients between 2006 and 2008. The amount of improvement in the OKS and SF-12 questionnaires were found to be correlated to the likelihood of post-operative satisfaction. The principal pre-operative predictors of outcome were depression, low SF-12 mental component score, pain in other joints, and low SF-12 and OKS scores at baseline (pre-operatively). The total length of time in hospital was also found to be predictive, although less statistically significant, and was regarded as a measure of post-operative complications. Patient satisfaction was measured at one year following surgery, with a rate of 81.4%.

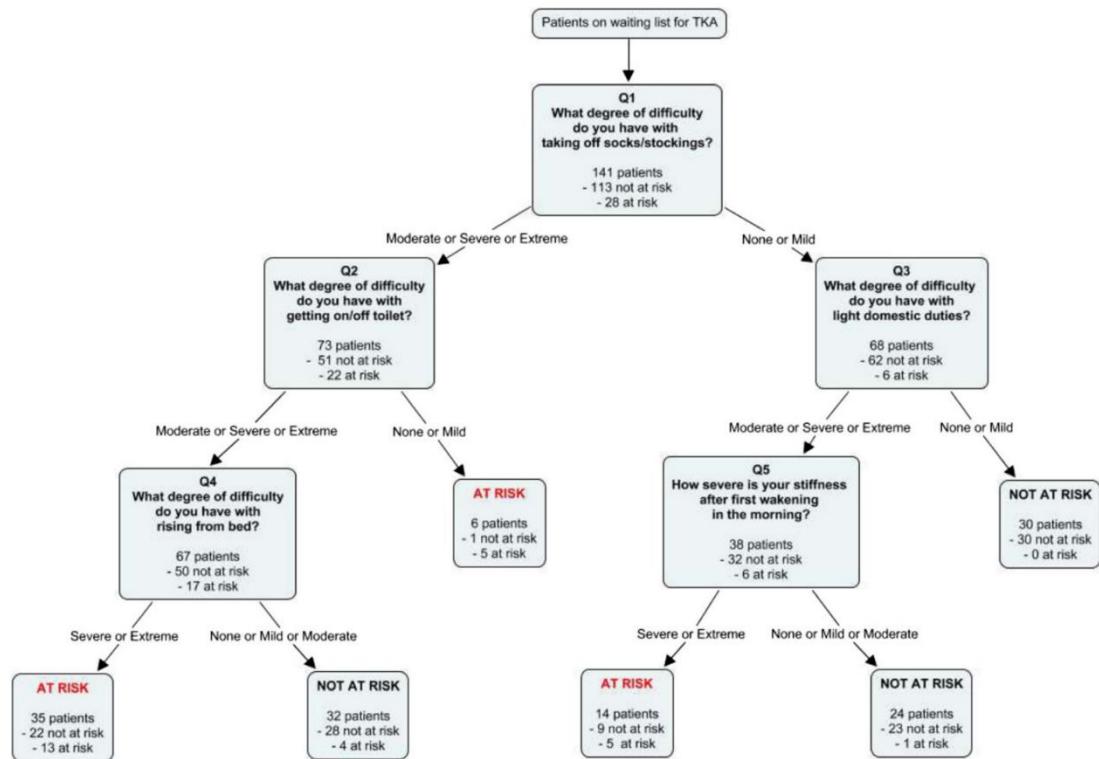


Figure 5.2 The final prediction model of Lungu *et al.* (2014), designed to pre-operatively identify patients at risk of poor outcomes following TKR.

Lungu *et al.* (2014) assessed predictions of poor outcomes six months following TKR surgery in 141 patients. The patients with the highest (more severe) 20% of post-operative WOMAC scores were defined as poor improvers, resulting in the 28 subjects with a post-operative WOMAC higher than 40.4% being selected. The authors then used this as training data to build a predictive model of good and poor TKR outcomes based on pre-operative data. The study applied recursive partitioning analysis to build a decision tree consisting of only five questions, which identified at-risk subjects with a positive predictive value of 41.8% and a negative predictive value of 94.2% (see Figure 5.2). Elaborating further, given an assumed prevalence of 20% of “poor post-operative outcomes” within the general population, it is predicted that of those subjects who are classified as at “risk”, there is only around 42% who will have poor outcome. Of the subjects classified as “not at risk”, only around 6% will have poor outcome.

Watling (2014) used a training body of 25 severe OA (pre-operative TKR) patients and 23 NP control subjects to classify between OA and NP gait biomechanics using the DST classifier. Of those 25 OA subjects, the gait biomechanics of 12 subjects were again

assessed approximately 12-months post-surgery. The pre and post-operative classification of gait function was used to quantify and compare biomechanical knee function before and after surgery. The research found that pre-operative age and biomechanical function were the best indicators of functional gait improvements following surgery, and that implant type, BMI and presence of comorbidities were poor indicators.

Worsley (2011) assessed the outcomes of 31 patients going for knee arthroplasty (15 UNI, 16 TKR), measuring changes in both subjective and objectively measured function. Objective changes in function were again measured using the Cardiff DST classification method by training the classifier on a cohort of the NP and pre-operative. Due to the uneven cohort size, the value of θ within the classification will have been biased towards the pathological group, which may have moderately adversely affected the classification. The study found a disparity between subjective and objective measures and therefore adds to the BOE in support of objective functional measures (Mizner *et al.*, 2011). This research is an extract of a much more in depth analysis published within a PhD thesis (Worsley, 2011), within which predictors of improvement in function were established. The principal predictors of functional improvement were the objective measures (muscle size, ROM, PC scores from biomechanics during various ADLS), as opposed to the subjective measures (WOMAC, OKS, VAS Pain).

Aim 2: Does functional recovery return following TKR?

Hypothesis 2: Recovery of biomechanical gait function will be variable across subjects. It is hypothesised that there may be two distinct groups of improvement, as found in preliminary work discussed within Section 5.2.

Aim 3: Does severity of pre-operative gait abnormality predict improvement in gait following surgery? Are there other biomechanical clinical factors which might predict improvement?

Hypothesis 3: Subjects with the most severe gait abnormalities pre-operatively, i.e. classifier as the most pathological within the DST classifier, will also have the worst post-operative biomechanical function.

5.2 Preliminary Work

Preliminary work was carried out to assess potential predictors of TKR outcome using data obtained within the research of Dr. Daniel Watling (Watling, 2014). This research study collected lower limb gait biomechanics from 25 OA (pre-TKR), 23 NP, and 12 post-TKR subjects. Subject demographics are shown in Table 5.1.

Table 5.1 Subject metrics for the healthy volunteer (NP), pre-operative (OA), and post-operative (TKR), total knee replacement subjects used for the classification of knee function (Watling, 2014).

Group	Number	Male	Female	Age/ years (SD)	Height/ m (SD)	Mass/ kg (SD)	BMI (SD)
NL	23	7	16	33.5 (11.9)	1.69 (0.07)	69.41 (12.12)	24.22 (3.49)
OA	25	11	14	69.0 (7.3)	1.68 (0.11)	93.25 (21.30)	32.81 (7.00)
TKR	12	5	7	69.1 (6.7)	1.65 (0.11)	94.58 (24.64)	34.04 (7.86)

The research methods for the collection of level gait biomechanics were identical to those used in this study, as they form part of the same longitudinal data collection. Of the 12 post-TKR subjects used within the work of Whatling, ten were also used within this PhD thesis. Two were excluded because their post-operative visits were less than ten months following surgery. This thesis also includes an additional 12 post-operative subjects which had been collected since Watling concluded his research. The research methods for the processing of biomechanical data is, however, slightly different; for further details the reader is directed to Watling (2014).

Watling included GRF, sagittal plane angles, and sagittal and coronal plane moments for the hip knee and ankle within the classification of OA and NP knee function. Coronal and transverse plane angles and transverse plane moments were excluded from the study.

In total, 27 PCs were included in the classification of OA gait function. The correlation coefficient was used as the definition for k , and the average variable value between both groups as the definition for θ . The resultant LOO classification accuracy was 97.9%.

The trained classifier was then used to classify the gait biomechanics of 12 of the subjects following surgery. The resultant classification is shown in Figure 5.3. Seven of the subjects significantly regained biomechanical function following surgery, while five appeared to have recovered very little.

It was then decided to respectively analyse whether there were any potential pre-operative predictors of post-operative recovery of biomechanical gait function. Patient demographics, temporal-spatial parameters, and biomechanical parameters taken pre-operatively were considered in the analysis. To limit the number of variables within the analysis, discrete parameters were taken from only the GRF and the sagittal knee angle. It was deemed that these were the most easily implementable within clinic, for example, using a single force plate and a high-speed digital camera. It is worth noting that the same level of accuracy may not be achievable in practice, particularly in the measurement of sagittal knee angles.

After much deliberation, it was decided that the percentage reduction in B(OA) was used, as opposed to the total reduction of the B(OA). i.e.

$$\text{Reduction} = \frac{B(\text{OA})_{\text{post}} - B(\text{OA})_{\text{pre}}}{B(\text{OA})_{\text{pre}}}$$

It was decided at the time that the percentage reduction in B(OA) was the most clinically relevant way of expressing the improvement in biomechanical function. This is discussed further in Section 5.4.2.

The results of the analysis are displayed in Table 5.2. There were four potential predictors which had a significance level $p < 0.05$, however, given that 20 variables were considered in the analysis and that the cohort size was very low, a great deal of caution is advised. Only one variable, cadence, reached the higher significance threshold of

$p<0.01$, hence an estimated change of less than one in 100 of being a false positive finding.

Interestingly, reduced AP force, reduced peak knee flexion and increased percentage of stance time might all potentially be explained by reduced gait velocity (Nilsson and Thorstensson, 1989, Oberg *et al.*, 1994). Reduced anterior-posterior force could also be a sign of quadriceps weakness, which has been shown to affect post-operative recovery (Mizner *et al.*, 2005).

Table 5.2 Preliminary work in assessing potential pre-operative predictors of post-operative outcome. Correlations satisfying $p<0.05$ are shaded light grey, and $p<0.01$ in dark grey.

	Percent reduction B(OA)		
	Correlation	Sig.	N
Age	-.660*	.020	12
BMI	-.032	.923	12
Cadence	.747**	.005	12
Stance percent	-.695*	.012	12
KOS pre-op operative leg	.102	.753	12
KOS pre-op non-operative leg	.280	.379	12
OKS pre-op operative leg	.253	.428	12
OKS pre-op non-operative leg	.247	.439	12
Knee flexion at HS	-.228	.475	12
Knee flexion at TO	.017	.959	12
Knee sagittal ROM	.519	.084	12
P Stance flex	.020	.950	12
P swing Flex	.636*	.026	12
Vertical GRF peak 1	-.158	.625	12
Vertical GRF peak 2	.197	.539	12
Dip in vertical GRF	.254	.426	12
Posterior GRF peak	.541	.070	12
Anterior GRF peak	-.637*	.026	12
ML GRF Max	-.492	.104	12

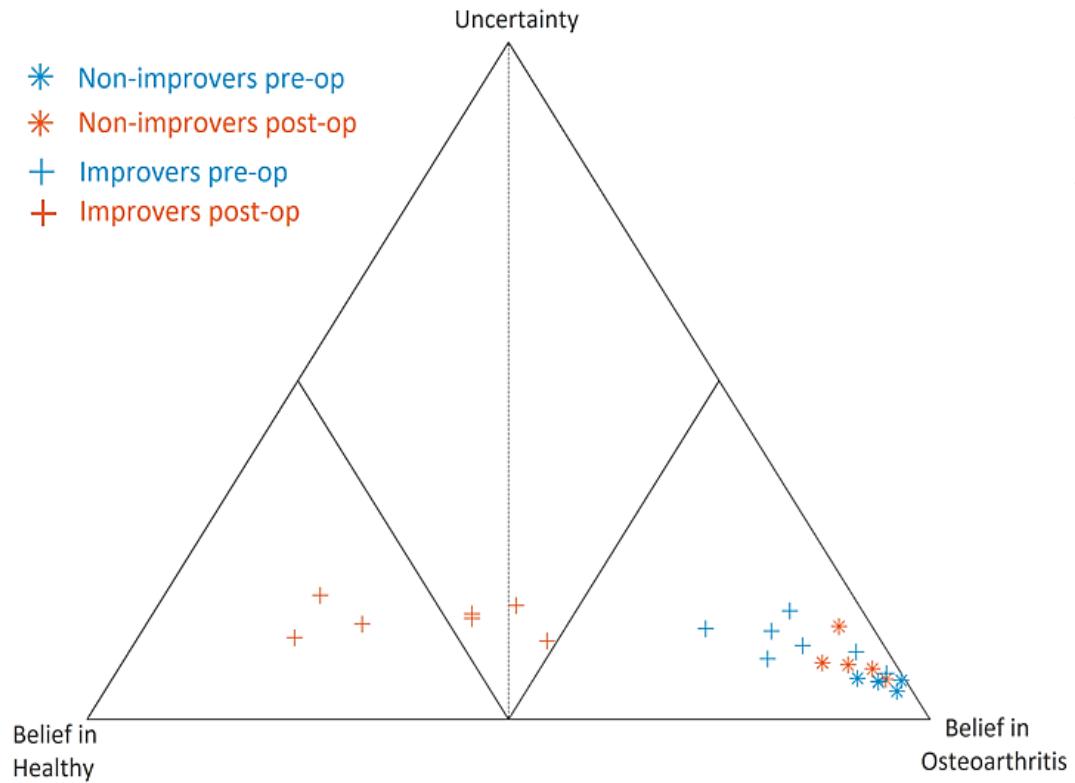


Figure 5.3 The resultant gait biomechanics classification of the 12 subjects pre (blue) and post (orange) operatively. Improvers are labelled as a cross; non-improvers are labelled as an asterisk. This illustration was created using the data and trained classifier from Watling (2014).

5.3 Methods

5.3.1 Participants

Of the 41 OA subjects recruited included within the previous chapter, 22 returned for a post-operative visit, which was initially planned to take place at 12 months post-surgery. There were several practical issues which prevented subjects from being recruited back at this point, which included:

- Participant unable to attend due to medical reasons
- Unable to contact participant
- Participant going on long holiday e.g. cruise
- Unforeseeable circumstances which lead to a temporary halt of all participant data collections within the ARUK centre.

The median time after surgery was therefore 13.2 months, however, these ranged between 9.3 months and 22.8 months. Subject demographics are shown in Table 5.3.

Table 5.3 Demographics of the 22 subjects included within this chapter

	Pre-op age	Height (m)	Pre-op weight (kg)	post-op weight (kg)
Mean	68.7	1.65	92.1	92.9
STD	8.3	0.1	23.5	23.6

5.3.2 Data Analysis, Processing and Classification

3D human motion analysis data was captured during level walking for each of the subjects pre and post-operatively using the methods previously described in Section 3.2. Joint kinematics and kinetics were calculated using the same techniques described in the previous chapter. The level gait biomechanical information was then outputted from Visual 3D, in preparation for PCA. The eigenvectors of the top 18 ranked PCs were taken from the previous chapter and used to calculate the PC scores pre and post-operatively from the corresponding waveforms.

At this stage, an input file could be created which used the same input variables as the data classifier in the previous section. A final column was added to the data called the

“class label” and was assigned a value of “1” for subjects pre-operatively, and “2” for subjects post-operatively.

The control variables k , θ , A and B of the trained classifier defined within the previous chapter were used to classify the level gait biomechanics of the 22 subjects pre and post-operatively.

5.3.3 Patient-reported Outcome Measures

Patients undergoing the same biomechanical assessment SOP as previously mentioned in other chapters, were asked to complete several commonly used PROMs. During the study, additional questionnaires were adopted, following suitable amendments to the ethical approval. In addition to this, the OKS and KOS scores were amended such that the participant responded considering each knee separately. The number of completed questionnaires are shown in Table 5.4.

Table 5.4 Number of completed questionnaires for collected PROMs

Pre/post-op	Leg	KOS	OKS	KOOS	PACS
Pre	Op	22	22	12*	16
	Other	15	15		
Post	Op	21	22	18*	18
	Other	15	16		

*Sports and recreation subscale from the KOOS could only be calculated on nine subjects pre-operatively and nine post-operatively.

5.3.4 Temporal-spatial Parameters

Spatial-temporal gait parameters were calculated within Visual3D. To represent between-leg differences, a level of symmetry between the two legs was calculated for each parameter. It was decided that this symmetry calculation, alongside the discrete parameter for the operative leg, would be sufficient in also describing the non-operative leg.

There are several different methods of calculating a symmetry index which have been applied to gait. The most commonly cited gait symmetry calculation is referred to as the Symmetry Index, which uses the following calculation:

$$\text{Symmetry Index} = \frac{|X_L - X_R|}{0.5(X_L + X_R)} \cdot 100\%$$

To translate this into a mathematical verbal expression – you take the magnitude of the difference between the left and the right leg, and represent it as a percentage of the average value between the two legs. It therefore gives no information as to which leg is dominant.

Another technique used is the symmetry ratio:

$$\text{Symmetry ratio} = \left(\frac{X_R}{X_L} \right) \cdot 100\%$$

This is a very easily interpretable measure, and is simply the percentage ratio of the right leg relative to the left. This can also be represented as the non-operative leg relative to the operative leg, depending on the context of the study (Patterson *et al.*, 2010).

One of the problems with this measure is its non-symmetrical nature. Figure 5.4 displays the value of the symmetry ratio for different values of the ratio r between two legs.
Where:

$$RI = \left(\frac{X_{op}}{X_{other}} \right) \cdot 100\%$$

As the Ratio Index is simply the ratio between the two legs expressed as a percentage, it can be seen that the Ratio Index starts to increase rapidly as the ratio favours to numerator (the operative leg).

This can be addressed simply by taking the log of the Ratio Index, referred to in the literature as the Gait Asymmetry:

$$\text{Gait Asymmetry} = \ln \left(\frac{X_{op}}{X_{other}} \right) \cdot 100\%$$

Where when $GA>0$ the numerator (operative limb) is dominant, and when $GA<0$ the contralateral limb is dominant.

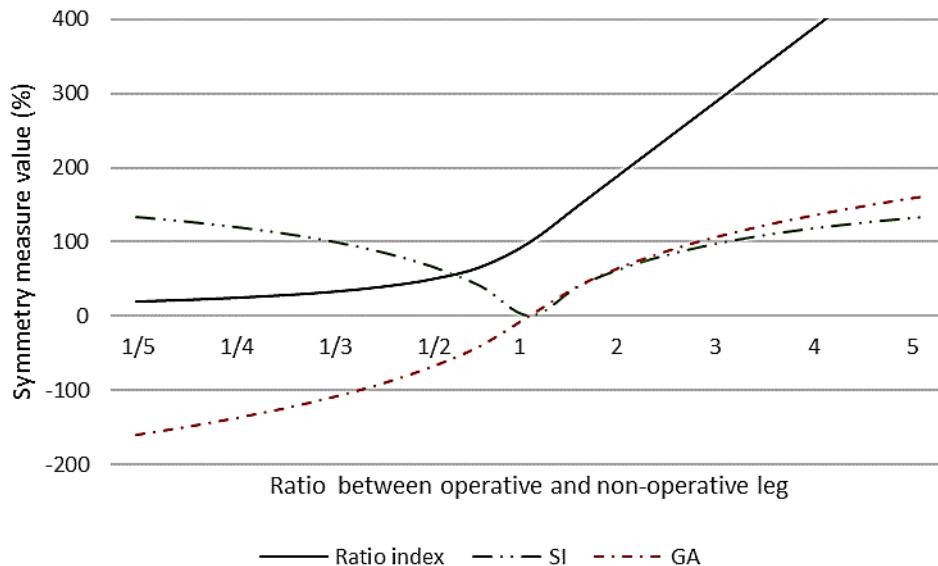


Figure 5.4 Comparison of the Ratio Index, Symmetry Index (SI) and Gait Asymmetry (GA) in representing between-leg differences.

The GA therefore appears preferable as a measure as it is possible to identify the dominant side by the sign of the number, and the absolute magnitude of the measure is the same, no matter which side is dominant.

5.3.5 Objective Improvement in Function

The simplex plot used to display and interpret the belief values calculated using the Cardiff Classifier allows an intuitive visual representation of uncertainty, however is less suited to statistical analysis. To define an objective improvement in function, it would be useful if the changes of the three belief values were reduced to a single discrete variable, in a meaningful way, without discarding important information. One method which has previously been used to define post-operative improvement, is to simply calculate the change in belief of OA.

Table 5.5 Example of differences in the defining functional gait changes using reduction in B(OA) and the proposed functional gait improvement (FGI) method for three subjects.

	Preoperative			Post-operative			B(OA) reduction	FGI
	B(OA)	B(NP)	B(U)	B(OA)	B(NP)	B(U)		
Example 1	80	10	10	50	10	40	30	15
Example 2	70	10	20	40	30	30	30	25
Example 3	20	30	20	10	70	20	10	25

This method, however, completely discards information on the changes in B(U) and B(NP). If we compare the post-operative changes of Example 1 and Example 2 within Table 5.5, both would be deemed to have improved equivalent amounts – a reduction of B(OA) of 30%. Example 2, however, also has an increase in belief B(NP), and therefore, when interpreting the simplex plot during visual inspection, would have been deemed to have achieved greater recovery than Example 1.

It therefore seems apparent that there may be some value in including the level of recovery of healthy function within the definition of biomechanical improvement. An additional calculation is proposed which will be referred to as the Functional Gait Improvement (FGI).

$$\text{Functional Gait Improvement (FGI)} = \frac{\Delta B(NP) - \Delta B(OA)}{2} \quad (5.1)$$

Where:

- 100% represents a change from 100%OA to 100% NP
- 0% could represent no change, or could represent an equal increase in B(OA) and B(NP)
- 50%, for example, could represent a decrease of B(OA) of 30% and an increase in B(NP) of 70%

Table 5.5 also shows the calculated FGI of those two subjects. Notice how Example 2 is now deemed to have recovered more than Example 1. Another subject within this table, Example 3, appeared to predominantly have an increase in B(NP) post-operatively, hence reduction of B(OA) was only 10%. Notice, however, that the FGI is the same; i.e. reduction in B(OA) is weighted the same as an increase in B(NP).

5.4 Results and Discussion

5.4.1 Does Functional Recovery Return Following TKR?

The simplex plots illustrating the CBOE of all 22 subjects before and 9+ months post-surgery are shown in Figure 5.5. Pre-operatively, 19 of the 22 subjects were within the dominant OA triangle where the belief of OA is greater than that of the combination of the other two belief values. Post-operatively, however, only ten of the 12 subjects remained dominant B(OA), with all but three subjects displaying a reduction in B(OA).

The two greatest improvements were subjects 8 and 18, who moved from dominant OA to dominant NP, followed by subject 3, who moved from dominant OA to non-dominant NP.

Figure 5.6 shows that the variable which had the most marked reduction in B(OA) across the 22 subjects was the second PC of the hip adduction moment, which reconstructed the rate of loading at early and at late stance, and the amount of 'double peak' in the waveform. Of the 15 PCs of the biomechanical variables which resulted in a reduced B(OA) following surgery, eight had statistically significant reductions in the PC score using the *p* value threshold of <0.01 , and two further PCs only reached the first threshold of $p<0.05$ and should therefore be interpreted with caution. It can be seen from this figure that on average, uncertainty increased in 13 of the 18 input variables. This can be confirmed visually in the simplex plots in Figure 5.5, as post-operatively subjects are closer to the uncertainty vertex. All statistically significant changes were mean changes towards that of NP subjects.

Interestingly, despite significant improvements in PC2 of the hip adduction moment; which reconstructs a biphasic waveform, changes in the second PC of the hip adduction angle were much more modest and didn't reach statistical significance. The best interpretation of PC2 of the hip adduction angle is that of a more abducted hip angle during stance phase, and adducted position during swing phase of the gait cycle. The first PC reconstructed a large magnitude offset throughout the gait cycle, hence PC2 reconstructs smaller changes relative to that offset.

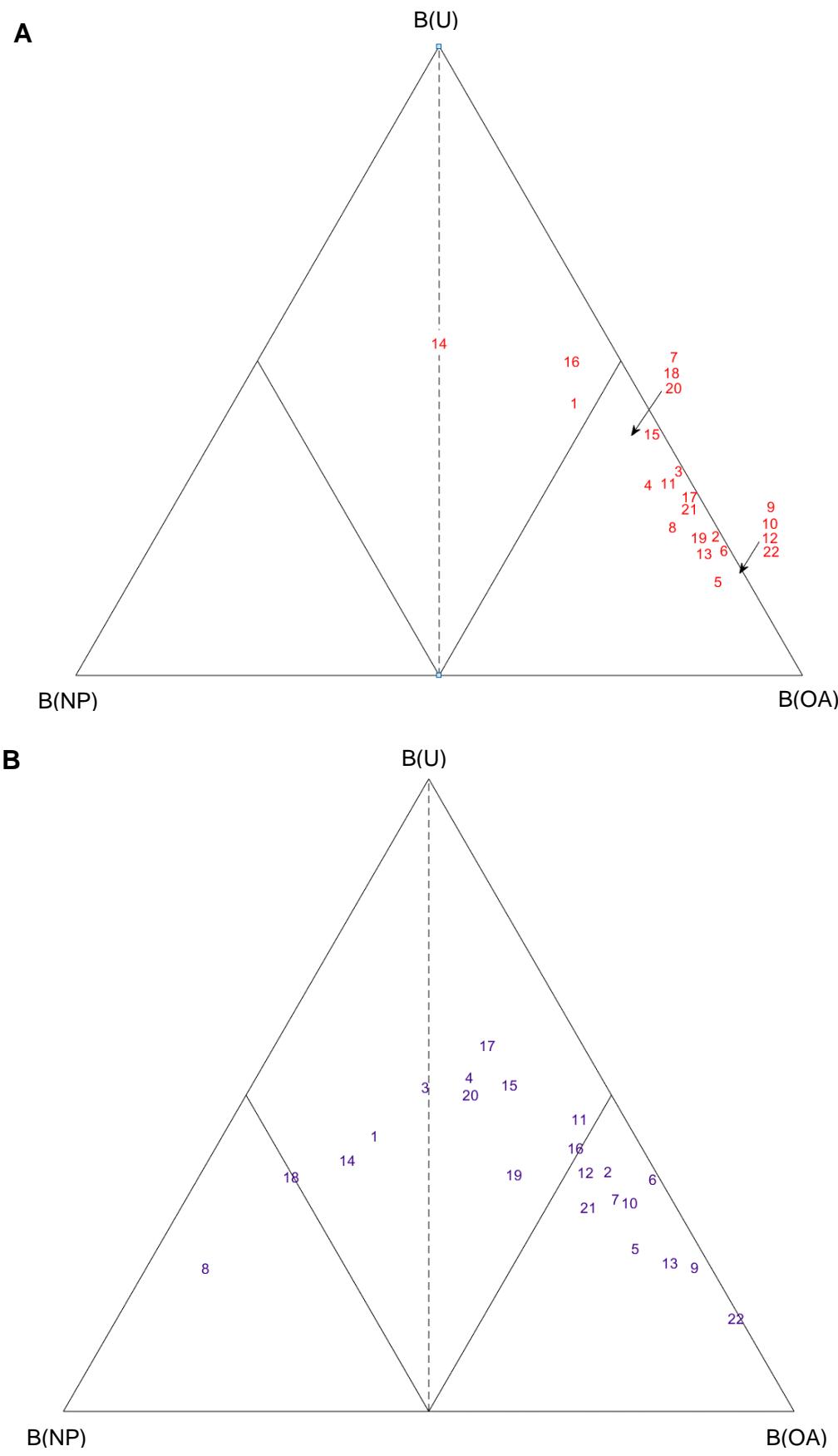


Figure 5.5 Simplex plot of the CBOE for all 22 subjects (numbered) both pre (A) and post (B) surgery.

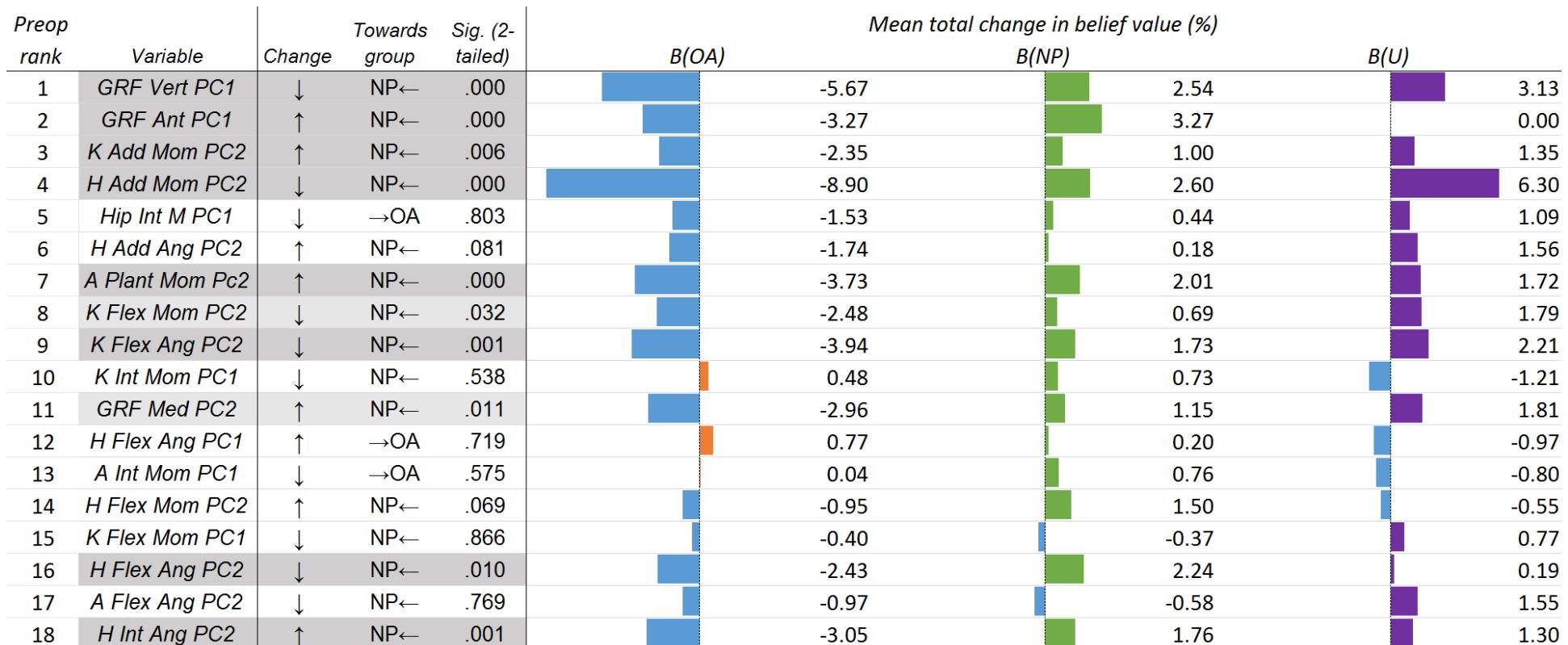


Figure 5.6 Illustration of the mean total change in belief values of each of the inputs of the classifier between pre and post-operative classification. Variables are ordered according to the ranked order of importance of contributions of the classification (see Section 4.3.4). Also shown is the level of statistical significance of each post-operative change in input variable. Statistically significant changes of $p < 0.01$ are shown in dark grey, and $p < 0.05$ shown light grey.

The subject with the highest pre-operative value of the second PC of the hip adduction angle is shown on the left within Figure 5.7, and the subject with the second highest can be seen on the right. During stance phase, the pelvis drops on the side of the leg. This effectively reduces the coronal plane angle of the femur relative to the pelvis during stance phase, and increases the adduction angle during swing phase. In NP subjects, the pelvis often drops a small amount towards that of the leg in swing phase, as opposed to that of the leg in stance. This is often exaggerated in subjects with hip OA and weak hip abductors, and is commonly referred to as ‘Trendelenburg gait’ (Whatling *et al.*, 2015). Interestingly, despite Trendelenburg gait being found in one study to be related to severe medial knee OA (Mündermann *et al.*, 2005), the second PC reconstructs pelvic drop in the opposite direction. To hold the pelvis in this position, the subjects might either have relatively strong hip abductors, or be leaning their trunk over towards the leg in stance. One potential explanation for this gait adaptation is that the pelvic drop towards the limb in stance and helps to lift the limb in swing phase. This is often referred to as “hip hiking” (Kerrigan *et al.*, 2000), and is a strategy that can increase ground clearance when hip flexion, knee flexion, or ankle dorsiflexion during swing phase is reduced.

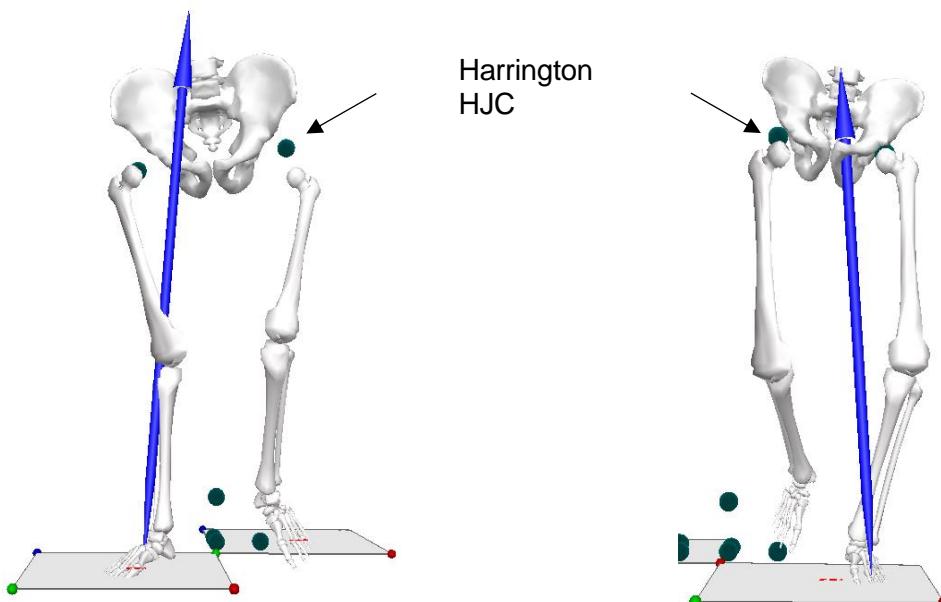


Figure 5.7 Coronal view of the two subjects with the largest hip adduction PC2 score during stance phase. Subject 13 (left), and Subject 8 (right). The Harrington HJC derived from the static file is derived during the static file.

It can also be seen within Figure 5.7 that the HJC derived from the pelvis coordinate system is very different to that derived from the tracking markers of the thigh. Due to STA during walking of bony landmarks used to define the pelvis (ASIS and PSIS), the virtual landmark of the Harrington HJC is only used within the model definition and not during tracking. Such a large discrepancy between coordinate systems, however, might also create a lack of trust in the estimation of the femoral head using the thigh marker cluster. Subject 13 did have the second highest BMI of the cohort at 44.5 kg/m^2 , as opposed to 26.1 kg/m^2 for subject 8.

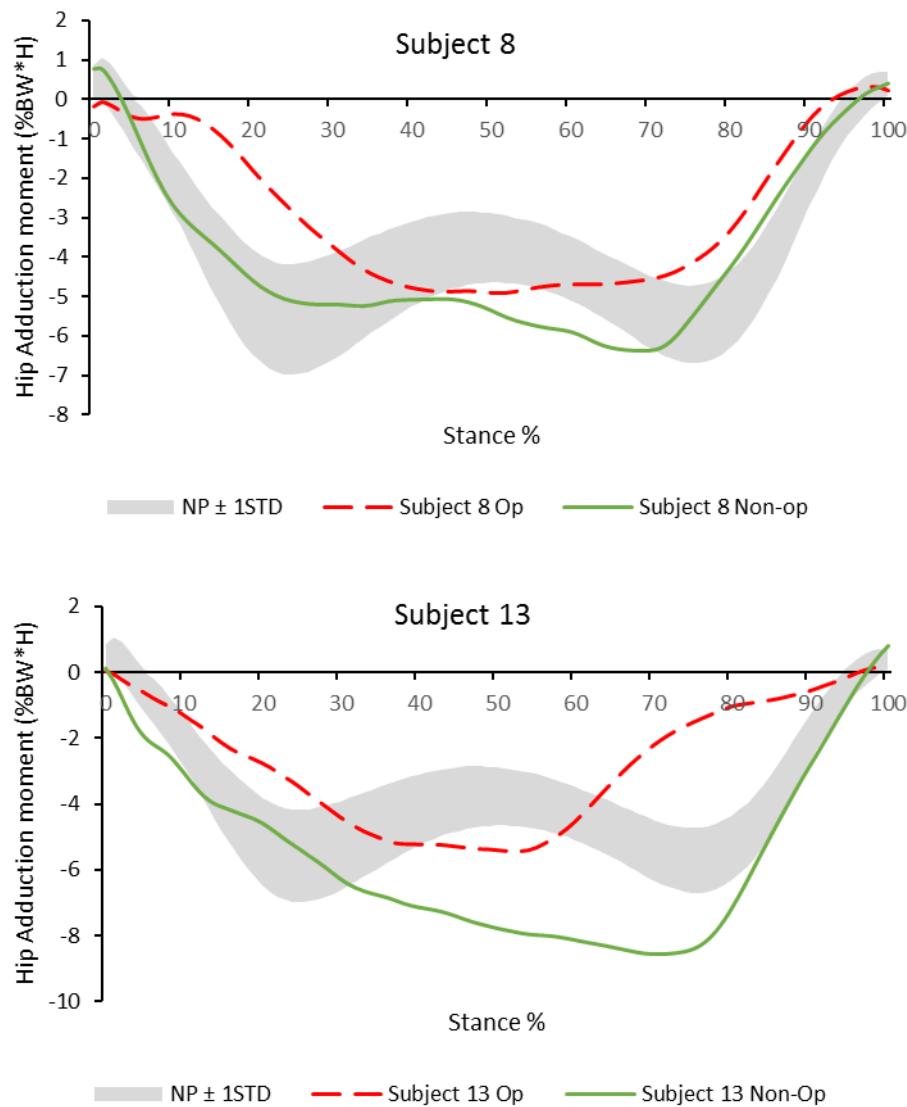


Figure 5.8 Pre-operative hip adduction moments for both operative (Op) and non-operative (Non-Op) legs of subjects 8 and 13.

It isn't possible to deduce why this biomechanical characteristic of osteoarthritic gait doesn't significantly improve following TKR surgery. One potential cause could be that the subject had walked in this manner for a long time, and hence this might have a long-term effect on hip abductor strength. Figure 5.7 shows the hip adduction moments of the operative and non-operative limb before surgery for the same two subjects. Anecdotally, it appears that both subjects have a reduced hip adduction moment on the operative limb, and increased on the contralateral side.

5.4.2 Defining Functional Improvement

Within Section 5.4.2 it was suggested that the technique used by previous authors to define changes in classification of OA function; simply looking at the magnitude reduction of the B(OA) value, might not be sufficient in summing functional improvement. Figure 5.9 shows that within this cohort there was a very strong linear correlation ($r=0.948$) between the suggested new measure and that used by previous authors. The new measure, named the FGI, would in theory provide a superior representation of functional changes in situations where a reduction of B(OA) was not strongly associated with an

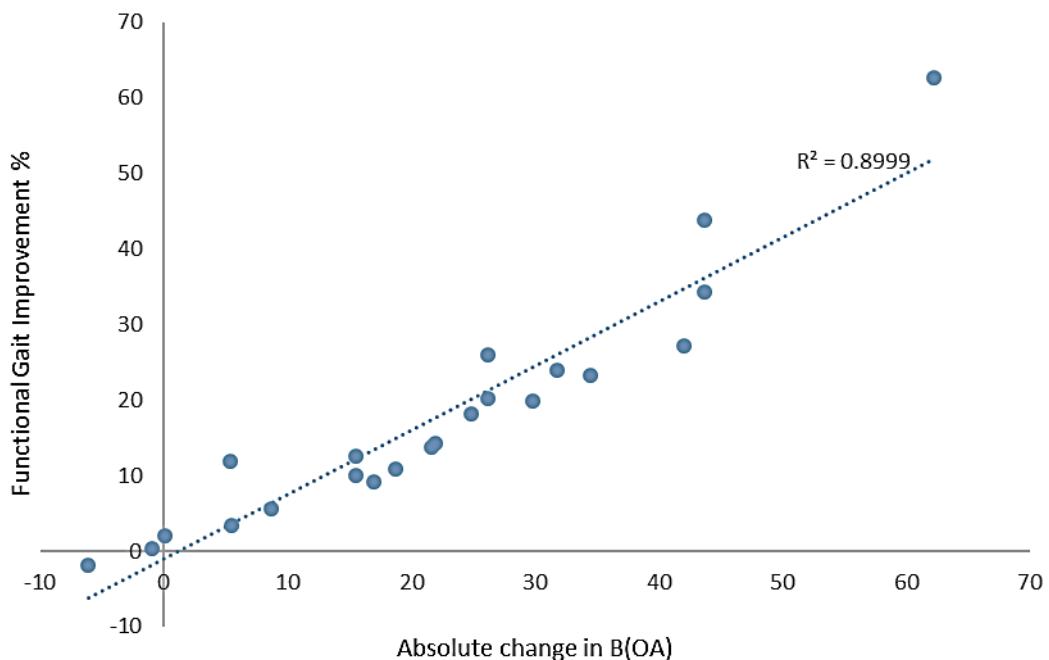


Figure 5.9 The linear correlation between absolute change in B(OA), and the new summative measure proposed in this chapter called; the Functional Gait Improvement (FGI).

increased B(NP). Considering the very strong relationship between the two measures, it was decided that to maintain optimal interpretability of findings, functional improvement shall be defined as the reduction of B(OA) throughout the rest of the analysis.

5.4.3 Functional Improvement of Each Limb

The reduction in B(OA), of both the operative and the non-operative (contralateral) limbs, is summarised within Figure 5.10. It is much clearer to see in this arrow chart that the B(OA) reduced post-operatively for all but 3 subjects on the operative limb, and reduced for all but 6 subjects on the contralateral side (non-operative). There was a strong correlation ($r=0.807$, $p<0.001$) between reduction of B(OA) on the operative side, and a reduction on the contralateral side.

Poorest functional improvement:

Subject 16: This subject was a 73-year-old female with a BMI of 24.6 (73kg, 1.65m), who reported severe OA in both knees and was undergoing a TKR on the right knee. Before surgery, the right (operative) leg had a B(OA) of 42%, and an OKS percentage score of 65%, while the left had a B(OA) of 57% and an OKS of 67%. The B(OA) therefore suggests biomechanical function may have been poorer on the non-operative limb. The subject didn't suffer from OA in any other joints, however did have back pain, and they reported swollen toe joints.

At six months post-surgery the subject 16 reported being very happy with the right knee TKR, despite numbness around the patella and clicking/knocking when rising from being seated. The subject reported that the left knee would likely require a TKR soon but that it was not imminently necessary. The OKS score was 73% for the right, and 69% for the left.

At nine months following surgery the subject reported stiffness in the right TKR. The subject also reported that there was an 'odd' feeling of numbness and that it no longer 'felt like *her* leg'. The subject was still, however, happy with the TKR. The subject had been told that a TKR on the opposite leg was a possibility if they wanted one, however

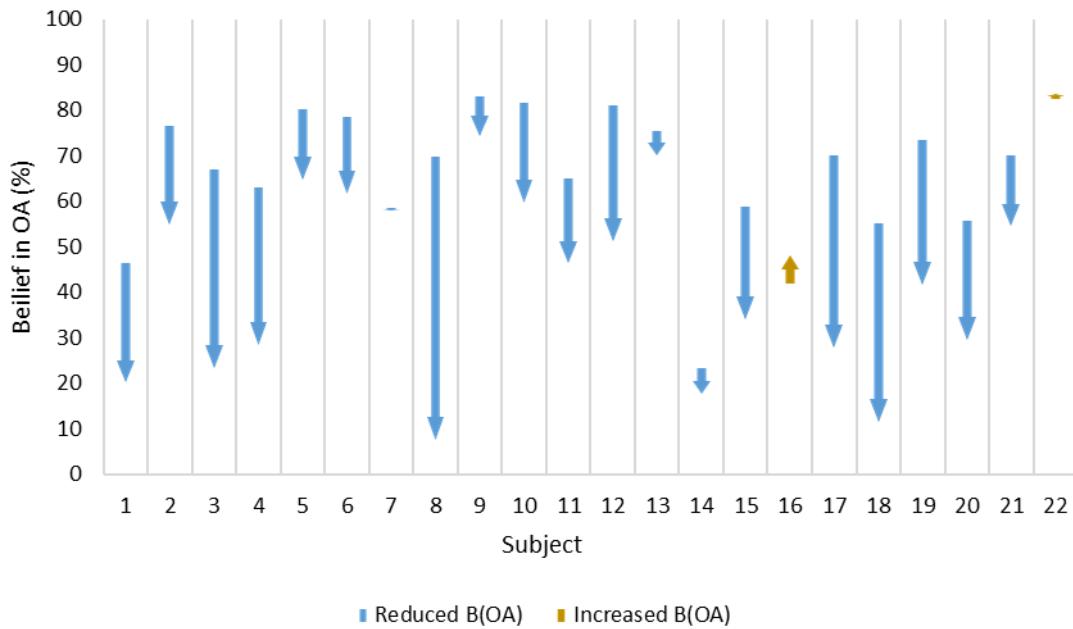
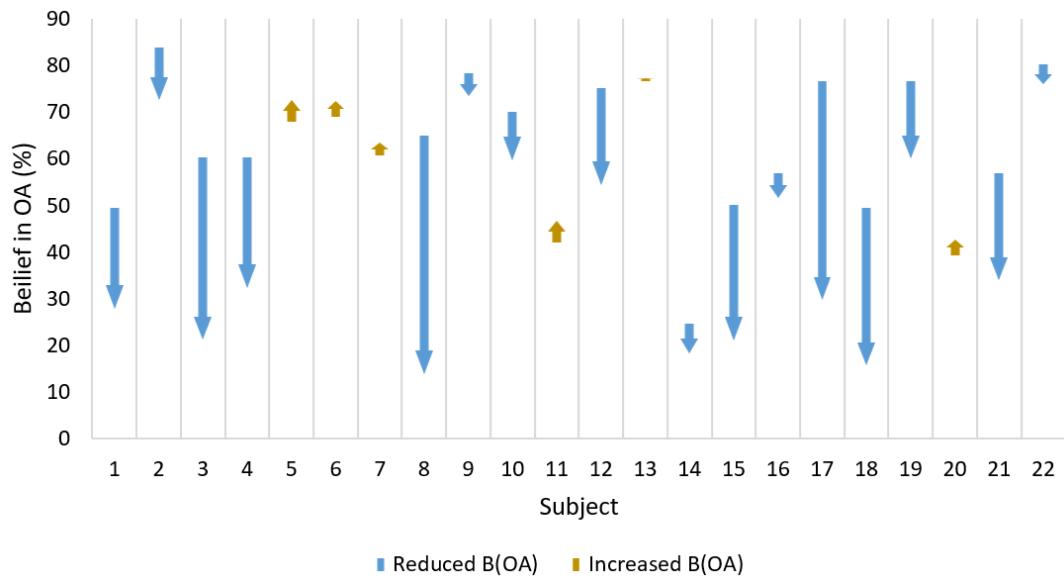
A) Operative leg:

B) Non-operative leg:


Figure 5.10 Arrow chart of individual changes in B(OA) for all 22 subjects from pre to post-surgery of the operative leg (A), and non-operative leg (B). Blue arrows represent a reduction in B(OA), and orange an increase. Subject 7 had a negligible reduction, and subject 22 had a very small increase in B(OA) on the operative limb.

they were trying to avoid it because they didn't want the 'six-week aggro'. The OKS for the right leg was 79%, and 75% for the left. The subject was taking medication for back pain.

During the final visit, 12 months post-surgery, subject 16 reported there had been no change in right TKR since the previous visit. They reported that the right knee remained stiff and the left knee painful – particularly during stair-climbing. The OKS for the right leg was 90% and the 69% for the left. The B(OA) of the right leg was 48%, and the 52% for the left.

In summary, both the OKS and B(OA) suggested that the functional status of both knees was similar pre-operatively; which matches that reported by the subject. The OKS score indicated a gradual improvement of 25 percentage points on the operative knee, however the DST classifier indicated modest worsening in function with B(OA), increasing from 42% to 48% on the operative limb. The OKS score had improved slightly from 67% to 69% on the other leg, and the classification also indicated a modest reduction in B(OA) from 57% B(OA) to 52% B(OA).

Figure 5.11 and Figure 5.12 display changes in the individual belief values assigned to each of the input variables of the classifier for subject 16. It shows that the main input variable to improve on the operative leg was the second PC of the ankle plantarflexion moment. Referring back to Section 4.3.4, the first PC reconstructed large magnitude differences between the plantarflexion moment throughout stance, particularly towards toe-off, while the second PC represented a relative increase in the plantarflexion moment in early stance in comparison to that of late stance. In order to confirm the findings of the combined PCA and classification, the pre and post-operative plantarflexion moments for the subject are shown in Figure 5.13. It can be seen that, while the absolute magnitude of the ankle plantarflexion moment hasn't improved, the magnitude of the plantarflexion moment during the first half of stance, relative to that during the second half, was increased pre-operatively and improved post-operative. This feature is quite subtle and might have gone un-noticed had it not been for the application of PCA.

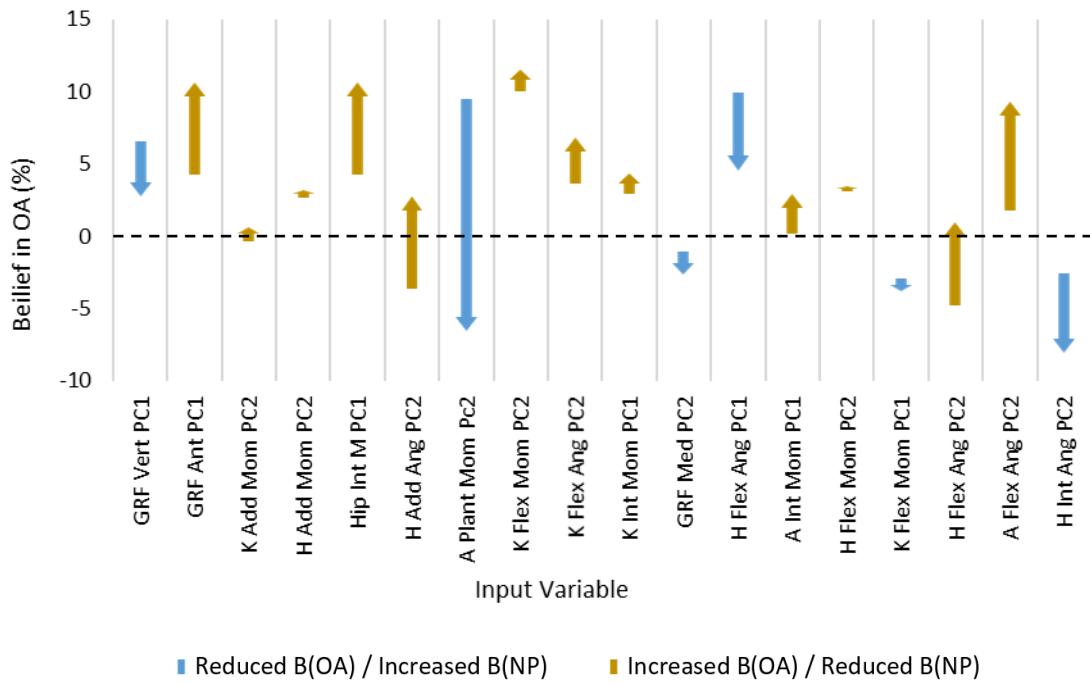


Figure 5.11 Changes in individual bodies of evidence for each of the 18 input variables for the operative leg of subject 16

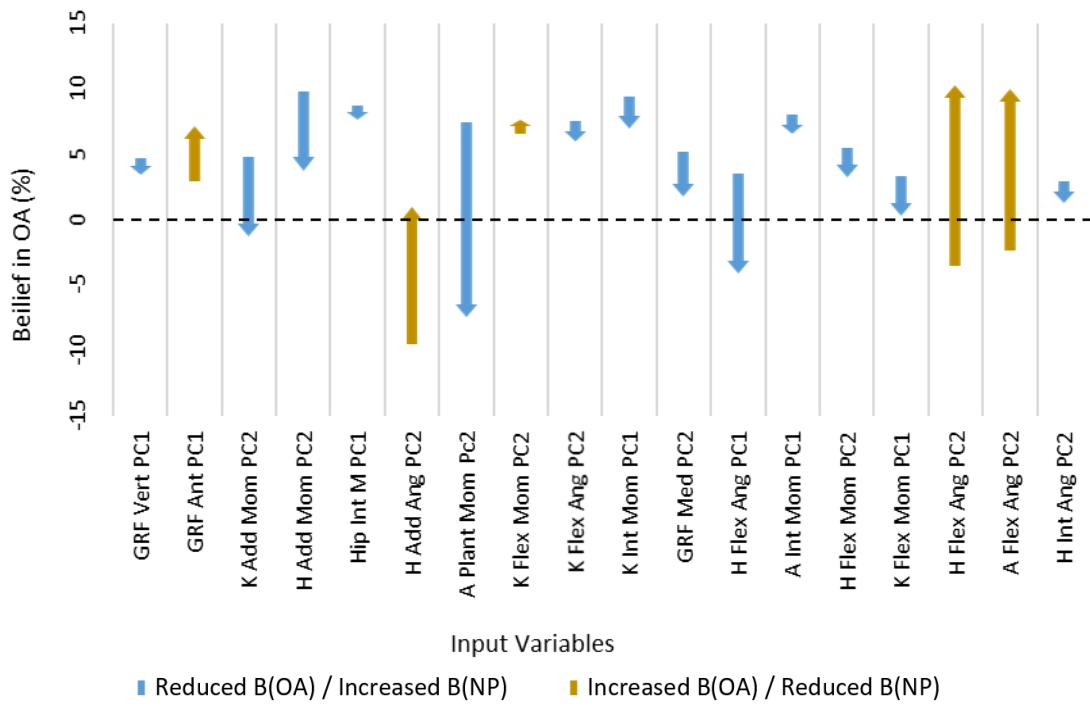


Figure 5.12 Changes in individual bodies of evidence for each of the 18 input variables for the non-operative leg of subject 16

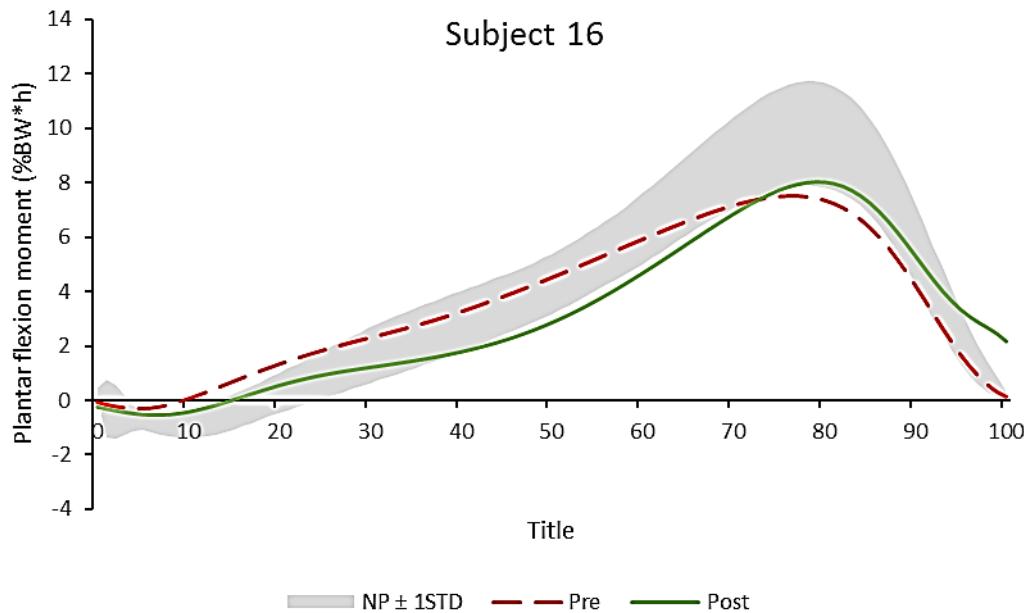


Figure 5.13 Ankle plantarflexion moment of subject 16 pre (red dashed) and post (solid green) TKR surgery.

Figure 5.12 illustrates how each input variable contributed to pre and post-operative B(OA) on the non-operative (left) limb. H Add PC2, previously contextualised as “hip hiking”, has resulted in an increase in B(OA) bilateral. This could perhaps be related to the reported “stiffness” of the operative limb, or even due to the reported OA and pain on the non-operative limb.

5.4.4 Greatest Functional Improvement

It can be seen from Figure 5.10A that the greatest reduction in B(OA) was achieved by subject 8, changing from 69.8% B(OA) to only 7.5%. Subject 8 was a 67-year-old female with a BMI of 26.14 (1.54m, 62kg). Both hips were painful due to the left knee pain, the left hip causing the most pain. This hip pain had been present for months, as opposed to years. This had resulted in the use of crutches as of two weeks prior to the pre-operative visit, specifically to ease hip pain. The subject also reported back pain.

Pre-operative, the subject had an OKS of 16.7% on the left knee and 83.3% on the right. At three months following surgery, the subject reported a large improvement of the left leg following surgery, reporting it as “straighter” and easier to walk on. They reported

that, alongside the standard physio care, they were also seeing a personal trainer every other week. The OKS for the left leg was now 73%, and 90% for the right. Despite these improvements, the subject noted swelling of the left ankle. The subject added that they may have previously fractured the left foot laterally.

At seven months post-surgery, the subject reported more “crunching” in the knee since their last visit, however not much pain. They felt that their movement had improved, and that they were back to normal. They did still have pain in their lower back. The OKS was 98% for both knees. The subject also reported numbness around the scar, and more swelling around the ankle. The ankle swelling was accompanied by painful veins above the ankle.

At 12 months post-surgery, the subject reported that the knee swells slightly if not rested, or if they sleep in certain positions. They were only conscious of the knee when they go to kneel, and felt they could perform all-over ADLs normally, despite clicking noises from the joint. They reported that they had no post-operative pain and didn’t remember having much pain before TKR surgery. They were still seeing a personal trainer, and found the exercise bike to be the best activity. They tended to avoid doing squats as they felt this wasn’t good for the knee.

Considering the individual BOE of the subject, nearly every input variable of the operative leg improved other than the second PC of the ankle plantar/dorsi flexion angle. Referring to Section 4.3.4, this PC reconstructs the slightly reduced ankle dorsi-flexion during the first half of stance phase, but primarily the delayed initiation of plantarflexion, and the reduced peak plantarflexion during toe-off/early swing. This is interesting considering the subject had noted swelling of the left ankle (operative side) following surgery.

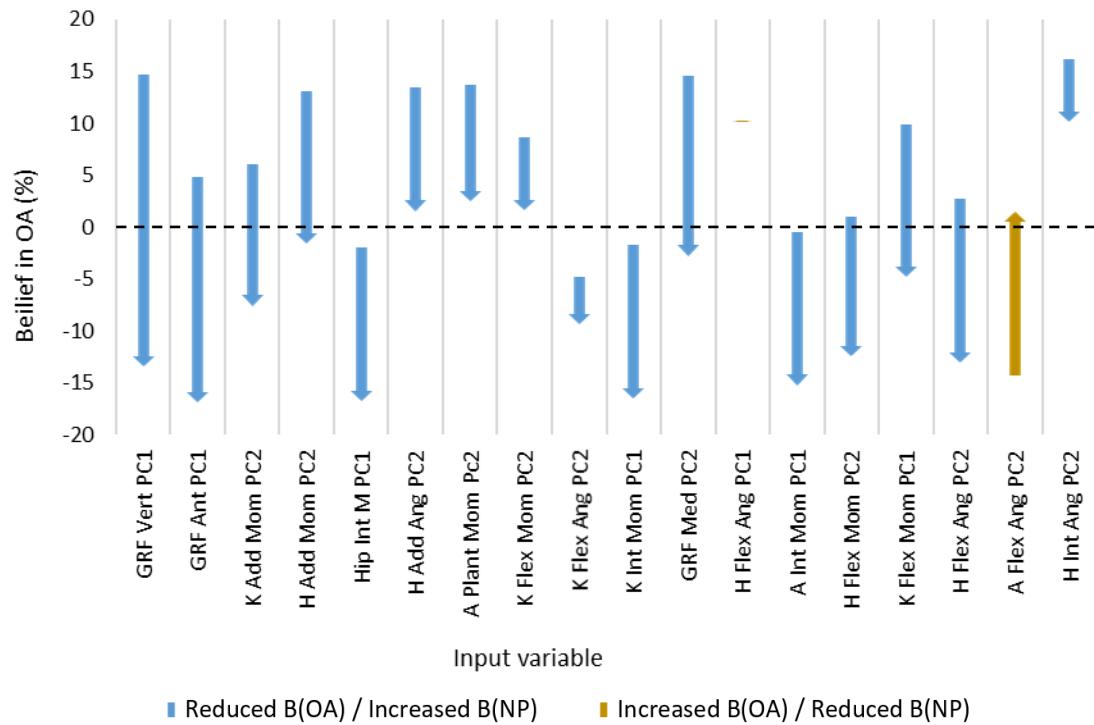


Figure 5.14 Changes in individual bodies of evidence for each of the 18 input variables for the operative leg of subject 8

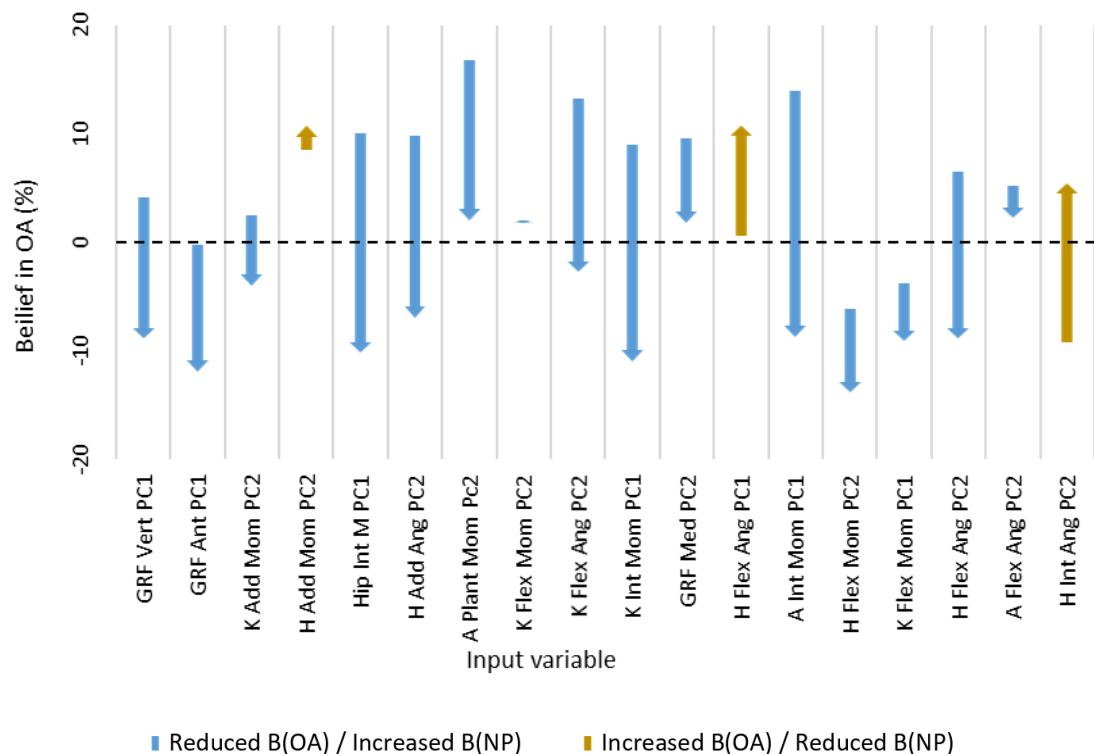


Figure 5.15 Changes in individual bodies of evidence for each of the 18 input variables for the non-operative leg of subject 8

5.4.5 Do Pre-, Post-, and the Relative Change in Subjective Outcome Measures Correlate With Changes in Biomechanical Gait Classification?

Correlations between the objectively measured gait function, B(OA) and the various PROMs, are shown in Table 5.6. For further clarity: pre-operative B(OA) were correlated to pre-operative PROMs, post-operative B(OA) to post-operative PROMs, and changes in B(OA) correlated to changes in PROMs between the pre-and post-operative visits. Box and whisker plots have also been included to illustrate the changes within each measure following surgery (Figure 5.16). Within this figure, the B(OA) has been subtracted from 100%, such that the interpretation matches that of the other PROMS; i.e. 0% is pathological, 100% is healthy.

KOS and OKS: The KOS and OKS PROMs correlated moderately with pre-operative B(OA), post-operative B(OA), and the change in B(OA) following TKR surgery. The statistical significance of these correlations was consistently high ($p<0.01$), indicating a strong likelihood that there is a consistent relationship between these measures, and the objective classification of gait biomechanics. The pre-operative relationship is illustrated within Figure 5.17, and helps visualise the strength of the correlation. The large number of subjects who are between 75% to 85% B(OA) appear to also have quite a large range of OKS and KOS scores.

Table 5.6 Correlations between B(OA) and each PROMS pre-operatively, post-operatively, and the post-surgical change of B(OA) and PROMs.

		KOS	OKS	PACS Pain	KOOS						
Pre B(OA)	Corr				Symp toms	Pain	ADLs	Sport/ Rec	QOL	Total	
	Sig.	.004	.008	.343	.303	.146	.134	.484	.255	.165	
	N	22	22	16	12	12	12	9	12	12	
Post B(OA)	Corr	-.599**	-.652**	-.509*†	-.443*	-.469*	-.591**	-.501*	-.584**	-.551**	
	Sig.	.002	.001	.015	.033	.025	.005	.029	.005	.009	
	N	21	22	18	18	18	18	15	18	18	
Change B(OA)	Corr	-.637**	-.697**	0.175	-.388	-.364	-.624*	.031	-.401	-.509*	
	Sig.	.001	.000	.275	.106	.122	.015	.471	.098	.046	
	N	21	22	14	12	12	12	8	12	12	

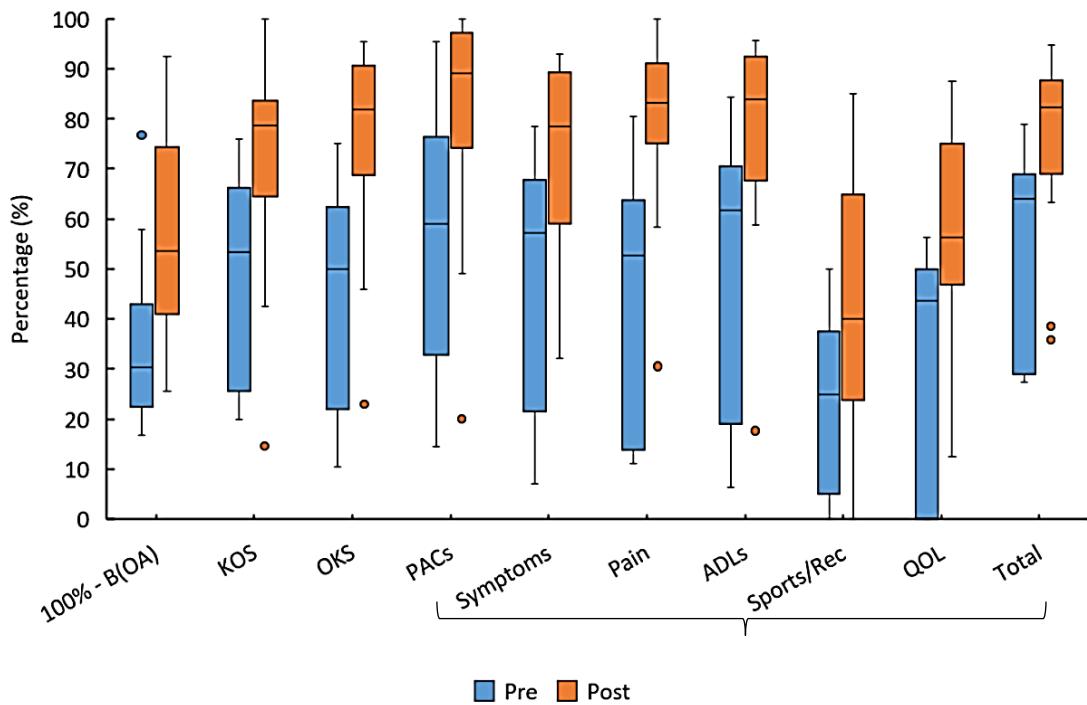


Figure 5.16 Box and whisker plots of B(OA) and PROMs before and after surgery. B(OA) has been represented as 100% - B(OA), such that magnitude interpretation matches that of the other PROMs.

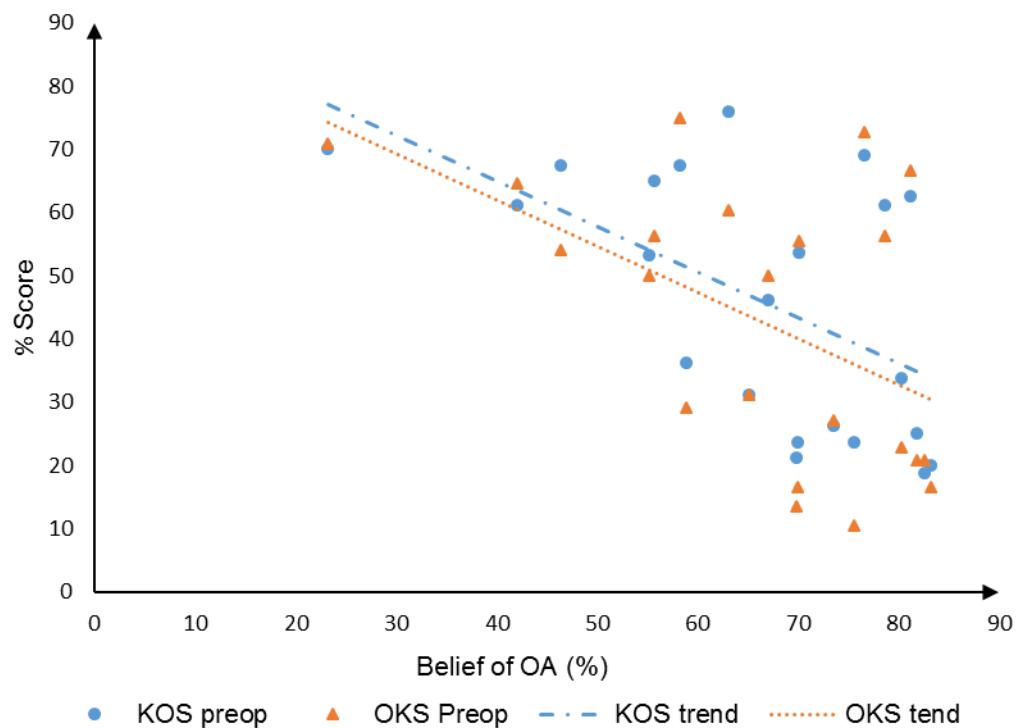


Figure 5.17 Illustration of the strength of linear relationship between the B(OA) score, and both the KOS (blue circle, linear trend dash-dotted) and OKS (orange triangle, linear trend dotted) pre-operatively.

The OKS has been specifically designed to assess the outcome of joint replacement, while being influenced as little as possible by other comorbidities (Murray *et al.*, 2007). The score was also recommended by the department of health in the UK for quantifying TKR outcomes (Smith *et al.*, 2005), and a number of studies have shown that post-operative OKS can be used to predict satisfaction (Clement *et al.*, 2013). It is therefore a particularly positive finding that functional gait classification appears to be moderately correlated to the OKS in particular. Had a very strong correlation been found, the conclusion might have been that the functional gait classification was adding little to the understanding of outcome. Had a very weak correlation been found, this would have laid question as to whether changes in gait classification are clinically relevant to the patient.

PACS There appeared to be no relationship between PACs pain score and pre-operative gait classification (pre-B(OA)). However, it appears that post-operative gait function was more strongly correlated to post-operative pain. Thus, post-operative change in B(OA) doesn't appear to be correlated with post-operative reductions in pain. These results suggest that biomechanical gait adaptations of late-stage osteoarthritic subjects are beyond that of typical antalgic gait adaptations, which aim to reduce weight bearing. It is always challenging to identify potential causal relationships between correlated variables and these hypotheses should therefore be treated with caution until further tested. This is particularly true in interpreting the relationship between pain and biomechanics following surgery, as post-operative pain may alter function, poor biomechanical function following surgery could lead to pain, or level of function restored may well be simply a good predictor of level of pain reduction.

KOOS Pre-operatively there were no statistically significant correlations between B(OA), and either the KOOS cumulative score, or any of the KOOS sub-scores. There was, however, reduced statistical power due to only having pre-operative KOOS scores for 12 out of the 22 subjects. Much like PACs, there appears to be a strong relationship between B(OA) and a KOOS scores post-operatively. This relationship reaches the greater threshold of significance of $p<0.01$ when correlating to KOOS-ADLs, KOOS – QOL and

the KOOS-Total. Perhaps due in part to a weak pre-operative relationship, only post-operative changes in KOOS-ADLs and KOOS-Total appeared related to post-operative changes in B(OA). Walking is a very common activity of daily living, and has cross-overs in terms of functional requirements with many other common activities. It is therefore not surprising that this sub score is most correlated with improvement in gait function.

Summary: The level of B(OA), within the Cardiff Classifier appears to be consistently moderately correlated to KOS and OKS scores. There appears to be greater correlation between patient-reported function and pain, and objectively measured gait biomechanics post-TKR surgery. This finding is contrary to the initial hypothesis, which proposed that patient-reported outcomes would be much greater than objectively measured function following surgery.

5.4.6 Predicting Post-Operative Improvement

Clinical Measures

The results of the analysis of potential biomechanical predictors of post-operative recovery are displayed in Table 5.2. Surprisingly, within this cohort, neither pre-operative age, weight or BMI were significantly linearly correlated to either the change in B(OA) following TKR surgery or improvement in OKS. The effects are, however, going in the expected direction – as literature suggests, increased weight and BMI are risk factors for poor TKR outcome (AAOS, 2015), and preliminary work identified advanced age as a potential pre-operative risk factor (See Section 5.2). All three were also correlated more with improvement in OKS, as opposed to B(OA), with age almost reaching the first threshold for significance with a p value of 0.07.

Considering the AAOS, established that there was a “strong” level of evidence of a relationship between obesity and TKR outcome, it is surprising that this wasn’t reflected within the findings of this study. One potential explanation is that, while BMI might be a risk factors of poor TKR outcome, the effect is unlikely to be linear. Previous studies which have identified high BMI as a risk factor have categorised BMI using defined threshold (Spicer *et al.*, 2001). The relationship between the magnitude of improvement in B(OA) and OKS, and the categorised BMI is shown within a stem and leaf plot in Figure 5.18. A more recent review of the effect of BMI on TKR outcome, which considered the findings of 50 articles, concluded that functional improvements appear to be equivalent between high and low BMI patients, however, the risk of complications was higher and implant survival was lessened in high BMI patients (Rodriguez-Merchan, 2014). This study supports these findings: high BMI subjects appear to have achieved similar functional gains at around 12-months post-operatively, however the longevity of the prosthesis in this cohort is yet to be determined. The findings presented in this thesis only consider relatively short-term outcome, and may not be representative of longer-term outcomes.

Table 5.7 Summary of linear correlations between changes in B(OA) and OKS, with pre-operative clinical, temporal-spatial, kinematic, kinetic and summative (classification) data.

Statistical significances below $p<0.05$ are shown in light grey, and $p<0.01$ in dark grey

		Change in BOA			Change in OKS		
		Corr	Sig.	N	Corr	Sig.	N
Clinical	Age	-0.19	0.41	22	-0.41	0.07	21
	Weight	-0.12	0.60	22	-0.26	0.25	22
	BMI	-0.15	0.50	22	-0.24	0.28	22
	Operative leg OKS	-0.17	0.45	22	-0.36	0.10	22
	Non-operative leg OKS	0.29	0.29	15	-0.13	0.65	15
Temporal- spatial	Cadence	0.07	0.75	22	0.13	0.57	22
	Stance width	0.18	0.44	20	-0.19	0.42	20
	Stance percent	-0.39	0.08	22	-0.37	0.09	22
	Stance time symmetry	-0.32	0.14	22	-0.32	0.15	22
	Double limb support	-0.16	0.52	20	-0.26	0.27	20
Kinematic	Knee flexion at HS	-0.07	0.75	22	0.19	0.40	22
	Knee extension late stance	0.05	0.82	22	0.37	0.09	22
	Knee peak flexion angle	-0.01	0.96	22	0.00	1.00	22
	Knee peak percent	-0.44*	0.04	22	-0.437*	0.04	22
	Knee ROM	0.03	0.89	22	-0.16	0.47	22
	Hip ROM	0.20	0.38	22	0.17	0.45	22
	Ankle ROM	-0.45*	0.04	22	0.32	0.15	22
	Hip adduction HS	-0.40	0.06	22	-0.27	0.23	22
	Knee adduction HS	-0.48*	0.03	22	-0.26	0.25	22
	Ankle Eversion HS	-0.59**	0.00	22	-0.549**	0.01	22
Kinetics	V GRF P1	-0.20	0.37	22	-0.16	0.48	22
	VGRF P2	-0.18	0.42	22	0.12	0.59	22
	V Drip ratio	0.30	0.18	22	0.34	0.13	22
	Posterior GRF peak	0.12	0.58	22	0.11	0.64	22
	Anterior GRF peak	-0.07	0.76	22	-0.11	0.64	22
	ML Peak 1	-0.09	0.69	22	-0.15	0.52	22
	ML Peak 2	-0.03	0.90	22	0.02	0.95	22
Summative	Pre-op B(OA) operative	-0.17	0.64	22	0.05	0.84	22
	Pre-op B(OA) non-operative	-0.13	0.95	22	-0.44	0.85	22

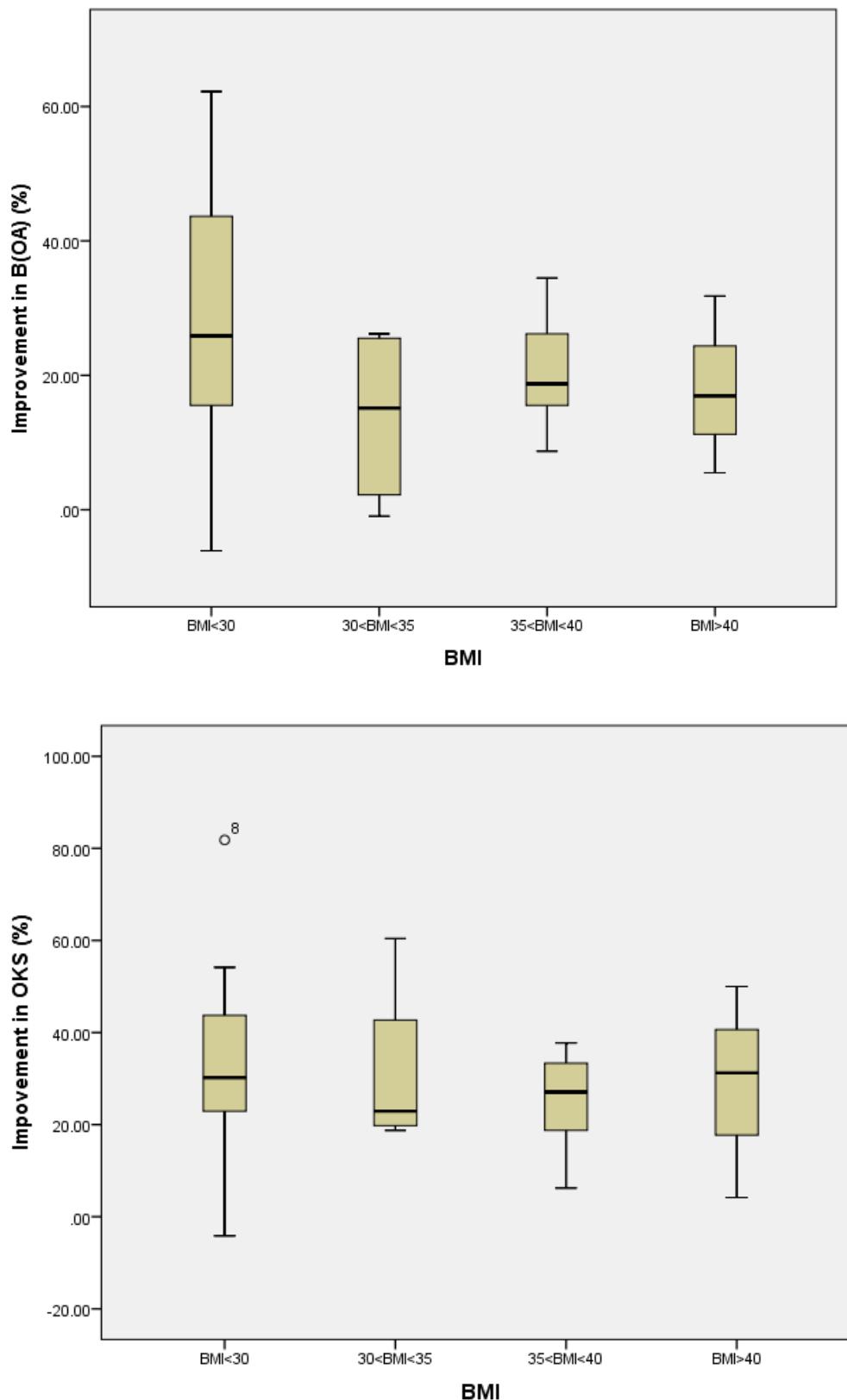


Figure 5.18 Stem leaf diagrams of the level of A) the post-operative reduction in B(OA), and B) The improvement in OKS relative to typical BMI category thresholds.

Neither the pre-operative OKS nor the B(OA) score of either the operative or non-operative leg appeared to have a significant linear effect on post-operative improvement. This differs from the findings of previous researchers using the DST classifier to discriminate osteoarthritic function; who found that poor function before surgery resulted in worse functional outcomes following TKR surgery (Watling, 2014, Worsley, 2011, Metcalfe, 2014).

Temporal-spatial

None of the temporal-spatial parameters considered within the analysis were significantly linearly correlated to improvement in B(OA) or OKS. In contrast to preliminary findings discussed in Section 5.2, neither pre-operative cadence or stance percent were indicators as potential predictors of recovery. The percentage of time spent in stance phase did almost reach significance, with a p value of 0.08, therefore there might not be sufficient statistical power to estimate this effect.

Kinematic

In keeping with preliminary findings, neither pre-operative sagittal knee ROM, knee flexion during loading response, nor flexion/extension towards terminal stance and toe-off appear to be significantly linearly correlated to post-operative improvement. Surprisingly, in contrast with preliminary findings, peak knee flexion during swing phase appears to have no effect on post-operative recovery. A new measure which wasn't previously included was the timing of the peak knee flexion, which indicated that subjects who reach peak knee flexion at a later stage might recover less following TKR surgery.

Kinetic

None of the kinetic measures included within the analysis were found to be significantly linearly correlated to improvement in B(OA) or of the OKS score. The closest result to reaching statistical significance was the dip ratio of the vertical GRF. The level of "double dip" within the vertical GRF is influenced by the vertical movement of the COM of the subject during gait.

In contrast to previous findings, no linear correlation was observed between the peak of the anterior GRF at push-off pre-operatively, and the post-operative improvement. This is in keeping with the finding in this cohort that the magnitude of peak knee flexion during swing, knee extension at terminal stance, and also cadence pre-operatively, had little or no impact on outcome. These are all gait variables that are largely affected by walking speed and therefore do not support preliminary findings that biomechanical attributes associated with slow gait might be predictive of poor post-operative outcomes.

Further Discussions

In theory, the core goal of walking is to progress the COM of the body within the sagittal plane. As such, much of the loading and motion of our joints also occurs within this plane. Preliminary findings suggested that biomechanical variables are associated with slower progression of the COM in sagittal plane. This wasn't, however, observed within the cohort presented in this chapter.

The variable identified as most significantly correlated with changes in B(OA) within this chapter was the coronal plane ankle angle at heel strike. The ankle angle was defined by the angle of the shank relative to a virtual foot segment (see Section 4.2.4). It is therefore assumed that the foot segment is at zero inversion/eversion at HS, and any coronal plane angle is measured from that reference position. Figure 5.19 shows the change in B(OA) against the ankle eversion angle at HS pre-operatively.

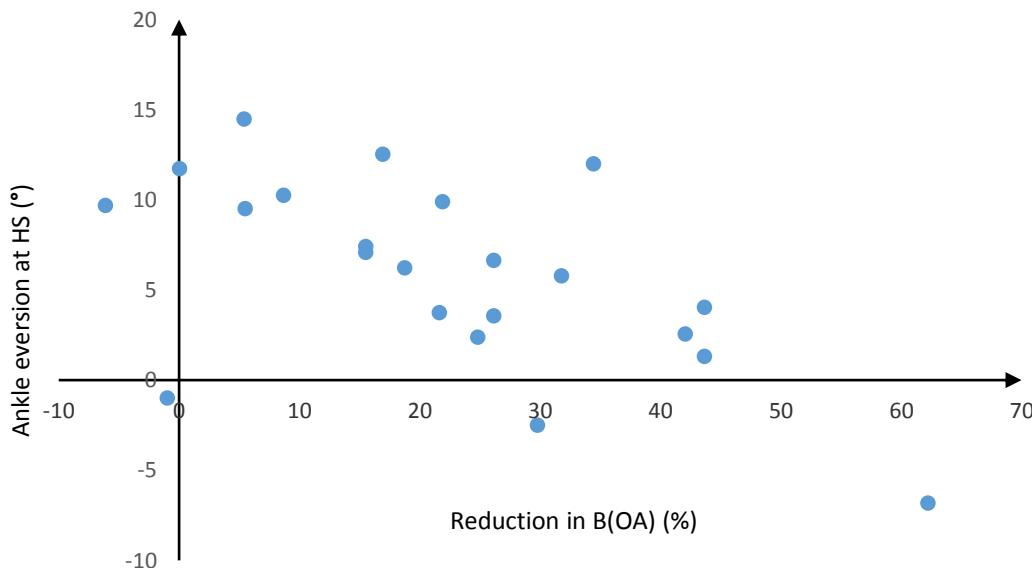


Figure 5.19 The relationship between ankle eversion at HS pre-operatively, and reduction in B(OA) following surgery.

Although the ankle eversion at HS was approximately normally distributed (Shapiro-Wilk test), subject 8 may well be having a large effect on the strength of the linear correlation found within this variable. Subject 8 was not detected as an outlier. A re-run of the statistical analysis without this subject results in a much smaller correlation of $r= 0.43$ and $p=0.052$. In fact, with subject 8 removed from the analysis, there are no variables which show a statistically significant linear correlation with the improvement in B(OA). The interpretation of the relationships does remain similar, but the effect sizes are not large enough to reach significance with a cohort of this size.

5.5 Conclusions

Aim 1: Do pre, post-, and the relative change in subjective outcome measures correlate with changes in biomechanical gait classification?

The trained DST classifier defined in the previous chapter was used to objectively summarise the gait biomechanics of 22 subjects pre-and post-TKR surgery. The changes in the degree of belief of osteoarthritic function B(OA) was used to objectively define post-operative improvements in biomechanical gait function towards that of a healthy individual.

It was hypothesised that the objective B(OA) would correlate well with PROMs before surgery. However, considering functional outcomes appear to be over-reported in TKR patients (Mizner *et al.*, 2011), it was expected that B(OA) would correlate less to PROMs following TKR surgery. Within this study, the opposite appeared true. The OKS, which is the primary outcome measure of TKR surgery in the UK (Clement *et al.*, 2013), was found to correlate pre and post-operatively to B(OA), as was the KOS.

Post-operatively, nearly all PROMs included in this study correlated with the post-operative changes in B(OA). This was a contrary to the study hypothesis and it appears that the objective quantification of gait function defined within this study was linearly correlated to patient-reported knee function. It is difficult to say whether improvement in function was generally under or over-reported, due to the nature of the measures themselves. The scale of each of the outcome measures is difficult to compare: for example, a subject recovering from 80% B(OA) to 40% B(OA) can't be directly compared to a subject whose PROM score has gone from 80% to 40%.

Aim 2: Does functional recovery return following TKR?

This study found overall biomechanical improvements in the operative limb of 19 of 22 subjects, and in the non-operative limb of 16 of 22 subjects. The post-operative reduction in B(OA) on the operative limb was strongly correlated to that of the contralateral limb. It was hypothesised in Section 5.1 that there would be a clearly distinguishable group of

improvers, and non-improvers following surgery. In contrast to the preliminary findings presented in Section 5.2. There was no clear and discernible boundary of improvement to separate the subjects into these two categories.

Aim 3: Does severity of pre-operative gait abnormality predict improvement in gait following surgery? Are there other biomechanical clinical factors which might predict improvement.

An analysis was performed to explore potential predictors of post-operative improvement in both B(OA) and OKS measures. This included several measures and only analysed for statistically significant linear correlators. Preliminary findings of gait velocity, AP force, and knee flexion during swing, as predictors of outcome; found within a further analysis of data from Watling (2014), were not observed within this study. Instead, ankle eversion and knee adduction at HS, ankle ROM, and timing of the peak knee flexion angle during swing, were found to be linearly correlated. A large number (29) of variables were examined in this exploratory analysis, therefore these results should be interpreted with caution. Furthermore, statistical significance was dependant on the inclusion of subject 8. Although subject 8 was not identified as an outlier, this does increase the likelihood that these findings are false positive findings.

This chapter has delved much deeper into not only the changes in overall B(OA) pre and post-operatively, but also how the individual input variables have contributed to those classifications at both time points. This has allowed further insight as to which input biomechanical variables changed following surgery, for both the whole cohort and for an individual subject.

5.6 Clinical Summary

Satisfaction following TKR surgery is comparatively poor, and objective performance-based measures might elucidate and help predict the level of functional recovery following surgery. The previous chapters build the foundations of a classification of biomechanical features which accurately discriminate OA gait. This chapter applies this classification to the biomechanics of 22 subjects before and around 12 months following TKR surgery.

Biomechanical recovery following TKR surgery was variable in this cohort, with no clear improvers/improvers. This is in contrast to previous studies considering HMA data (Metcalfe et al., 2017, Watling, 2014).

Biomechanical recovery, and improvement in OKS score following TKR surgery didn't appear significantly affected by pre-operative BMI within this cohort. This supports findings by (Rodriguez-Merchan, 2014), suggesting short-term functional outcomes appear similar within these cohorts.

Pre-operative functional status of the operative, or the contralateral limb, didn't predict recovery of either the OKS score, or improvement in biomechanics following surgery. There is evidence to suggest one of the biggest predictors of outcome is pre-operative function (Lingard et al., 2004), however this chapter has focused specifically post-operative change, or improvement, as opposed to considering the post-operative status alone. This might explain why this effect was not observed. It might also be that the cohort within this study was not heterogenous enough in pre-operative functional status to observe a significant effect, perhaps due to the inclusion criteria adopted clinically for consideration of TKR surgery.

PROMs might be more reflective of biomechanical function post-operatively. There is growing evidence that measuring functional changes following TKR requires both patient-reported and objectively assessed measures (Mizner et al., 2011, Naili et al., 2016). Findings from this study suggest that biomechanical joint function was less

strongly correlated PROMs, particularly the PACs pain and KOOS scores, pre-operatively. The sole use of patient-reported measures changes in functional outcome therefore might not reflect objectively assessed changes. Some studies have specifically found that patients report their improvements in physical function to be higher than it seems during objective assessments (Stratford and Kennedy, 2006, Worsley, 2011, Naili *et al.*, 2016). One interpretation of the results of the current study is that function might be under-estimated before surgery using PROMs, and hence contributed to over-reported function gains.

Ankle alignment might be important in predicting post-operative outcome. This study found the strongest predictor of outcome following TKR surgery was the ankle eversion angle at heel-strike. It is well known that varus/valgus alignment of the tibial component during TKR surgery might lead to changes in load distribution, increase likelihood of implant loosening, and hence decrease implant survivorship (Werner *et al.*, 2005). While the effect of knee alignment on TKR outcomes is well researched, there is much less evidence in the literature regarding alignment of the ankle. One study found highlighted that planovalgus foot might be a predictor of poor outcomes following surgery, finding posterior tibial tendon insufficiency within 12 of 48 revision cases (Meding *et al.*, 2005). Interestingly, posterior tibial tendon insufficiency appears to result in rearfoot eversion throughout the stance phase of gait (Tome *et al.*, 2006). The cohort within this study was small (22 subjects), however there may be reason to believe that either pre-operative ankle alignment itself, or the underlying mechanisms causing the malalignment, might affect post-operative outcomes.

Chapter 6 - Discussions

The research presented within this thesis will be further discussed within this chapter, and referred to the original objectives listed in Section 1.2.

6.1 Objective 1: Assess the validity and robustness of Jones' application of PCA dimensionality reduction and DST classification in characterising OA gait.

Within Chapter 3 it was introduced that since the DST classification work of Jones (2004), data collection has continued in a way which was compatible with these original methods. This provides an opportunity to test and validate the proposed techniques with a much-expanded cohort. For several reasons introduced within this chapter, it was deemed appropriate to expand upon pre-existing MATLAB code in order to process knee kinematics, and to include joint moment calculations. Part of this involved the addition of scripts which facilitated batch processing of biomechanical data using a central subject database. This provided further practical advantages in the ability to automatically detect the correct analogue channels and force plate calibration to use, fast reprocessing when updating biomechanical scripts and adding knee kinetic calculations, and the fast collation of data into a single file in preparation for waveform feature extraction. The updated and modified code was used to increase the cohort of subjects when using Jones' original input variables from 20 OA and 22 NP subjects to 85 OA and 38 NP subjects.

6.1.1 Dimensionality Reduction

Throughout this PhD thesis, PCA has been used as a dimension reduction technique for temporal biomechanical data. Previous research has demonstrated that this technique is an efficient way of objectively describing differences between OA and NP gait biomechanics (Jones *et al.*, 2006, Deluzio and Astephen, 2007, Landry *et al.*, 2007). When using PCA for data reduction, it is necessary for the researcher to decide how many PCs to retain. Three methods are discussed in this thesis:

1. **Kaiser's rule** - Previous researchers have used Kaiser's rule as an initial retention method, in which PCs which describe less than 1% of the total variance are excluded. This technique often retains a vast amount of PCs, with many having little significance (Jackson, 1993, Ferré, 1995). The first three PCs were initially considered within Section 4.3.3, and of those, only the first two PCs were ever considered for further analysis. Of those 18 PCs considered further, 16 of them described at least 10% of variance, and the remaining two still described over 5% variance. While the adoption of Kaiser's rule was not directly implemented and compared within this study, the findings suggest that PCs which describe between 1-5% variance would have been unlikely to have clinical importance in describe OA biomechanics. Kaiser's rule may therefore be too conservative a technique when applying PCA to temporal joint biomechanics obtained from HMA.
2. **Factor loadings rule** - Another technique which has been used, is to consider only PCs which have a factor loading of above 0.71 or below -0.71 at some point of the gait cycle (see Section 3.4.8). It was discussed that this factor loading is essentially the correlation between the original data points and the reconstructed points for each percentage of the gait cycle. This threshold might be therefore be better explained as the $r^2 > 0.5$, where at least 50% of the variance is reconstructed by that component. This thesis has demonstrated how using this threshold can result in potentially important PCs being discarded within biomechanical data. Within the top 18 variables within Section 4.3.4 which proved to accurately discriminate OA gait, four of them were PCs which otherwise wouldn't have been considered for analysis. One of them, PC2 of the hip adduction angle, was ranked 6th with a LOO classification accuracy of 86.1%. This PC was discussed further in Section 5.4.1, and appears to distinguish between pelvic drop (Trendelenburg gait) and hip hiking during gait. This feature didn't represent greater than 50% of the variance at any point of the gait cycle

due to large magnitude offsets in the adduction angle throughout the gait cycle, which were reconstructed by PC1.

3. **Novel method** – A choice was made within Chapter 4 to initially include at least the first three PCs of each waveform, and then to also include any additional components which fulfilled the squared factor loading $r^2 > 0.5$ rule. This rule is discussed further in Section 4.2.8. This rule was found to be initially more conservative than Kaiser's rule, however retained important and clinically meaningful features which were previously discarded by the factor loadings method. This technique resulted in 70 PCs being retained from 69 variables, with a 4th PC being retained through the additional factor loading rule on only one occasion. These PCs were then further reduced by considering their classification ranking using two datasets (each containing half the total data). Of the 18 PCs retained, only the top two PCs were ever considered. This latter rule only ever considered the first and second PCs, suggesting that the third PCs often either contained information on biomechanical changes which were only present in a select number of the OA subject, or, more likely, contained information of small magnitudes of variation and of little clinical significance.

Future studies are therefore strongly advised to use the factor loadings rule as an initial PCA retention techniques as it may discard important biomechanical features, and that in many cases only the first two PCs contain relevant information regarding biomechanical differences between OA and NP subjects. Previous studies have also considered retaining PCs by assigning a threshold of variance which must be described by each waveform i.e. select however many variables required to represent x% variance for each waveform (Deluzio and Astephen, 2007). Following this technique, discriminant analysis was then used to investigate the most distinguishing biomechanical features. This applies a methodology as the novel method used in this study, where the initial rule was less conservative, and then an additional multivariate analysis technique is used to further rank and retain components.

Alongside discarding important gait features, the use of factor loadings alone to interpret and contextualise PCs can be misleading. An example of this are the PCs of the knee flexion/extension angle displayed within Section 3.6.2 . The flexion/extension waveforms and their reconstructions have been replotted in Figure 6.1 (see Figure 3.30 for the original). When interpreting waveforms considering only the factor loadings, an approach which has been adopted in previous studies (Jones, 2004, Whatling, 2009, Metcalfe, 2014, Watling, 2014), PC1 would have been interpreted as representing variance at 0-57% of the gait cycle, and PC2 would have been interpreted as representing between 59-64% of the cycle. Within the region represented by PC1, OA subjects appear to have reduced peak knee flexion at midstance, and reduced peak extension during terminal stance – i.e. reduced ROM over stance phase. Within the region represented by PC2, OA subjects appear to have delayed and reduced peak knee flexion. Within the final

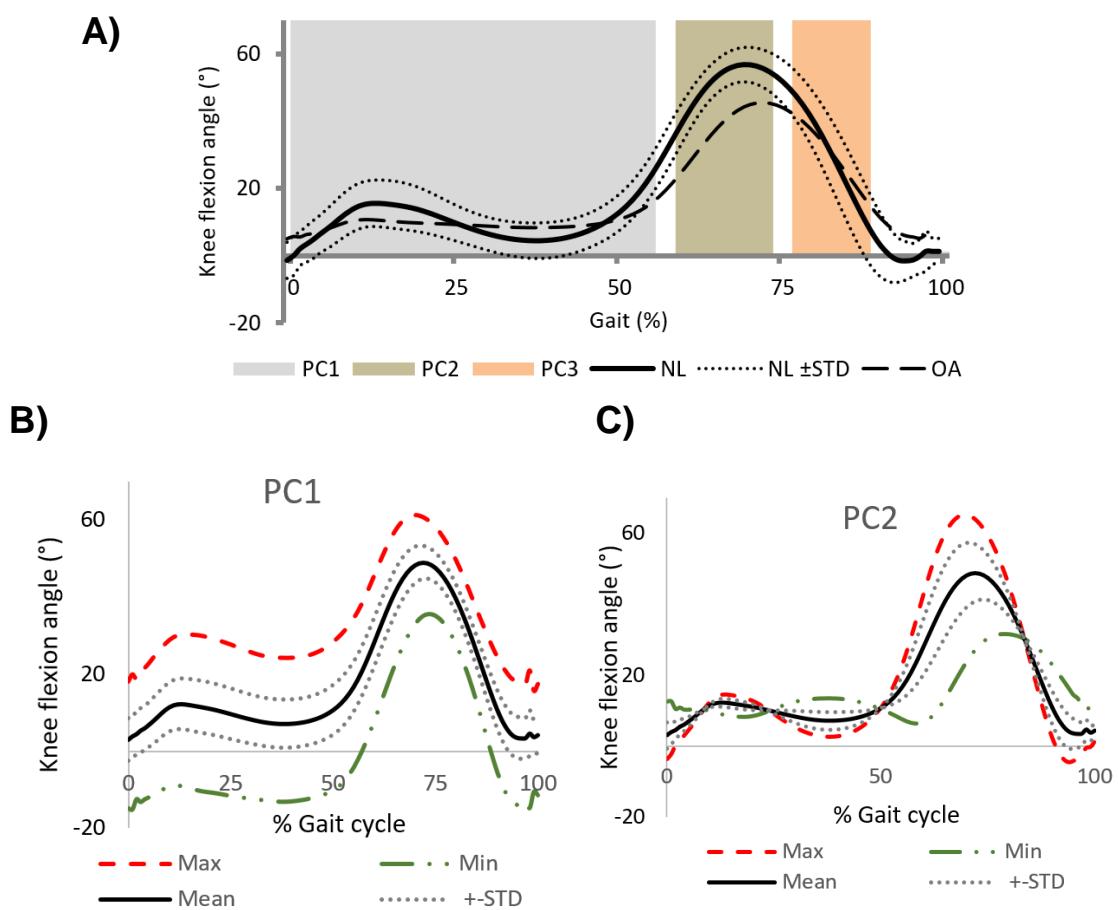


Figure 6.1 The average knee flexion extension waveforms for NP and OA subjects and the regions where retained PCs represent >50% of the variance. B) The reconstruction of these waveforms using PC1. C) the reconstruction using PC2.

classification, PC2 was shown to classify much more accurately (81.6%) than PC1 (52.6%). The interpretation would therefore be that reduced peak knee flexion is much more indicative of OA gait than reduced ROM during stance phase. If, however, we consider the reconstructed waveforms using each PC, shown in Figure 6.1B&C, it is PC2, not PC1, which reconstructs reduced ROM during stance phase. The corrected interpretation is therefore quite the opposite, the reduced ROM during stance is indicative of OA, and is highly correlated with reduced peak knee flexion.

6.1.2 Challenges in PC Reconstruction:

There have been a few instances in which the interpretation of reconstructed PCs has been a challenge. Level gait biomechanics are generally considered to be periodic, hence temporal alignment can be achieved by normalising between gait events. When considering gait kinematics, it is very common to normalise as a percentage of the whole gait cycle, from initial contact to the following foot contact of the same leg. While this is a common technique, functionally important gait events within a gait cycle often vary within this gait cycle (Helwig *et al.*, 2011), which has led to alternative techniques of time normalising such as curve regression (Sadeghi *et al.*, 2000). Figure 6.2 shows the ensemble average flexion/extension waveform for a single subject and the reconstruction of that waveform using the first two PCs displayed in Figure 6.1. Despite overestimating the peak knee flexion during stance and swing, the reconstruction is reasonably faithful to the original waveform. Within Figure 6.2B, the same waveform was manipulated such that toe-off occurred 5% earlier, and an additional 5% of the gait cycle was therefore spent in swing phase. Notice how the reconstruction of this waveform using the first two PCs is much less faithful to the original curve and how, in fact, the reconstruction accounts for very little of the temporal shift. To summarise, it appears that, when a waveform is not perfectly periodic, faithfulness of PCA interpretation and the subsequent interpretation becomes less reliable. Magnitude differences in the data

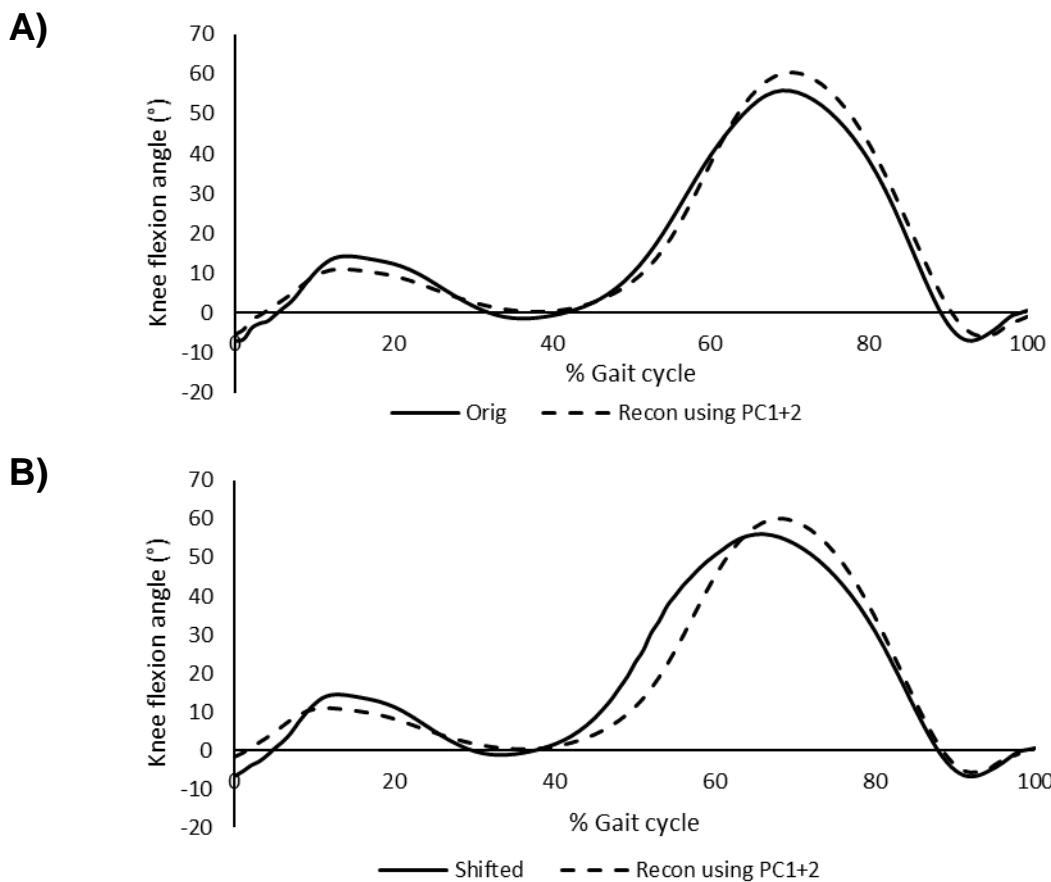


Figure 6.2 Illustration of the effect of temporal shifts has on the resultant PC reconstruction.

Figures **A)** Displays the original flexion/extension angle for a single subject (bold), and the reconstruction of that angle using PC1 and PC2 (dashed). Within **B)** the waveform has been shifted and interpolated such that toe-off occurs 5% of earlier within the gait cycle. The reconstruction of this shifted waveform is noticeably less faithful to the original curve.

(shifts in the y-axis) therefore appear well reconstructed; however temporal shifts within the data (shifts in the x-axis) seem more challenging to reconstruct.

6.1.3 DST Classification Control Parameters

Before using this technique to further test and validate the DST classification method, two theoretical disadvantages with the control parameters recommended by Beynon (2002), and also used by a number of authors adopting these techniques (Jones, 2004, Whatling, 2009, Worsley, 2011, Watling, 2014, Metcalfe, 2014), were highlighted. The first of these is the definition of θ , the variable value at which $cf(\theta) = 0.5$. This value was contextually compared to the sigmoid input bias of an ANN; however unlike within an ANN this value is not iteratively optimised and is instead explicitly defined. Two new

methods of defining θ , referred to as θ_A and θ_S , are proposed and tested throughout this chapter. Within the tested data sets it is apparent that both new methods result in large improvements in LOO classification in comparison to the original definition, with improvements in classification accuracies as large as 7.3% (9 less misclassified subjects). The definition θ_S , which is adjusted to account for difference in variance between the two data groups, resulted in a more modest improvement over θ_A .

The second theoretical disadvantage was that of the definition of k , which defines the steepness of the sigmoid activation function. It was first highlighted that this value had a similar function to the inputs weights of an ANN, again being defined explicitly as opposed to being optimised iteratively. The use of the correlation coefficient k_c to define the steepness results, as recommended by Jones (2004), results in a bias such that input variables of a larger scale, e.g. height in millimetres, result in larger belief functions than equivalent input variables of smaller scale e.g. height in metres. This is because the classifier input variables aren't typically standardised. It was shown that the use of the other definition tested by Jones, k_s , which defines the steepness relative to the standard deviation of the input variable, corrects for this. This, however, merely acts as standardisation, and doesn't weight input variables based on their ability to discriminate between the two groups. A new method, $k_{C/S}$, was proposed and tested, and found to result in modest improvements in classification accuracy. The relatively small improvements in the performance of $k_{C/S}$ over k_c is likely because the input variables were derived from PCA. The adopted methodology for calculating PCs involves the calculation of z-scores. While the standardisation of data within PCA doesn't necessarily result in PC scores of equivalent scales, anecdotally the scales of selected PC scores did appear similar.

6.1.4 Robustness of Classification

Data classification techniques are susceptible to over-fitting, in which performance on training data might not well reflect performance on unseen data (see Section 3.5.4). Previous use of the DST classification method within this research group has adopted

the use of the LOO classification technique, which optimises the size of the training cohort by only removing one subject at a time from the training body, treating this one subject as “unseen data”, and then attempting to classify them. This study has compared the LOO technique with leave-p-out techniques; with different values of p being considered. In summary, the LOO technique appeared to be no less conservative in its accuracy estimations even when p was as large as 20, in a cohort of 85OA and 38 NP subjects.

Within Section 3.6.5 it was also demonstrated that LOO classification accuracy increased quite dramatically from cohort sizes of up to ten subjects in each group (20 total), and then began to plateau particularly towards cohorts of 30 within each group. This may be useful in future studies, as well as adding confidence that the training bodies used within this thesis were large enough to model OA biomechanics.

The objective reduction of input variables based on their discriminatory power within the training body was demonstrated to bias the classification performance. Therefore, within Chapter 4 the training body was split into two halves, the input variables were ranked, and then the ranked variables were compared. The top 18 ranked variables were consistent between the two halves of the training body, and hence were selected for further analysis. This reduces the risk of classification bias in the selection of input variables to retain in the classification training body.

6.2 Objective 2: Determine the biomechanical changes in the ankle, knee and hip and due to late-stage osteoarthritis using the methods developed in Objective 1.

6.2.1 Ground Reaction Forces

The GRF repeatedly demonstrated to be remarkably reliable at distinguishing between OA and NP gait. Within each classification, the vertical and the AP GRF were the two top-ranked input variables. The feature consistently identified from the vertical force is that of a slower rate of load acceptance and offloading, and a reduced “double peak” of reaction force.

The features consistently identified within the AP force were the magnitude of the posterior force in the first half of stance, and the anterior force in the second half of stance phase. The magnitude of this braking and propulsive forces are very much related to the speed at which the subject is walking. If we take Newton's second law, $F=ma$, and consider that standing limb must decelerate after contact and accelerate at push-off, apparent that faster walking requires larger accelerations and hence increased force. Walking speed is a confounding variable which challenges interpretation of the results within this study. The reduced gait velocity could be in avoidance of pain or instability associated with increased braking or propulsion during gait.

The mediolateral GRF wasn't initially included by Jones (2004) and was added to the classification training body in Section 3.6.6. While not ranking as highly as AP and vertical forces, the second PC of the mediolateral force also proved reliable in distinguishing OA gait. In both instances, the second PC of the ML force appeared to also reconstruct the level of double peak of the reaction force, as opposed to PC1 which constructed larger magnitude differences.

6.2.2 Knee Kinematics

The knee kinematic which appeared to best discriminate osteoarthritic function was the knee flexion angle. This is perhaps not surprising as the goal of gait is to move the COM within the sagittal plane, and hence both force and movement tend to be greater within this direction. Because of the increased magnitude of sagittal angles necessary for gait, both intra-subject variability and inter-subject differences tend to also be of larger magnitudes. STA and marker placement error are well established as the most significant sources of error during HMA. These appear to result in the smallest magnitude of error within knee flexion extension angle, and the largest within the internal/external rotation angle (Leardini *et al.*, 2005) (Della Croce *et al.*, 2005). This is reflected within the current study, which found that knee coronal and transverse angles were poorly ranked within the classification, and often the first principal component of the transverse

angles represented a large offset throughout the waveform which might indicate errors in the definition of the joint axis.

The specific feature of the knee flexion waveform that well distinguished OA was the one which represented a reduction in ROM during the stance phase of gait, associated with reduced peak knee flexion during swing phase. The peak knee flexion also tended to occur later, and was associated with slower extension during terminal swing. This feature was consistently represented within the second PC, with the component describing most variance reconstructing differences in magnitude of flexion during stance phase as opposed to the ROM. This first PC appears to be much less associated with pathological gait.

6.2.3 Knee Kinetics

The two PCs which have been used to calculate joint moments have found differing results as to the most important knee joint moment in distinguishing OA. The first technique used within Chapter 3 found PC1 of the flexion/extension moment to best discriminate OA function, which reconstructed a reduced external flexion moment during the first half of stance and reduced extension moment within the second. When using the larger cohort and methodology of Chapter 4 however, a very similar feature was identified as PC2 of the flexion moment and was only the second highest ranking knee joint moment. The highest ranked was that of PC2 of the knee adduction moment, which reconstructed the reduction of the “double peak” of the knee adduction moment, with a consequently larger moment at midstance and a slower rate of offloading towards toe-off. The EKAM was introduced within Section 2.1 as a proposed surrogate measure for medial contact forces, a biomechanical factor related to radiographic alignment, and has often been reported to increase in more severe OA (Foroughi *et al.*, 2009). It is interesting that the feature which appeared to very consistently distinguish OA gait was not related to magnitude of the moment, but instead to level of “double peak”.

6.2.4 Hip Kinematics

At the hip, kinematic features within the coronal plan were better at discerning OA gait biomechanics than the other two planes. The feature which discerned OA gait was the second PC if the hip adduction angle, which characterises a relatively more adducted hip during the first half of stance, and a relatively more abducted hip during swing phase. It was proposed that this could be a sign that there was pelvis/trunk lean towards the side of the leg in stance, which might be a sign of “hip hiking” to increase ground clearance.

6.2.5 Hip Kinetics

Similarly, the coronal plane moment, the hip adduction moment, was the kinetic feature of the hip which best distinguished OA gait. The second PC of the hip adduction moment was ranked 4th with 87.5% LOO classification accuracy, and again reconstructed the level of “double peak” of the hip adduction moment.

6.2.6 Ankle Kinematics

The kinematics of the ankle appeared less powerful at discerning OA gait. In fact, only one kinematic feature of the ankle was included in the classification defined within Chapter 4 – the second PC of the ankle plantar/dorsiflexion angle. This PC was ranked 17th with a LOO classification of 76.3% and reconstructed a delayed ankle dorsiflexion during stance phase, a late initiation of plantarflexion at midstance, and a reduced peak plantarflexion at toe-off. From a slightly plantarflexed position after foot contact, the ankle would normally begin to dorsiflex as the knee progresses in the anterior direction; however, with OA subjects, this appears slightly delayed. In NP subjects, at around 50% of the gait cycle the heel will often start to rise, which initiates increased plantarflexion towards toe-off - OA subjects in this study tended to have a delayed initiation of this plantarflexion, perhaps indicating late heel rise. Furthermore, NP subjects appeared to reach a peak plantarflexion angle of 5-15 degrees around toe-off; however, this appeared reduced in OA subjects. This might indicate that OA subjects have their knee in a more anterior position at toe-off, hence moving the ankle in to a relatively more dorsiflexed

position, or perhaps that the angle of incidence of between the foot and the floor was less at heel strike. If the latter is true, it might be that the COP at toe-off was relatively more posterior in the OA subjects than in the NP subjects, which should be explored in more detail in future work.

6.2.7 Ankle Kinetics

Of the three ankle-joint moments, the most distinguishing factor of OA gait was PC2 of the ankle plantarflexion moment. It has already been established that OA subjects had a reduced “push-off” force towards toe-off; however, the PC which reconstructed large magnitude differences of the ankle plantarflexion moment peak, PC1, was poorly ranked within initial classifications (31st) and was therefore not retained for further analysis. The second PC reconstructed an increased plantarflexion moment throughout the first half of stance, which was associated with a reduced peak plantarflexion moment throughout the second half of stance. Figure 6.3 displays alongside the mean \pm 1 STD of the plantarflexion moment during stance phase for the 41 OA and 31 NP subjects included within Chapter 4. It is noticeable that when treated as a single group, OA subjects don't appear to display an increased plantarflexion moment during the first half of stance phase.

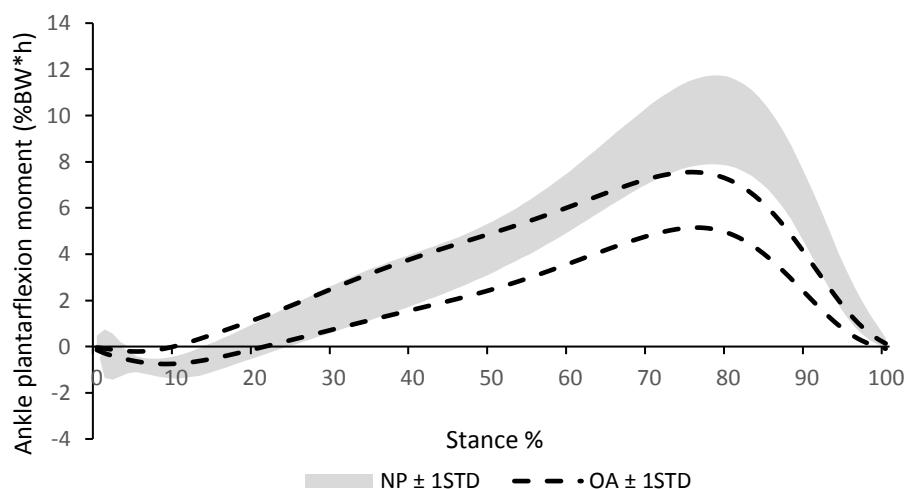


Figure 6.3 Displays the mean \pm 1 STD of the plantarflexion moment of the 31 NP and 41 OA participants of Chapter 4 during stance phase.

6.3 Objective 3: Objectively measure biomechanical changes following TKR surgery, and elucidate the relationship between pre and post-operative gait biomechanics, and patient-reported outcome.

Overall there were very clear signs of functional improvement of the operative leg within the majority of subjects following TKR surgery. Watling (2014) used the DST classifier to measure changes in function of 12 TKR subjects, finding that five of these subjects (41.7%) received little to no functional benefit following TKR. Watling also identified two additional groups – mixed recovery and poor recovery. Within this small cohort, there appeared to be some separation between the poor recovery group and the other two groups. Recovery within the poor outcome group ranged between an increase of 1% B(OA), to a decrease of 12% B(OA), while the worst of the “mixed” recovery group had reduced by 0.28 B(OA). Within this study however there appears to be much less of a divide between good and poor recovery. Had the same threshold been considered, a reduction of 12% B(OA), this study would have identified six of the 22 subjects who had poor functional improvement (27%).

Metcalfe also identified two clusters of improvement using similar methodology in 14 unilateral knee OA subjects undergoing TKR surgery, with seven subjects being classified as predominantly healthy post-operatively. High pre-operative B(OA), increased age, and decreased pre-operative walking speed all appeared to be highly correlated to functional improvement following surgery.

This study has explored the classification results in greater depth than previous studies adopting the DST classification method. More specifically, the contribution of individual bodies of evidence to the belief values have been analysed pre and post-surgery, in order to determine which biomechanical features have been altered most by TKR surgery. In summary, kinematic changes were most apparent in the sagittal plane of the hip and knee; however, there also appeared to be apparent changes in the transverse hip angle. In terms of kinetics, changes were most apparent in the coronal plane of the hip and knee, and in the sagittal plane of the hip, knee and ankle. Changes in the GRF were largest in the vertical, AP, and ML forces respectively. Statistically significant

changes were also found in the biomechanical features following surgery. Considering the $p<0.01$ threshold, eight out of the 18 input variables resulted in significant improvements (changes which brought the PC scores closer to that of the NP control group).

6.3.1 Comparison With PROMs

Within this study, linear correlations between B(OA) and PROMs were explored. Using the correlation interpretation of Evans (1996), this study found moderate correlations between B(OA) with OKS and KOS pre-operatively, a strong correlation with OKS post-operatively, and moderate correlations with all PROMs post-operatively. The previous research of Watling (2014) found B(OA) to strongly correlate with OKS and KOS scores, however considered pre- and post-operative subjects within the same analysis. Had this study used the same approach as Watling, strong correlations would have been found within every outcome measure, with the highest being with KOS ($r=0.712$) and KOS ($r=0.710$). An issue with taking this approach shall be outlined using Figure 6.4. Consider two separate variables, outcome measure 1 and 2, which quantity very different aspects of function. When pre and post-operative results are considered separately, it is clear

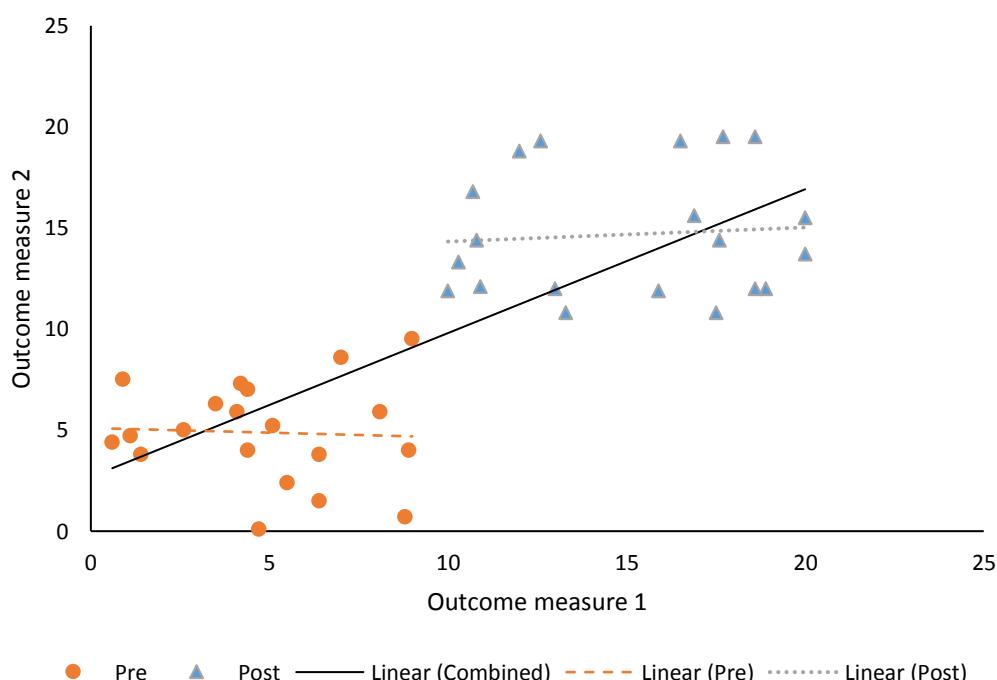


Figure 6.4 Illustrative example of why different outcome measures shouldn't be correlated both pre and post-operatively within a single analysis.

from the linear regression that there is very little relationship between the two measurements. Now let's consider that both the measures aspects of function improve following surgery. Considering both pre and post-operative measures within a single analysis, it now appears as if there is a strong correlation between the two outcome measures ($r=0.754$).

6.4 Contributions to Knowledge

The section will briefly summarise the primary novel contributions to knowledge made by this research. Within this thesis, novel developments were made to the application of PCA and a DST classifier to objectively quantify and monitor pathological changes in level gait biomechanics. The key contributions have been summarised and categorised as: methodological if they are focused on the genetic techniques of summarating temporal biomechanical information to objectively measure differences and changes in function, and clinical if they are focused on the interpretation of novel biomechanical findings using these techniques.

Methodological

- Alternative definitions of two control parameters within the DST, which were theoretically demonstrated to remove input bias, and when implemented significantly improved the accuracy of the classification.
- The robustness of the original DST gait classification presented by (Jones, 2004) was validated through the significant expansion of the cohort size and the cross-validation across different test set sizes.
- The effect of cohort size on classification accuracy was investigated, finding only 10 subjects in each group were required to achieve a stable classification accuracy of $92\% \pm 2.5\%$.
- In contrast to the approach taken previously by (Watling, 2014), theoretical advantages were highlighted for the inclusion of a broad age range when discriminating OA and NP gait biomechanics using a DST classifier.

- Changes in the individual BOE within DST classification has, for the first time, been compared to traditional statistical methods. The analysis found that the biomechanical features which were deemed to have improved most following TKR surgery using DST classification were detected as significantly different changes. This increases the confidence that the DST classification method is deriving intuitive and interpretable relationships between biomechanical inputs and individual 'beliefs of OA'.

Clinical

- Despite receiving little focus in the literature, transverse plane hip, knee and ankle moments have been found to be a very strong discriminator of late-stage OA gait. Furthermore, transverse moments did not appear to recover towards that of NP subjects following surgery. It is recommended that transverse plane moments are considered in future studies, and their implications on knee joint loading, control and function are explored.
- A PC of the hip adduction angle interpreted as hip hiking was a strong discriminator of late-stage OA gait, despite only accounting for 11% of the variance. From interpretation of the PC it is suggested changes in hip adduction relative to HS might be a more sensitive measure at detecting this biomechanical feature than the magnitude of the angle itself.
- This research suggests that OKS and KOS scores are correlated to gait function, but that patient-reported measures seem more strongly related following TKR surgery. Patient-reported improvements in function have previous been shown to be over-reported following TKR surgery (Stratford and Kennedy, 2006, Worsley, 2011, Naili *et al.*, 2016), this might be due to under-estimated function before surgery, perhaps due to pre-operative pain, outlook, or perhaps a measure of confidence or stability not accounted for in this biomechanical assessment.
- A more everted ankle at heel strike pre-operatively appeared to predict poorer outcome post-operatively. To the author's knowledge, the relationship between

foot angulation and biomechanical outcomes have not previously been reported. A more everted ankle during stance is suggestive of posterior tibial tendon deficiency and a planovalgus foot (Tome et al., 2006), which has previously been suggested to influence surgical outcome (Meding et al., 2005). These findings were a result of an explorative retrospective analysis and warrants further investigation.

Chapter 7 - Limitations

7.1 Variability

Within this analysis multiple walks were considered, and the ensemble average of these walks were taken. Within the methodology, it was aimed to get six clean gait cycles for each limb, with clear single-limb force plate strikes at heel strike and toe-off. In practice, when processing the data, one or more of these might turn out not to be suitable for processing. It might take several cycles to achieve clean force plate data, and therefore pathological patients will occasionally start to feel pain during the session and hence unable to complete all six walks. In these instances, the ensemble average is being calculated on a smaller sample size and therefore might be less representative of the subject's true average gait cycle.

While ensemble averaging can reduce measurement errors, the validity of this technique relies on a low level of variability in the movement itself. An example of a subject who had a high degree of variability in their knee flexion angle is shown in Figure 7.1A. It can be seen from this figure that through taking the ensemble average of each walk, a lot of information is lost. For example, the subject ranges from hyperextending during midstance, to flexing 10 degrees. There are also large differences in the timing and magnitude of peak knee flexion during the swing phase of gait. Figure 7.1B displays the knee flexion angles for three walks in which this was observed during data processing, which are labelled as Walk 2, Walk 6, and Walk 8. Walks one and walks four didn't contain suitable force plate data and therefore weren't included for analysis, however the original naming of the walks is preserved. It can be noticed that in this example the knee flexion angle during gait appears substantially different from the first to the last walk included in the analysis.

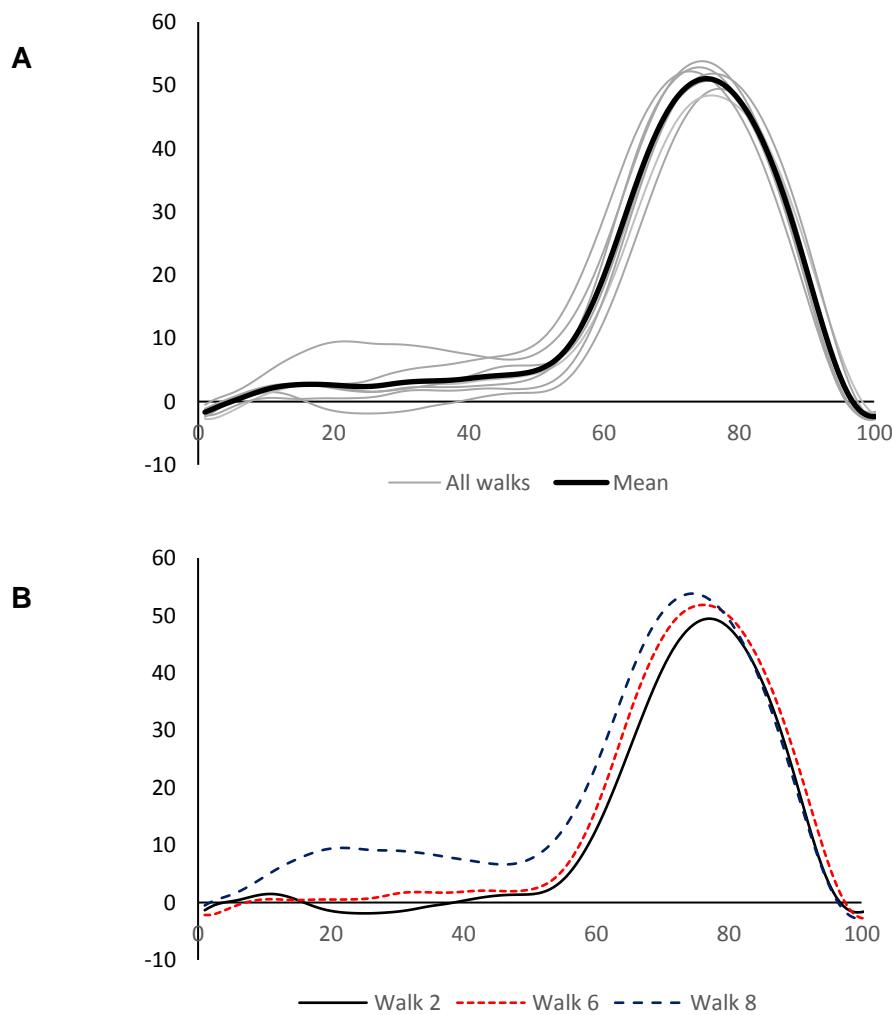


Figure 7.1 Variability in knee flexion waveforms for a single subject post-TKR surgery A) shows the all walks in grey, and the mean (ensemble average) in black. B) Displays only walks 2, 6 and 8.

A subject's movement might change during the course of a study for numerous reasons. Subjects might have been seated for a long time before the gait analysis initiated, often remaining seated while they fill out PROMs and for much of marker placement. Many of the OA subjects reported stiff knees after sitting, and it is therefore possible that their knees might have become less stiff over the course of the data collection. The rigid marker clusters used within this study were affixed to the subjects using cohesive bandage tape, and were fixated tight enough to reduce relative movement between the skin and the cluster while still being comfortable for the subject. While upmost effort is made to provide a comfortable environment for the subject, gait analysis can feel like a

very unnatural experience, and this is likely exacerbated by the inclusion of markers, rigid clusters, EMG electrodes etc.

Subjects were asked to walk along a 10m walkway at a self-selected pace, and dummy force plates were included within the floor to avoid force plate targeting by the subject. While it might not have been possible to distinguish between force plates and dummy plates, the flooring within the centre of the walkway was visually distinct. Figure 7.2 displays the example where ankle plantarflexion at the first foot contact upon the force platform was distinctly different to that of the following foot contact. Figure 7.2C shows that during the actual force platform contact, the ankle is in a much more plantarflexed position. In fact, it appears that as opposed to striking with the heel first, the toes come

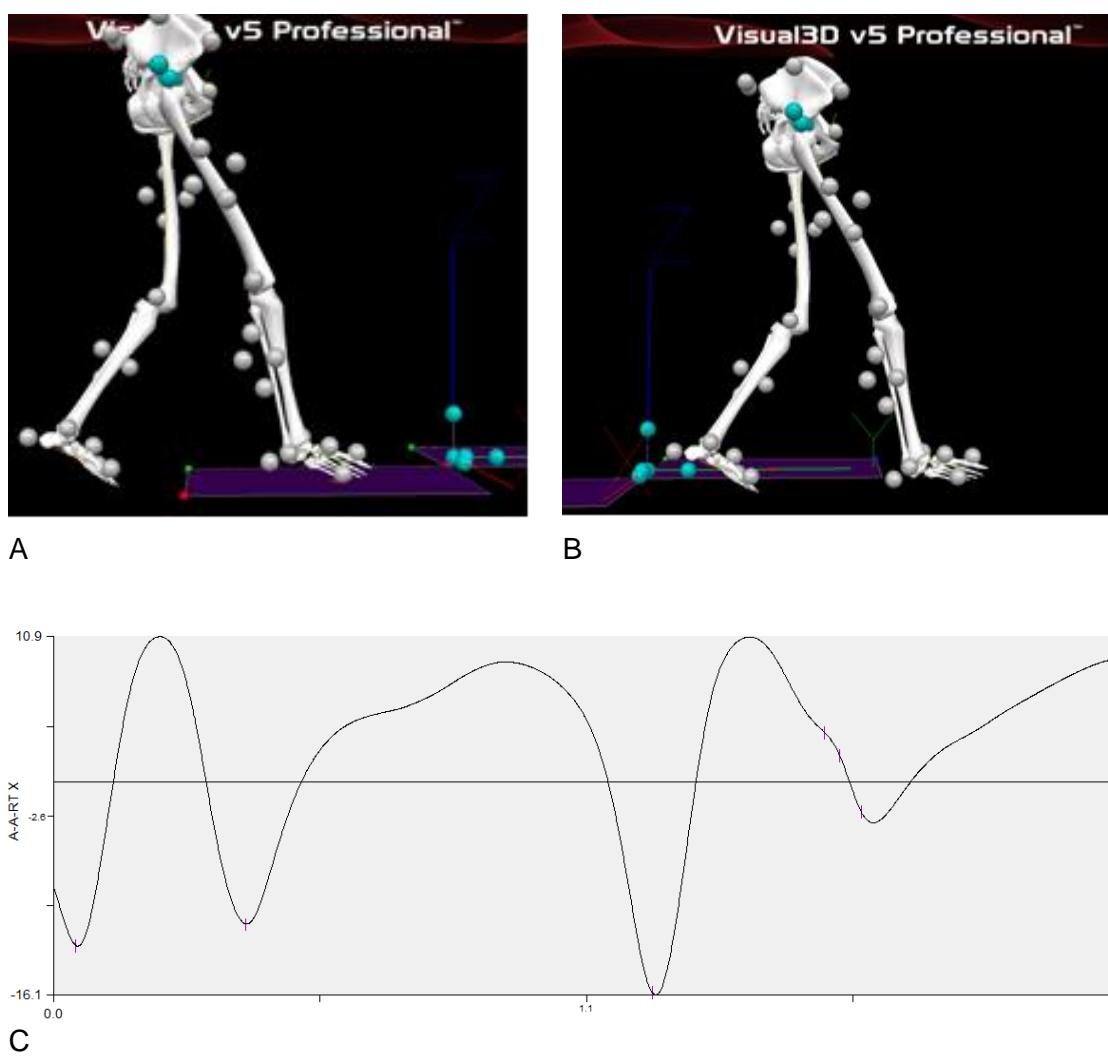


Figure 7.2 An example of suspected fear of force platforms. A) and B) show the model segment view during foot contact 1 & 2 respectively, and C) displays the dorsiflexion angle at these points. It can be seen that the subject is much more plantarflexed at the first (force platform) foot contract.

into contract first. This may well be due to the variability of the subject's movement itself, or a variable pain avoidance strategy. It could also be because of apprehension of foot contact, perhaps due to the subject not being confident stepping into the area in which the force platforms are hidden.

7.2 Patient Cohort

7.2.1 Heterogeneity

Knee OA is a bilateral disease (Metcalfe, 2014) which most commonly effects an elderly population. TKR surgery is recommended to late-stage osteoarthritic subjects who will likely have been suffering with knee pain, swelling, instability for several years. Gait biomechanics have been shown to increasingly change during OA disease progression, and hence the joints of both the affected and contralateral leg, as well as the spine, will have suffered from altered joint mechanics over a sustained period. The subjects within this study reported a diverse range of symptoms which might affect the way they move, such as contralateral knee pain or OA, contralateral arthroplasty, back pain, ankle swelling, injury or pain, hip pain or replacement. Information regarding physiotherapy and rehabilitation exercises was taken, but not analysed. It was however noticed that there were large variations in the rehabilitation offered, and levels of compliance to both clinic visits and home-exercises. On top of this, participants were treated by different surgeons, had different pre-operative KL grades, and received different implants. Even in cases where the implant design matches, alignment and sizing of the implant will vary across subjects. These clinical and surgical factors have been collected for several the subjects, however due to the difficult nature of accessing and recording this information the dataset is incomplete and hasn't been considered within the analysis. Of the surgical factors which have been collated, all had fixed-bearing knees, there seemed roughly an equal number of cruciate-retaining and posterior-stabilised, OA most often affected all three compartments, and patella surfacing was inconsistently carried out.

The heterogeneity of the cohort improves the generalisability of the results, i.e. they are they are not specific to TKR patients performed by a particular surgeon, using a specific

implant, rehabilitated by a particular clinician, however it is acknowledged the influence of these factors may have been influential in determining functional outcomes following TKR surgery.

7.2.2 Sample Bias

The sample of late-stage OA subjects considered within this study were those who had been recommended TKR surgery, and could walk at least 10m without a walking aid. The results therefore cannot be generalised to all patients with severe OA (KL grade 3-4), or even all those who undergo TKR surgery. The participation in the study was of course optional, and therefore there may have been sample bias in those who wished to volunteer. Anecdotally when attempting to recruiting volunteers, those who worked full time often noted they would have to take multiple days or half-days annual leave, and were typically less flexible when booking the motion analysis sessions.

7.3 Hardware Changes

It was noted that in Section 3.1 that there were hardware changes within the motion analysis laboratory during the course of the subject data collections included within this study. The resultant effect of upgrading the force platforms hasn't been quantified, however anecdotally there was no obvious changes in the GRF data, particularly following the low-pass Butterworth filtering. The change which most likely had the largest effect was the upgrade from ProReflex infrared cameras (Qualisys, Sweden), with a capture resolution of 680 x 500 pixels, to the Oqus 3 cameras (Qualisys, Sweden), which have a resolution of 1280 x 1024 pixels. Average residual marker trajectory errors can be calculated during initial calibration, and these errors were decreased from often being around 1.2mm, to being generally lower than 0.8mm with the updated cameras. It isn't known what effect this might have had on calculations of dynamic joint biomechanics, however it is acknowledged that these changes are much smaller than the errors induced through STA, which have been reported as high as 30mm on the thigh and 15mm on the tibia (Peters *et al.*, 2010).

7.4 Inter-operator Errors

It was discussed within Section 3.1 that the subjects used within this study were part of ongoing data collection over several years. Within this time, numerous researchers have helped to collect this data. Clear SOPs have continually been in place, and new researchers have undergone training and assessment before being permitted to help collect this data. It is however possible that this studies suffers from inter-operator variability, on top of the unavoidable intra-operator variability of motion analysis techniques. The study of Della Croce *et al.* (1999) identified intra and inter-operator errors when identifying anatomical landmarks in the range 6-21 mm and 13-25 mm, respectively.

7.5 Sensitivity and Specificity

The DST classifier was shown to accurately discriminate between healthy and late-stage osteoarthritic (pre-TKR) level gait biomechanics. The high classification accuracies generally found within this study indicate that the trained model was effective as estimating the likelihood of a subject belonging to either the NP or the OA cohort. Often, when data classification is used within a clinical context, sensitivity and specificity are highly important as they indicate the likelihood of false positives and false negative results. This can be particularly important when a test is being used as a diagnostic tool or aid. Within this study sensitive and specificity of the classification technique has not been reported, and the classification results do not reflect the accuracy that could be achieved in distinguishing OA gait biomechanics from other lower leg pathologies.

Chapter 8 Recommendations for future work

8.1 PCA Using Multiple Waveforms in One State Space.

Within this study, temporal data has been objectively reduced using PCA. This technique has been very successful in representing a large proportion of the total variance using only 1-4 PCs. The success of PCA in reconstructing level gait biomechanics infers that there are strong linear correlations within the waveforms. The application of PCA as described within this study, therefore, takes advantage of the inter-relation between, say, the value of knee flexion at 60% of gait, and its value at 70% of the gait cycle. These

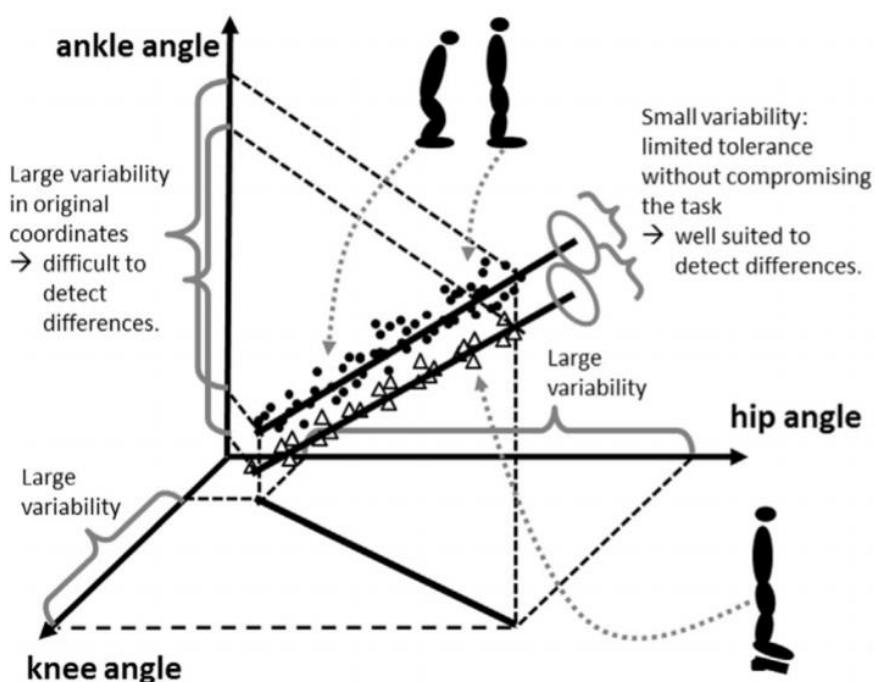


Figure 8.1 Schematic figure reprinted from Federolf et al. (2013). Variables from the hip, knee and ankle angle are plotted for subjects who have simply been asked to stand (solid circles). While each axis contains a high degree of variability, when plotted in three dimensions there is only limited variability about the PC (black line). Subjects then repeat the task standing on a shoe with an elevated heel (hollow triangle). It would have been difficult to distinguish differences considering each axis individually, however a new three-dimensional PC clearly differentiates the two groups.

interdependencies within a single waveform make it possible to distinguish biomechanical features using PCA.

While there are certainly many independencies *within* a biomechanical waveform during gait, there are also interdependencies *between* different biomechanical variables. This is well illustrated by the example of Federolf *et al.* (2013), which has been reprinted within Figure 8.1. The illustrative example demonstrates that while there may a high degree of variability in which different subjects approach a task, the internal biomechanical constraints of that task can result in PCs reflecting interdependencies between variables, and much less variability amongst subjects about this PC. Federolf *et al.* considered all the marker trajectories and GRF waveforms within a single state space, and hence considered both interdependencies within and between these variables. In order to normalise the raw marker trajectories, each marker was expressed as a position relative to the centre four pelvis markers. Another way of expressing interdependencies within biomechanical variables would be to combine joint angles within a single PCA analysis. Either all variables could be combined, or specific variables which are known to be interdependent such as hip knee and ankle coronal angles combined with the knee adduction moment.

8.2 Non-linear PCA

The application of PCA identifies multi-dimensional factors through the analysis of linear correlations between input variables and applies a linear transformation of the input data to represent data along the newly defined components. In summary, it is most effective when relationships between input variables are linear. It might be the case however that in some instances, non-linear relationships exist between the input variables. For example, perhaps when knee flexion during swing phase is $\frac{3}{4}$ of the size, knee ROM is 75% of an average NP subjects, however, if it is $\frac{1}{2}$ the size, the ROM is only 20% of a healthy individual. A few researchers have explored the application of non-linear PCA analysis, often referred to “kernel” principal component analysis. These techniques

should be explored relative to standard PCA within level gait biomechanics to gleam If there is added value to this technique.

8.3 Subgrouping Using PCA

The OA and TKR subjects within this study have been considered as a single group, despite being a potentially heterogeneous cohort. Biomechanical changes and compensatory adaptations are not expected to be uniform within the cohort. More specifically, there might be subgroups within each cohort who are biomechanically distinct from each other. It is thought in particular that the identification of clinical phenotypes of knee OA could prove highly relevant to disease management/treatment (Knoop et al., 2011). The study of (Knoop et al., 2011), for example, used a k-means clustering algorithm to analyse 842 subjects considering KL grade, muscle strength, depression and BMI. The study identified five different clusters within the data, which were termed phenotypes, and these phenotypes were then contextualised based on the features which differentiate them e.g. “obese and weak muscle phenotype”. Distinctly different clinical outcomes were observed between the five phenotypes.

In a similar fashion to K-means clustering, PCA could be adopted in a slightly different way to identify subgroupings. A synthesised example is shown in Figure 8.2. PCA can be applied to the dataset of the variables which are thought to identify subgroupings. This data might be discrete metrics calculated from HMA data, and/or any relevant clinical measures such as strength, age, BMI, KL grade. The principal components of variance are then calculated within this data set, and hence the PC scores for each subject. Individual PCs can then be plotted in 2D, or 3D, in order to identify sub-groupings within the data. In this example, PC scores for the first two PCs have been plotted. There appears to be a clear clustering of NP subjects and two distinct clusters of OA subjects. Following the identification of the two clusters, the eigenvectors of the PCs can be investigated to identify the contributions of each input variable to the direction of the PCs of interest.

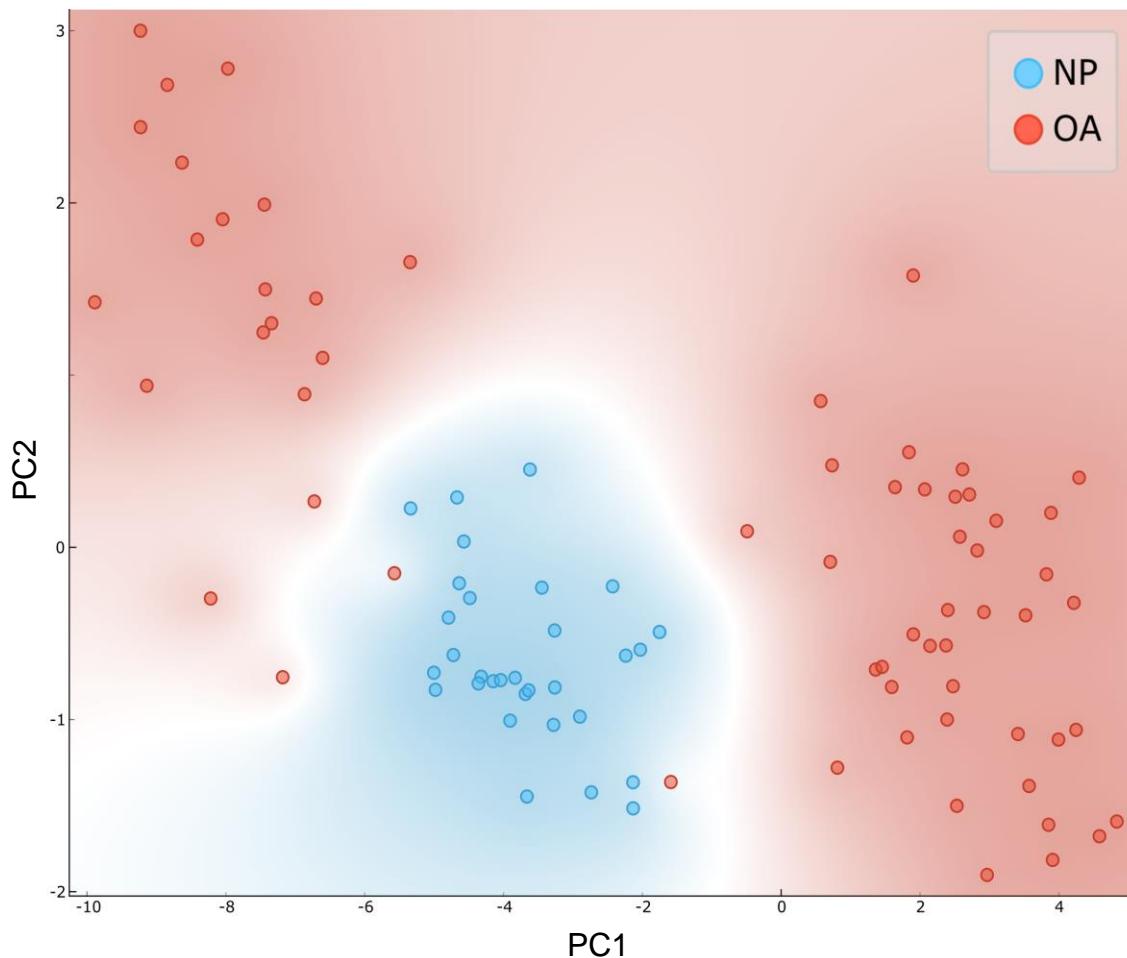


Figure 8.2 A synthesised dataset had been used to demonstrate how subgroups may be identifiable by using PCA as a clustering technique. Two distinct groups of OA (red) subjects can be seen, one with high PC1 scores, and the other with low PC1 scores but high PC2 scores relative to NP (blue) subjects.

Chapter 9 References

- AAOS. 2015. Surgical Management of Osteoarthritis of the Knee - Evidence-based clinical practice guideline. Available: http://www.aaos.org/uploadedFiles/PreProduction/Quality/Guidelines_and_Reviews/SMOAK%20CPG_12.4.15.pdf [Accessed 11/10/2016].
- Ackerman, I. N. & Bennell, K. L. 2004. Does pre-operative physiotherapy improve outcomes from lower limb joint replacement surgery? A systematic review. *Australian Journal of Physiotherapy*, 50, 25-30.
- Alpaydin, E. 2014. *Introduction to machine learning*, MIT press.
- Anakwe, R. E., Jenkins, P. J. & Moran, M. 2011. Predicting dissatisfaction after total hip arthroplasty: a study of 850 patients. *The Journal of arthroplasty*, 26, 209-213.
- Andersen, M. S., Mellon, S., Grammatopoulos, G. & Gill, H. S. 2013. Evaluation of the accuracy of three popular regression equations for hip joint centre estimation using computerised tomography measurements for metal-on-metal hip resurfacing arthroplasty patients. *Gait & posture*, 38, 1044-1047.
- Andriacchi, T. P. & Mündermann, A. 2006. The role of ambulatory mechanics in the initiation and progression of knee osteoarthritis. *Current opinion in rheumatology*, 18, 514-518.
- Angeloni, C., Riley, P. O. & Krebs, D. E. 1994. Frequency content of whole body gait kinematic data. *IEEE Transactions on Rehabilitation Engineering*, 2, 40-46.
- Artz, N., Elvers, K. T., Lowe, C. M., Sackley, C., Jepson, P. & Beswick, A. D. 2015. Effectiveness of physiotherapy exercise following total knee replacement: systematic review and meta-analysis. *BMC musculoskeletal disorders*, 16, 15.
- ARUK. 2013. Osteoarthritis in General Practice - Data and perspectives. *Arthritis Research UK* [Online]. Available: <http://www.arthritisresearchuk.org/policy-and-public-affairs/reports-and-resources/reports.aspx> [Accessed 23/03/2015].
- Assi, A., Ghanem, I., Lavaste, F. & Skalli, W. 2009. Gait analysis in children and uncertainty assessment for Davis protocol and Gillette Gait Index. *Gait & posture*, 30, 22-26.
- Astaphen, J. L., Deluzio, K. J., Caldwell, G. E. & 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, 332-341.
- Astaphen, J. L., Deluzio, K. J., Caldwell, G. E., Dunbar, M. J. & Hubley-Kozey, C. L. 2008b. Gait and neuromuscular pattern changes are associated with differences in knee osteoarthritis severity levels. *Journal of biomechanics*, 41, 868-876.
- Bader, D. L., Salter, D. M. & Chowdhury, T. T. 2011. Biomechanical influence of cartilage homeostasis in health and disease. *Arthritis*, 2011, 979032.
- Baker, P., Deehan, D., Lees, D., Jameson, S., Avery, P., Gregg, P. & Reed, M. 2012. The effect of surgical factors on early patient-reported outcome measures (PROMS) following total knee replacement. *J Bone Joint Surg Br*, 94, 1058-1066.
- Baker, P., Van der Meulen, J., Lewsey, J. & Gregg, P. 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. *Journal of Bone & Joint Surgery, British Volume*, 89, 893-900.
- Barton, G., Lees, A., Lisboa, P. & Attfield, S. 2006. Visualisation of gait data with Kohonen self-organising neural maps. *Gait & Posture*, 24, 46-53.
- Behrens, F., Kraft, E. L. & Oegema, T. R. 1989. Biochemical changes in articular cartilage after joint immobilization by casting or external fixation. *Journal of Orthopaedic Research*, 7, 335-343.
- Bell, A. L., Brand, R. A. & Pedersen, D. R. 1989. Prediction of hip joint centre location from external landmarks. *Human Movement Science*, 8, 3-16.
- Bell, A. L., Pedersen, D. R. & Brand, R. A. 1990. A comparison of the accuracy of several hip center location prediction methods. *Journal of biomechanics*, 23, 617-621.

- Benedetti, M., Catani, F., Leardini, A., Pignotti, E. & Giannini, S. 1998. Data management in gait analysis for clinical applications. *Clinical Biomechanics*, 13, 204-215.
- Bennell, K. L., Bowles, K.-A., Payne, C., Ciccuttini, F., Williamson, E., Forbes, A., Hanna, F., Davies-Tuck, M., Harris, A. & Hinman, R. S. 2011. Lateral wedge insoles for medial knee osteoarthritis: 12 month randomised controlled trial. *Bmj*, 342, d2912.
- Berry, D., Wold, L. & Rand, J. 1993. Extensive osteolysis around an aseptic, stable, uncemented total knee replacement. *Clinical orthopaedics and related research*, 204-207.
- Beynon, M., Curry, B. & Morgan, P. 2000. The Dempster-Shafer theory of evidence: an alternative approach to multicriteria decision modelling. *Omega*, 28, 37-50.
- Beynon, M. J. 2005. A novel technique of object ranking and classification under ignorance: An application to the corporate failure risk problem. *European Journal of Operational Research*, 167, 493-517.
- Beynon, M. J. & Buchanan, K. L. 2004. Object classification under ignorance using CaRBS: the case of the European barn swallow. *Expert systems with applications*, 27, 403-415.
- Beynon, M. J., Jones, L. & Holt, C. A. 2006. Classification of osteoarthritic and normal knee function using three-dimensional motion analysis and the Dempster-Shafer theory of evidence. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 36, 173-186.
- Beynon, M. J., Jones, L., Holt, C.A. Classification of osteoarthritic and normal knee function using three dimensional motion analysis and the Dempster-Shafer theory of evidence. Proceedings of the International Society of Biomechanics – 7th Symposium on Human Motion Analysis, 2002 Newcastle, UK. pp. 85-88.
- Bierma-Zeinstra, S. & Verhagen, A. P. 2011. Osteoarthritis subpopulations and implications for clinical trial design. *Arthritis Res Ther*, 13, 213.
- Bishop, C. M. 1995. *Neural networks for pattern recognition*, Oxford university press.
- Black, C., Clar, C., Henderson, R., MacEachern, C., McNamee, P., Quayyum, Z., Royle, P. & Thomas, S. 2009. The clinical effectiveness of glucosamine and chondroitin supplements in slowing or arresting progression of osteoarthritis of the knee: a systematic review and economic evaluation.
- Bourne, R. B. 2008. Measuring tools for functional outcomes in total knee arthroplasty. *Clinical orthopaedics and related research*, 466, 2634-2638.
- Bourne, R. B., Chesworth, B. M., Davis, A. M., Mahomed, N. N. & Charron, K. D. 2010. Patient satisfaction after total knee arthroplasty: who is satisfied and who is not? *Clinical Orthopaedics and Related Research®*, 468, 57-63.
- Boyer, K., Federolf, P., Lin, C., Nigg, B. & Andriacchi, T. 2012. Kinematic adaptations to a variable stiffness shoe: Mechanisms for reducing joint loading. *Journal of biomechanics*, 45, 1619-1624.
- Brandon, S. C. & Deluzio, K. J. 2011. Robust features of knee osteoarthritis in joint moments are independent of reference frame selection. *Clinical Biomechanics*, 26, 65-70.
- Brandon, S. C., Graham, R. B., Almosnino, S., Sadler, E. M., Stevenson, J. M. & Deluzio, K. J. 2013. Interpreting principal components in biomechanics: Representative extremes and single component reconstruction. *Journal of Electromyography and Kinesiology*, 23, 1304-1310.
- Brown, G. A. 2013. AAOS Clinical Practice Guideline: Treatment of Osteoarthritis of the Knee: Evidence-Based Guideline. *Journal of the American Academy of Orthopaedic Surgeons*, 21, 577-579.
- Buckwalter, J. A., Anderson, D. D., Brown, T. D., Tochigi, Y. & Martin, J. A. 2013. The roles of mechanical stresses in the pathogenesis of osteoarthritis: implications for treatment of joint injuries. *Cartilage*, 4, 286.
- Cappozzo, A., Catani, F., Della Croce, U. & Leardini, A. 1995. Position and orientation in space of bones during movement: anatomical frame definition and determination. *Clinical biomechanics*, 10, 171-178.

- Cappozzo, A., Della Croce, U., Leardini, A. & Chiari, L. 2005. Human movement analysis using stereophotogrammetry: Part 1: theoretical background. *Gait & posture*, 21, 186-196.
- Chang, A., Hurwitz, D., Dunlop, D., Song, J., Hayes, K. & Sharma, L. 2007. The relationship between toe-out angle during gait and progression of medial tibiofemoral osteoarthritis. *Annals of the rheumatic diseases*, 66, 1271-1275.
- Chau, T. 2001a. A review of analytical techniques for gait data. Part 1: fuzzy, statistical and fractal methods. *Gait & Posture*, 13, 49-66.
- Chau, T. 2001b. A review of analytical techniques for gait data. Part 2: neural network and wavelet methods. *Gait & Posture*, 13, 102-120.
- Chen, K., Li, G., Fu, D., Yuan, C., Zhang, Q. & Cai, Z. 2013. Patellar resurfacing versus nonresurfacing in total knee arthroplasty: a meta-analysis of randomised controlled trials. *International orthopaedics*, 37, 1075-1083.
- Cheng, M.-H., Ho, M.-F. & Huang, C.-L. 2008. Gait analysis for human identification through manifold learning and HMM. *Pattern recognition*, 41, 2541-2553.
- Chester, V. L. & Wrigley, A. T. 2008. The identification of age-related differences in kinetic gait parameters using principal component analysis. *Clinical Biomechanics*, 23, 212-220.
- Chiari, L., Della Croce, U., Leardini, A. & Cappozzo, A. 2005. Human movement analysis using stereophotogrammetry: Part 2: Instrumental errors. *Gait & posture*, 21, 197-211.
- Chohan, A., Dey, P., Sutton, C., Thewlis, D. & Richards, J. 2013. Do different methods of hip joint centre location impact on kinetics and kinematics in obese adults? *Gait & Posture*, 38, S109.
- Christensen, R., Bartels, E. M., Astrup, A. & Bliddal, H. 2007. Effect of weight reduction in obese patients diagnosed with knee osteoarthritis: a systematic review and meta-analysis. *Annals of the rheumatic diseases*, 66, 433-439.
- Clement, N. D., Macdonald, D. & Burnett, R. 2013. Predicting patient satisfaction using the Oxford knee score: where do we draw the line? *Archives of orthopaedic and trauma surgery*, 133, 689-694.
- Coggon, D., Reading, I., Croft, P., McLaren, M., Barrett, D. & Cooper, C. 2001. Knee osteoarthritis and obesity. *International Journal of Obesity & Related Metabolic Disorders*, 25.
- Comrey, A. L. & Lee, H. B. 2013. *A first course in factor analysis*, Psychology Press.
- Cretual, A., Bervet, K. & Ballaz, L. 2010. Gillette gait index in adults. *Gait & posture*, 32, 307-310.
- De Leva, P. 1996. Adjustments to Zatsiorsky-Seluyanov's segment inertia parameters. *Journal of biomechanics*, 29, 1223-1230.
- Della Croce, U., Cappozzo, A. & Kerrigan, D. C. 1999. Pelvis and lower limb anatomical landmark calibration precision and its propagation to bone geometry and joint angles. *Medical & biological engineering & computing*, 37, 155-161.
- Della Croce, U., Leardini, A., Chiari, L. & Cappozzo, A. 2005. Human movement analysis using stereophotogrammetry: Part 4: assessment of anatomical landmark misplacement and its effects on joint kinematics. *Gait & posture*, 21, 226-237.
- Deluzio, K. & Ast Stephen, J. 2007. Biomechanical features of gait waveform data associated with knee osteoarthritis: an application of principal component analysis. *Gait & posture*, 25, 86-93.
- Deluzio, K. J., Wyss, U. P., Costigan, P. A., Sorbie, C. & Zee, B. 1999. Gait assessment in unicompartmental knee arthroplasty patients: Principal component modelling of gait waveforms and clinical status. *Human Movement Science*, 18, 701-711.
- Deluzio, K. J., Wyss, U. P., Zee, B., Costigan, P. A. & Serbie, C. 1997. Principal component models of knee kinematics and kinetics: normal vs. pathological gait patterns. *Human Movement Science*, 16, 201-217.
- Dempster, A. P. 1968. A generalization of Bayesian inference. *Journal of the Royal Statistical Society. Series B (Methodological)*, 205-247.

- Dennis, D. A., Komistek, R. D., Scuderi, G. R. & Zingde, S. 2007. Factors affecting flexion after total knee arthroplasty. *Clinical orthopaedics and related research*, 464, 53-60.
- Dezert, J., Wang, P. & Tchamova, A. On the validity of Dempster-Shafer theory. Information Fusion (FUSION), 2012 15th International Conference on, 2012. IEEE, 655-660.
- Ditmyer, M. M., Topp, R. & Pifer, M. 2002. Prehabilitation in preparation for orthopaedic surgery. *Orthopaedic Nursing*, 21, 43-54.
- Dye, S. F. 1996. The knee as a biologic transmission with an envelope of function: a theory. *Clinical orthopaedics and related research*, 325, 10-18.
- Emrani, P. S., Katz, J. N., Kessler, C. L., Reichmann, W. M., Wright, E. A., McAlindon, T. E. & Losina, E. 2008. Joint space narrowing and Kellgren-Lawrence progression in knee osteoarthritis: an analytic literature synthesis. *Osteoarthritis and Cartilage*, 16, 873-882.
- Erhart, J. C., Mündermann, A., Elspas, B., Giori, N. J. & Andriacchi, T. P. 2010. Changes in knee adduction moment, pain, and functionality with a variable-stiffness walking shoe after 6 months. *Journal of Orthopaedic Research*, 28, 873-879.
- Escalante, Y., Saavedra, J. M., García-Hermoso, A., Silva, A. J. & Barbosa, T. M. 2010. Physical exercise and reduction of pain in adults with lower limb osteoarthritis: a systematic review. *Journal of Back and Musculoskeletal Rehabilitation*, 23, 175-186.
- Evans, J. D. 1996. *Straightforward statistics for the behavioral sciences*, Brooks/Cole.
- Fallah-Yakhdan, H. R., Abbasi-Bafghi, H., Meijer, O. G., Bruijn, S. M., van den Dikkenberg, N., Benedetti, M.-G. & van Dieën, J. H. 2012. Determinants of co-contraction during walking before and after arthroplasty for knee osteoarthritis. *Clinical Biomechanics*, 27, 485-494.
- Federolf, P., Boyer, K. & Andriacchi, T. 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, 2173-2178.
- Fernandes, L., Hagen, K. B., Bijlsma, J. W., Andreassen, O., Christensen, P., Conaghan, P. G., Doherty, M., Geenen, R., Hammond, A. & Kjeken, I. 2013. EULAR recommendations for the non-pharmacological core management of hip and knee osteoarthritis. *Annals of the rheumatic diseases*, 72, 1125-1135.
- Fernandez, J., Akbarshahi, M., Kim, H. & Pandy, M. 2008. Integrating modelling, motion capture and x-ray fluoroscopy to investigate patellofemoral function during dynamic activity. *Computer methods in biomechanics and biomedical engineering*, 11, 41-53.
- Ferré, L. 1995. Selection of components in principal component analysis: a comparison of methods. *Computational Statistics & Data Analysis*, 19, 669-682.
- Foroughi, N., Smith, R. & Vanwanseele, B. 2009. The association of external knee adduction moment with biomechanical variables in osteoarthritis: a systematic review. *The Knee*, 16, 303-9.
- Gandhi, R., Davey, J. R. & Mahomed, N. N. 2008. Predicting patient dissatisfaction following joint replacement surgery. *The Journal of rheumatology*, 35, 2415-2418.
- Garling, E. H., Kaptein, B. L., Mertens, B., Barendregt, W., Veeneg, H., Nelissen, R. G. & Valstar, E. R. 2007. Soft-tissue artefact assessment during step-up using fluoroscopy and skin-mounted markers. *Journal of biomechanics*, 40, S18-S24.
- Gaudreault, N., Mezghani, N., Turcot, K., Hagemeister, N., Boivin, K. & de Guise, J. A. 2011. Effects of physiotherapy treatment on knee osteoarthritis gait data using principal component analysis. *Clinical biomechanics*, 26, 284-291.
- Goodfellow, J., O'Connor, J. & Murray, D. 2010. A critique of revision rate as an outcome measure: re-interpretation of knee joint registry data. . *Journal of Bone & Joint Surgery, British Volume*, 92, 1628-1631.
- Grood, E. S. & 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, 136.

- Hamilton, D., Lane, J. V., Gaston, P., Patton, J., Macdonald, D., Simpson, A. & Howie, C. 2013. What determines patient satisfaction with surgery? A prospective cohort study of 4709 patients following total joint replacement. *BMJ open*, 3, e002525.
- Harrington, M., Zavatsky, A., Lawson, S., Yuan, Z. & Theologis, T. 2007. Prediction of the hip joint centre in adults, children, and patients with cerebral palsy based on magnetic resonance imaging. *Journal of biomechanics*, 40, 595-602.
- Hatze, H. 1981. The use of optimally regularized Fourier series for estimating higher-order derivatives of noisy biomechanical data. *Journal of Biomechanics*, 14, 13-18.
- Helwig, N. E., Hong, S., Hsiao-Wecksler, E. T. & Polk, J. D. 2011. Methods to temporally align gait cycle data. *Journal of biomechanics*, 44, 561-566.
- Hillman, S. J., Hazlewood, M. E., Schwartz, M. H., van der Linden, M. L. & Robb, J. E. 2007. Correlation of the Edinburgh gait score with the Gillette gait index, the Gillette functional assessment questionnaire, and dimensionless speed. *Journal of Pediatric Orthopaedics*, 27, 7-11.
- Hinman, R. S., Bowles, K. A., Metcalf, B. B., Wrigley, T. V. & Bennell, K. L. 2012. Lateral wedge insoles for medial knee osteoarthritis: effects on lower limb frontal plane biomechanics. *Clinical Biomechanics*, 27, 27-33.
- Hochberg, M. C., Altman, R. D., April, K. T., Benkhalti, M., Guyatt, G., McGowan, J., Towheed, T., Welch, V., Wells, G. & Tugwell, P. 2012. American College of Rheumatology 2012 recommendations for the use of nonpharmacologic and pharmacologic therapies in osteoarthritis of the hand, hip, and knee. *Arthritis care & research*, 64, 465-474.
- Hof, A., Gazendam, M. & Sinke, W. 2005. The condition for dynamic stability. *Journal of biomechanics*, 38, 1-8.
- Holt, C., Hayes, N., van Deursen, R. & O'Callaghan, P. 2001. Three-dimensional analysis of the tibiofemoral joint using external marker clusters and the JCS approach—comparison of normal and osteoarthritic knee function. *Computer methods in biomechanics and biomedical engineering*, 3, 289-294.
- Holt, C. A., Whatling, G. M., Wilson, C., Bonnet, C., Elford, C., Brakspear, K., Biggs, P. & Mason, D. J. 2016. Biological changes in tibial subchondral bone following high tibial osteotomy. *Osteoarthritis and Cartilage*, 24, S511.
- Hortobágyi, T., Westerkamp, L., Beam, S., Moody, J., Garry, J., Holbert, D. & DeVita, P. 2005. Altered hamstring-quadriceps muscle balance in patients with knee osteoarthritis. *Clinical Biomechanics*, 20, 97-104.
- Hui, C., Salmon, L. J., Kok, A., Williams, H. A., Hockers, N., van der Tempel, W. M., Chana, R. & Pinczewski, L. A. 2011. Long-term survival of high tibial osteotomy for medial compartment osteoarthritis of the knee. *The American journal of sports medicine*, 39, 64-70.
- Hunt, M., Birmingham, T., Bryant, D., Jones, I., Giffin, J., Jenkyn, T. & Vandervoort, A. 2008. Lateral trunk lean explains variation in dynamic knee joint load in patients with medial compartment knee osteoarthritis. *Osteoarthritis and Cartilage*, 16, 591-599.
- Hunt, M. A., Wrigley, T. V., Hinman, R. S. & 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, 1426-1432.
- Hurwitz, D., Ryals, A., Case, J., Block, J. & Andriacchi, T. 2002. The knee adduction moment during gait in subjects with knee osteoarthritis is more closely correlated with static alignment than radiographic disease severity, toe out angle and pain. *Journal of orthopaedic research*, 20, 101-107.
- Intrator, O. & Intrator, N. 2001. Interpreting neural-network results: a simulation study. *Computational statistics & data analysis*, 37, 373-393.
- Jackson, D. A. 1993. Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches. *Ecology*, 2204-2214.
- Jackson, J. E. 1991. *A User's Guide to Principal Components*, Wiley.

- Jones, L. 2004. *Method For The Classification Of Osteoarthritic And Normal Knee Function*. PhD, Cardiff University.
- Jones, L., Beynon, M. J., Holt, C. A. & 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, 2512-2520.
- Jones, L. & Holt, C. 2008. An objective tool for assessing the outcome of total knee replacement surgery. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 222, 647-655.
- Jones, L., Holt, C. A. & 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, 31-40.
- Jordan, R., Smith, N., Chahal, G., Casson, C., Reed, M. & Sprowson, A. 2014. Enhanced education and physiotherapy before knee replacement; is it worth it? A systematic review. *Physiotherapy*, 100, 305-312.
- Kainz, H., Carty, C. P., Modenese, L., Boyd, R. N. & Lloyd, D. G. 2015. Estimation of the hip joint centre in human motion analysis: A systematic review. *Clinical biomechanics*, 30, 319-329.
- Kaiser, H. F. 1960. The application of electronic computers to factor analysis. *Educational and psychological measurement*.
- Kakihana, W., Akai, M., Nakazawa, K., Takashima, T., Naito, K. & Torii, S. 2005. Effects of laterally wedged insoles on knee and subtalar joint moments. *Archives of physical medicine and rehabilitation*, 86, 1465-1471.
- Katz, D. & Ali, A. 2009. Preventive medicine, integrative medicine & the health of the public. Commissioned for the IOM summit on integrative medicine and the health of the public. *Institute of Medicine (IOM), Washington, DC*.
- Kerrigan, D. C., Frates, E. P., Rogan, S. & Riley, P. O. 2000. Hip hiking and circumduction: quantitative definitions. *American journal of physical medicine & rehabilitation*, 79, 247-252.
- Kerrigan, D. C., Lelas, J. L., Goggins, J., Merriman, G. J., Kaplan, R. J. & Felson, D. T. 2002. Effectiveness of a lateral-wedge insole on knee varus torque in patients with knee osteoarthritis. *Archives of Physical Medicine and Rehabilitation*, 83, 889-893.
- Kim, T. K., Chang, C. B., Kang, Y. G., Kim, S. J. & Seong, S. C. 2009. Causes and predictors of patient's dissatisfaction after uncomplicated total knee arthroplasty. *The Journal of arthroplasty*, 24, 263-271.
- Kirkwood, R. N., Resende, R. A., Magalhães, C., Gomes, H. A., Mingoti, S. A. & Sampaio, R. F. 2011. Application of principal component analysis on gait kinematics in elderly women with knee osteoarthritis. *Brazilian Journal of Physical Therapy*, 15, 52-58.
- Knoop, J., van der Leeden, M., Thorstensson, C. A., Roorda, L. D., Lems, W. F., Knol, D. L., Steultjens, M. P. & Dekker, J. 2011. Identification of phenotypes with different clinical outcomes in knee osteoarthritis: data from the Osteoarthritis Initiative. *Arthritis care & research*, 63, 1535-1542.
- Kraus, V. B., Blanco, F., Englund, M., Karsdal, M. & Lohmander, L. 2015. Call for standardized definitions of osteoarthritis and risk stratification for clinical trials and clinical use. *Osteoarthritis and Cartilage*, 23, 1233-1241.
- Kristianslund, E., Krosshaug, T. & Van den Bogert, A. J. 2012. Effect of low pass filtering on joint moments from inverse dynamics: implications for injury prevention. *Journal of biomechanics*, 45, 666-671.
- Kurz, B., Lemke, A. K., Fay, J., Pufe, T., Grodzinsky, A. J. & Schünke, M. 2005. Pathomechanisms of cartilage destruction by mechanical injury. *Annals of Anatomy-Anatomischer Anzeiger*, 187, 473-485.
- Kutzner, I., Trepczynski, A., Heller, M. O. & Bergmann, G. 2013. Knee adduction moment and medial contact force—facts about their correlation during gait. *PloS one*, 8, e81036.

- Lafortune, M., Cavanagh, P. R., Sommer, H. & Kalenak, A. 1992. Three-dimensional kinematics of the human knee during walking. *Journal of biomechanics*, 25, 347-357.
- Lafuente, R., Belda, J., Sanchez-Lacuesta, J., Soler, C. & Prat, J. 1998. Design and test of neural networks and statistical classifiers in computer-aided movement analysis: a case study on gait analysis. *Clinical Biomechanics*, 13, 216-229.
- Lai, D. T., Begg, R. K. & Palaniswami, M. 2009. Computational intelligence in gait research: a perspective on current applications and future challenges. *IEEE Transactions on Information Technology in Biomedicine*, 13, 687-702.
- Landry, S. C., McKean, K. A., Hubley-Kozey, C. L., Stanish, W. D. & Deluzio, K. J. 2007. Knee biomechanics of moderate OA patients measured during gait at a self-selected and fast walking speed. *Journal of biomechanics*, 40, 1754-1761.
- Lane, N., Brandt, K., Hawker, G., Peeva, E., Schreyer, E., Tsuji, W. & Hochberg, M. 2011. OARSI-FDA initiative: defining the disease state of osteoarthritis. *Osteoarthritis and Cartilage*, 19, 478-482.
- Leardini, A., Chiari, L., Della Croce, U. & Cappozzo, A. 2005. Human movement analysis using stereophotogrammetry: Part 3. Soft tissue artifact assessment and compensation. *Gait & posture*, 21, 212-225.
- Lelas, J. L., Merriman, G. J., Riley, P. O. & Kerrigan, D. C. 2003. Predicting peak kinematic and kinetic parameters from gait speed. *Gait & posture*, 17, 106-112.
- Li, N., Tan, Y., Deng, Y. & Chen, L. 2014. Posterior cruciate-retaining versus posterior stabilized total knee arthroplasty: a meta-analysis of randomized controlled trials. *Knee Surgery, Sports Traumatology, Arthroscopy*, 22, 556-564.
- Lingard, E. A., Katz, J. N., Wright, E. A. & Sledge, C. B. 2004. Predicting the outcome of total knee arthroplasty. *J Bone Joint Surg Am*, 86, 2179-2186.
- Lungu, E., Desmeules, F., Dionne, C. E., Belzile, É. L. & Vendittoli, P.-A. 2014. Prediction of poor outcomes six months following total knee arthroplasty in patients awaiting surgery. *BMC musculoskeletal disorders*, 15, 1.
- Madry, H., Kon, E., Condello, V., Peretti, G. M., Steinwachs, M., Seil, R., Berruto, M., Engebretsen, L., Filardo, G. & Angele, P. 2016. Early osteoarthritis of the knee. *Knee Surgery, Sports Traumatology, Arthroscopy*, 1-10.
- Mahomed, N. N., Liang, M. H., Cook, E. F., Daltroy, L. H., Fortin, P. R., Fossel, A. H. & Katz, J. N. 2002. The importance of patient expectations in predicting functional outcomes after total joint arthroplasty. *The Journal of rheumatology*, 29, 1273-1279.
- Maly, M. R., Costigan, P. A. & Olney, S. J. 2006. Determinants of self-report outcome measures in people with knee osteoarthritis. *Archives of physical medicine and rehabilitation*, 87, 96-104.
- Matassi, F., Carulli, C., Civinini, R. & Innocenti, M. 2013. Cemented versus cementless fixation in total knee arthroplasty. *Joints*, 1, 121.
- McAlindon, T. E., Bannuru, R. R., Sullivan, M., Arden, N., Berenbaum, F., Bierma-Zeinstra, S., Hawker, G., Henrotin, Y., Hunter, D. & Kawaguchi, H. 2014. OARSI guidelines for the non-surgical management of knee osteoarthritis. *Osteoarthritis and cartilage*, 22, 363-388.
- McClelland, J. A., Webster, K. E. & Feller, J. A. 2007. Gait analysis of patients following total knee replacement: a systematic review. *The Knee*, 14, 253-263.
- McLaughlin, T. M., Dillman, C. J. & Lardner, T. J. 1977. Biomechanical analysis with cubic spline functions. *Research Quarterly. American Alliance for Health, Physical Education and Recreation*, 48, 569-582.
- McMulkin, M. L. & MacWilliams, B. A. 2008. Intersite variations of the gillette gait index. *Gait & posture*, 28, 483-487.
- Meding, J. B., Keating, E. M., Ritter, M. A., Faris, P. M., Berend, M. E. & Malinzak, R. A. 2005. The planovalgus foot: a harbinger of failure of posterior cruciate-retaining total knee replacement. *J Bone Joint Surg Am*, 87, 59-62.
- Metcalfe, A., Stewart, C., Postans, N., Biggs, P., Whatling, G., Holt, C. & Roberts, A. 2017. Abnormal loading and functional deficits are present in both limbs before and after unilateral knee arthroplasty. *Gait & Posture*.

- Metcalfe, A., Stewart, C., Postans, N., Dodds, A., Holt, C. A. & Roberts, A. 2013. The effect of osteoarthritis of the knee on the biomechanics of other joints in the lower limbs. *Bone & Joint Journal*, 95, 348-353.
- Metcalfe, A. J. 2014. *Knee osteoarthritis is a bilateral disease*. Cardiff University.
- Michie, D., Spiegelhalter, D. J. & Taylor, C. C. 1994. *Machine Learning, Neural and Statistical Classification*, Ellis Horwood.
- Mizner, R. L., Pettersson, S. C., Clements, K. E., Zeni, J. A., Irrgang, J. J. & Snyder-Mackler, L. 2011. Measuring functional improvement after total knee arthroplasty requires both performance-based and patient-report assessments: a longitudinal analysis of outcomes. *The Journal of arthroplasty*, 26, 728-737.
- Mizner, R. L., Pettersson, S. C., Stevens, J. E., Axe, M. J. & Snyder-Mackler, L. 2005. Preoperative quadriceps strength predicts functional ability one year after total knee arthroplasty. *The Journal of rheumatology*, 32, 1533-1539.
- Mizner, R. L. & Snyder-Mackler, L. 2005. Altered loading during walking and sit-to-stand is affected by quadriceps weakness after total knee arthroplasty. *Journal of Orthopaedic Research*, 23, 1083-1090.
- Moreland, J. D., Richardson, J. A., Goldsmith, C. H. & Clase, C. M. 2004. Muscle weakness and falls in older adults: a systematic review and meta-analysis. *Journal of the American Geriatrics Society*, 52, 1121-1129.
- Mündermann, A., Dyrby, C. O. & 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, 2835-2844.
- Murray, D., Fitzpatrick, R., Rogers, K., Pandit, H., Beard, D., Carr, A. & Dawson, J. 2007. The use of the Oxford hip and knee scores. *Bone & Joint Journal*, 89, 1010-1014.
- Naili, J. E., Iversen, M. D., Esbjörnsson, A.-C., Hedström, M., Schwartz, M. H., Häger, C. K. & Broström, E. W. 2016. Deficits in functional performance and gait one year after total knee arthroplasty despite improved self-reported function. *Knee Surgery, Sports Traumatology, Arthroscopy*, 1-9.
- NICE. 2014. *Osteoarthritis: care and management* [Online]. Available: <https://www.nice.org.uk/guidance/cg177/chapter/2-Research-recommendations> [Accessed 20/04/2017 2017].
- Nilsson, J. & Thorstensson, A. 1989. Ground reaction forces at different speeds of human walking and running. *Acta Physiologica Scandinavica*, 136, 217-227.
- Noble, P. C., Conditt, M. A., Cook, K. F. & Mathis, K. B. 2006. The John Insall Award: Patient expectations affect satisfaction with total knee arthroplasty. *Clinical orthopaedics and related research*, 452, 35-43.
- Oberg, T., Karsznia, A. & Oberg, K. 1994. Joint angle parameters in gait: reference data for normal subjects, 10-79 years of age. *Journal of rehabilitation Research and Development*, 31, 199-213.
- Parisi, L., Biggs, P. R., Whatling, G. M. & Holt, C. A. A Novel Comparison of Artificial Intelligence Methods for Diagnosing Knee Osteoarthritis. XXV Congress of the International Society of Biomechanics, 2015. International Society of Biomechanics, 1227-1229.
- Partridge, T., Carluke, I., Emmerson, K., Partington, P. & Reed, M. 2016. Improving patient reported outcome measures (PROMs) in total knee replacement by changing implant and preserving the infrapatella fatpad: a quality improvement project. *BMJ quality improvement reports*, 5, u204088. w3767.
- Patterson, K. K., Gage, W. H., Brooks, D., Black, S. E. & McIlroy, W. E. 2010. Evaluation of gait symmetry after stroke: a comparison of current methods and recommendations for standardization. *Gait & posture*, 31, 241-246.
- Peters, A., Galna, B., Sangeux, M., Morris, M. & Baker, R. 2010. Quantification of soft tissue artifact in lower limb human motion analysis: a systematic review. *Gait & posture*, 31, 1-8.
- Pezzack, J., Norman, R. & Winter, D. 1977. An assessment of derivative determining techniques used for motion analysis. *Journal of biomechanics*, 10, 377-382.
- Powell, V. 2015. *Principal Component Analysis - Explained Visually* [Online]. Available: <http://setosa.io/ev/principal-component-analysis/> [Accessed 18/05/2016 2016].

- Price, A., Longino, D., Rees, J., Rout, R., Pandit, H., Javaid, K., Arden, N., Cooper, C., Carr, A. & Dodd, C. 2010. Are pain and function better measures of outcome than revision rates after TKR in the younger patient? *The Knee*, 17, 196-199.
- Prince, F., Corriveau, H., Hébert, R. & Winter, D. A. 1997. Gait in the elderly. *Gait & Posture*, 5, 128-135.
- Raja, K. & Dewan, N. 2011. Efficacy of knee braces and foot orthoses in conservative management of knee osteoarthritis: a systematic review. *American journal of physical medicine & rehabilitation*, 90, 247-262.
- Rankin, E. A., Alarcón, G. S., Chang, R. W. & Cooney Jr, L. M. 2004. NIH Consensus Statement on total knee replacement December 8-10, 2003. *Journal of Bone and Joint Surgery*, 86, 1328.
- Razmjou, H., Schwartz, C. E., Yee, A. & Finkelstein, J. A. 2009. Traditional assessment of health outcome following total knee arthroplasty was confounded by response shift phenomenon. *Journal of clinical epidemiology*, 62, 91-96.
- Reid, S. M., Graham, R. B. & Costigan, P. A. 2010. Differentiation of young and older adult stair climbing gait using principal component analysis. *Gait & posture*, 31, 197-203.
- Rodriguez-Merchan, E. C. 2014. The Influence of Obesity on the Outcome of TKR: Can the Impact of Obesity be justified from the Viewpoint of the Overall Health Care System? *HSS Journal®*, 10, 167-170.
- Sadeghi, H., Allard, P., Barbier, F., Sadeghi, S., Hinse, S., Perrault, R. & Labelle, H. 2002. Main functional roles of knee flexors/extensors in able-bodied gait using principal component analysis (I). *The Knee*, 9, 47-53.
- Sadeghi, H., Allard, P., Shafie, K., Mathieu, P. A., Sadeghi, S., Prince, F. & Ramsay, J. 2000. Reduction of gait data variability using curve registration. *Gait & Posture*, 12, 257-264.
- Safranek, R. J., Gottschlich, S. & Kak, A. C. 1990. Evidence accumulation using binary frames of discernment for verification vision. *Robotics and Automation, IEEE Transactions on*, 6, 405-417.
- Sangeux, M., Pillet, H. & Skalli, W. 2014. Which method of hip joint centre localisation should be used in gait analysis? *Gait & posture*, 40, 20-25.
- Schache, A. G., Baker, R. & Vaughan, C. L. 2007. Differences in lower limb transverse plane joint moments during gait when expressed in two alternative reference frames. *Journal of biomechanics*, 40, 9-19.
- Schöllhorn, W. 2004. Applications of artificial neural nets in clinical biomechanics. *Clinical Biomechanics*, 19, 876-898.
- Schutte, L., Narayanan, U., Stout, J., Selber, P., Gage, J. & Schwartz, M. 2000. An index for quantifying deviations from normal gait. *Gait & posture*, 11, 25-31.
- Schwartz, M. H. & Rozumalski, A. 2008. The Gait Deviation Index: a new comprehensive index of gait pathology. *Gait & posture*, 28, 351-357.
- Scott, C., Howie, C., MacDonald, D. & Biant, L. 2010. Predicting dissatisfaction following total knee replacement A PROSPECTIVE STUDY OF 1217 PATIENTS. *Journal of Bone & Joint Surgery, British Volume*, 92, 1253-1258.
- Shafer, G. 1976. *A mathematical theory of evidence*, Princeton university press Princeton.
- Sharma, L., Hayes, K. W., Felson, D. T., Buchanan, T. S., Kirwan-Mellis, G., Lou, C., Pai, Y.-C. & Dunlop, D. D. 1999. Does laxity alter the relationship between strength and physical function in knee osteoarthritis? *Arthritis & Rheumatism*, 42, 25-32.
- Sharma, L., Hurwitz, D. E., Thonar, E. J., Sum, J. A., Lenz, M. E., Dunlop, D. D., Schnitzer, T. J., Kirwan-Mellis, G. & Andriacchi, T. P. 1998. Knee adduction moment, serum hyaluronan level, and disease severity in medial tibiofemoral osteoarthritis. *Arthritis and rheumatism*, 41, 1233-40.
- Shimada, S., Kobayashi, S., Wada, M., Uchida, K., Sasaki, S., Kawahara, H., Yayama, T., Kitade, I., Kamei, K. & Kubota, M. 2006. Effects of disease severity on response to lateral wedged shoe insole for medial compartment knee osteoarthritis. *Archives of physical medicine and rehabilitation*, 87, 1436-1441.

- Sinclair, J., Taylor, P. J. & Hobbs, S. J. 2013. Digital filtering of three-dimensional lower extremity kinematics: An assessment. *Journal of human kinetics*, 39, 25-36.
- Singh, J., O'Byrne, M., Colligan, R. & Lewallen, D. 2010. Pessimistic explanatory style A PSYCHOLOGICAL RISK FACTOR FOR POOR PAIN AND FUNCTIONAL OUTCOMES TWO YEARS AFTER KNEE REPLACEMENT. *Journal of Bone & Joint Surgery, British Volume*, 92, 799-806.
- Smith, S. C., Cano, S., Lamping, D. L., Staniszewska, S., Browne, J., Lewsey, J., van der Meulen, J., Cairns, J. & Black, N. 2005. Patient-Reported Outcome Measures (PROMs) for routine use in Treatment Centres: recommendations based on a review of the scientific evidence.
- Snelling, S., Rout, R., Davidson, R., Clark, I., Carr, A., Hulley, P. & Price, A. 2014. A gene expression study of normal and damaged cartilage in anteromedial gonarthrosis, a phenotype of osteoarthritis. *Osteoarthritis and Cartilage*, 22, 334-343.
- Spicer, D., Pomeroy, D., Badenhausen, W., Schaper, J. L., Curry, J., Suthers, K. & Smith, M. 2001. Body mass index as a predictor of outcome in total knee replacement. *International orthopaedics*, 25, 246-249.
- Stanhope, S., Kepple, T., McGuire, D. & Roman, N. 1990. Kinematic-based technique for event time determination during gait. *Medical and Biological Engineering and Computing*, 28, 355-360.
- Stratford, P. W. & Kennedy, D. M. 2006. Performance measures were necessary to obtain a complete picture of osteoarthritic patients. *Journal of clinical epidemiology*, 59, 160-167.
- Tanamas, S., Hanna, F. S., Cicuttini, F. M., Wluka, A. E., Berry, P. & Urquhart, D. M. 2009. Does knee malalignment increase the risk of development and progression of knee osteoarthritis? A systematic review. *Arthritis care & research*, 61, 459-467.
- Toda, Y., Segal, N., Kato, A., Yamamoto, S. & Irie, M. 2001. Effect of a novel insole on the subtalar joint of patients with medial compartment osteoarthritis of the knee. *The Journal of rheumatology*, 28, 2705.
- Tome, J., Nawoczenski, D. A., Flemister, A. & Houck, J. 2006. Comparison of foot kinematics between subjects with posterior tibialis tendon dysfunction and healthy controls. *Journal of Orthopaedic & Sports Physical Therapy*, 36, 635-644.
- UK-NJR. 2014. 11th Annual Report. Available: http://www.njrcentre.org.uk/njrcentre/Portals/0/Documents/England/Reports/11th_annual_report/NJR%2011th%20Annual%20Report%202014.pdf [Accessed 30/04/2016].
- UK-NJR. 2015. 12th Annual Report. Available: <http://www.njrcentre.org.uk/njrcentre/Reports/PublicationsandMinutes/Annualreports/tabid/86/Default.aspx> [Accessed 30/04/2016].
- UKHF. 2014. Risk factor based modelling for Public Health England. Available: <http://www.ukhealthforum.org.uk/prevention/pie/?entryid43=38207> [Accessed 27/04/2016].
- Van den Bogert, A. & De Koning, J. On optimal filtering for inverse dynamics analysis. Proceedings of the IXth biennial conference of the Canadian society for biomechanics, 1996. 214-215.
- Van der Esch, M., Knoop, J., van der Leeden, M., Roorda, L., Lems, W., Knol, D. & Dekker, J. 2015. Clinical phenotypes in patients with knee osteoarthritis: a study in the Amsterdam osteoarthritis cohort. *Osteoarthritis and Cartilage*, 23, 544-549.
- Van der Voort, P., Pijls, B., Nouta, K., Valstar, E., Jacobs, W. & Nelissen, R. 2013. A systematic review and meta-regression of mobile-bearing versus fixed-bearing total knee replacement in 41 studies. *Bone & Joint Journal*, 95, 1209-1216.
- Van Meurs, J. & Uitterlinden, A. 2012. Osteoarthritis year 2012 in review: genetics and genomics. *Osteoarthritis and Cartilage*, 20, 1470-1476.
- van Raaij, T. M., Reijman, M., Brouwer, R. W., Bierma-Zeinstra, S. M. & Verhaar, J. A. 2010. Medial knee osteoarthritis treated by insoles or braces: a randomized trial. *Clinical Orthopaedics and Related Research®*, 468, 1926-1932.

- Vanwanseele, B., Eckstein, F., Knecht, H., Spaepen, A. & Stüssi, E. 2003. Longitudinal analysis of cartilage atrophy in the knees of patients with spinal cord injury. *Arthritis & Rheumatism*, 48, 3377-3381.
- von Tscharner, V., Enders, H. & Maurer, C. 2013. Subspace identification and classification of healthy human gait. *PLoS one*, 8, e65063.
- Warner, M. B., Whatling, G., Worsley, P. R., Mottram, S., Chappell, P. H., Holt, C. A. & Stokes, M. J. 2015. Objective classification of scapular kinematics in participants with movement faults of the scapula on clinical assessment. *Computer methods in biomechanics and biomedical engineering*, 18, 782-789.
- Watling, D. 2014. *Development of novel methodologies to quantify, analyse and classify in-vivo knee function affected by aging, osteoarthritis and total knee replacement*. Citeseer.
- Weinhandl, J. T. & O'Connor, K. M. 2010. Assessment of a greater trochanter-based method of locating the hip joint center. *Journal of Biomechanics*, 43, 2633-2636.
- Werner, F. W., Ayers, D. C., Maletsky, L. P. & Rullkoetter, P. J. 2005. The effect of valgus/varus malalignment on load distribution in total knee replacements. *Journal of biomechanics*, 38, 349-355.
- 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. Cardiff University.
- Whatling, G. M., Dabke, H. V., Holt, C. A., Jones, I., Madete, J., Alderman, P. M. & Roberts, P. 2008. Objective functional assessment of total hip arthroplasty following two common surgical approaches: the posterior and direct lateral approaches. *Proceedings of the Institution of Mechanical Engineers Part H-Journal of Engineering in Medicine*, 222, 897-905.
- Whatling, G. M., Holt, C. A. & Beynon, M. J. 2015. The application of NCaRBS to the Trendelenburg test and total hip arthroplasty outcome. *Annals of biomedical engineering*, 43, 363-375.
- Winter, D. A. 1990. *Biomechanics and Motor Control of Human Movement*, Wiley.
- Winter, D. A., Sidwall, H. G. & Hobson, D. A. 1974. Measurement and reduction of noise in kinematics of locomotion. *Journal of biomechanics*, 7, 157-159.
- Worsley, P. 2011. *Assessment of short-term knee arthroplasty function using clinical measures, Motion Analysis, and Musculoskeletal Modelling*. Univiersity of Southampton.
- Wren, T. A., Do, K. P., Hara, R., Dorey, F. J., Kay, R. M. & Otsuka, N. Y. 2007. Gillette Gait Index as a gait analysis summary measure: comparison with qualitative visual assessments of overall gait. *Journal of Pediatric Orthopaedics*, 27, 765-768.
- Wu, G., Siegler, S., Allard, P., Kirtley, C., Leardini, A., Rosenbaum, D., Whittle, M., D D'Lima, D., Cristofolini, L. & Witte, H. 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, 543-548.
- Zadeh, L. A. 1984. Review of a mathematical theory of evidence. *AI magazine*, 5, 81.
- Zatsiorsky, V. & Seluyanov, V. 1983. The mass and inertia characteristics of the main segments of the human body. *Biomechanics viii-b*, 56, 1152-1159.
- Zeni, J. A., Rudolph, K. & Higginson, J. S. 2010. Alterations in quadriceps and hamstrings coordination in persons with medial compartment knee osteoarthritis. *Journal of Electromyography and Kinesiology*, 20, 148-154.
- Zhang, W., Moskowitz, R., Nuki, G., Abramson, S., Altman, R., Arden, N., Bierma-Zeinstra, S., Brandt, K., Croft, P. & Doherty, M. 2007. OARSI recommendations for the management of hip and knee osteoarthritis, part I: critical appraisal of existing treatment guidelines and systematic review of current research evidence. *Osteoarthritis and cartilage*, 15, 981-1000.
- Zhang, W., Moskowitz, R., Nuki, G., Abramson, S., Altman, R., Arden, N., Bierma-Zeinstra, S., Brandt, K., Croft, P. & Doherty, M. 2008. OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. *Osteoarthritis and cartilage*, 16, 137-162.

- Zhang, W., Nuki, G., Moskowitz, R., Abramson, S., Altman, R. D., Arden, N., Bierma-Zeinstra, S., Brandt, K., Croft, P. & Doherty, M. 2010. OARSI recommendations for the management of hip and knee osteoarthritis: part III: Changes in evidence following systematic cumulative update of research published through January 2009. *Osteoarthritis and Cartilage*, 18, 476-499.