

**PREOPERATIVE SYMPTOM BURDEN IN  
THE KNEE REPLACEMENT POPULATION  
- THE ROLE OF THE OXFORD KNEE  
SCORE**

THESIS SUBMITTED IN PARTIAL FULFILMENT  
FOR THE DEGREE OF  
MASTER OF SCIENCE BY RESEARCH



BY  
DERFEL PARI WILLIAMS

WOLFSON COLLEGE, OXFORD

TRINITY TERM 2015

## **Acknowledgements**

The work presented in this thesis would not have been possible without the support and guidance of a number of individuals along the way. Foremost, I would like to express my gratitude to both my supervisors, Andrew Price and David Beard, for sharing their expertise and ideas with me, and providing the opportunity and encouragement to undertake this work in the first place. Furthermore, I am thankful for the support and guidance delivered throughout the process.

During my time at the Nuffield Department of Orthopaedic, Rheumatology and Musculoskeletal Sciences, I had the pleasure to work with several individuals who have contributed in some small way to my work, and without whom I may never have reached this point. In particular, would like to thank Kristina Harris for her support with the database management, to Ines Rombach for her expertise in providing statistical guidance during my research.

As a direct result of my studies, I was lucky enough to be given an opportunity to collaborate with other experienced and respected researchers, further developing my research knowledge and experiences. I am grateful to all the team at the Primary Joint Unit, Musgrove Park Hospital, Belfast, with particular mention to David Beverland, Seamus O'Brien, and Emer Doran. Similarly, to Richard Field and Gaye Hadfield at the Elective Orthopaedic Centre, Epsom.

I would like to dedicate this work to my family. A special thanks to my wife Ceri, for her support and hard work at home, without whom I could not have dedicated the time and effort to complete my studies. Finally, thank you to my children, Dyfrig and Mabli, for ensuring life remained fun, exciting and enjoyable during the stressful times.

## **Abstract**

The Oxford Knee Score (OKS) represents a validated Patient Reported Outcome Score (PROM), routinely utilised following knee replacement (TKR). Healthcare commissioners have proposed its application as a threshold criterion to define eligibility for TKR, despite an absence of supporting evidence. This thesis explores the role of preoperative OKS, using local and national PROMs data, as a measure of symptom burden and highlights potential issues as a threshold criterion tool.

Examination of the preoperative OKS profile in three orthopaedic centres showed a normal distribution profile, and wide-ranging preoperative symptom burden. Some patients appeared to proceed to TKR with a seemingly minimal symptom burden, raising question about TKR selection criteria. Further analysis of national PROMs data confirmed wide variability in preoperative symptom burden prior to TKR to be common practice. Some variation existed between centres and individual surgeons, but patient level factors were a greater contributor, suggesting poor selection criteria may not be to blame. Case note analysis of a TKR cohort from a single centre identified that those with a high preoperative OKS still exhibited a significant symptom burden, remaining clinically justified for TKR. An appropriate consultation process was evident, rejecting inappropriate selection criteria as an explanation. Analysis of linked national pre- and postoperative OKS data showed that patients with highest preoperative OKS had best overall outcomes following TKR, despite minimal change in OKS. Preoperative OKS < 16 was associated with worse outcomes and minimal further benefit as OKS decreased.

Preoperative OKS alone may not encapsulate the true clinical indications for TKR. Utilisation of OKS-based thresholds, may exclude equally eligible patients, risking greater inequity and variation in TKR practice. Restricting TKR to those with the worst preoperative OKS may compromise subsequent outcomes. Its role as an indicator of preoperative symptom burden and potential eligibility for TKR should be used with extreme caution.

# Table of Contents

<i>Acknowledgements</i> .....	<i>i</i>
<i>Abstract</i> .....	<i>ii</i>
<i>List of Figures</i> .....	<i>vii</i>
<i>List of Tables</i> .....	<i>x</i>
<b>CHAPTER ONE</b> .....	<b>1</b>
<b>1 Introduction</b> .....	<b>1</b>
1.1 The Burden of Osteoarthritis .....	1
1.2 Management Options in Knee Osteoarthritis .....	4
1.2.1 <i>Non-operative interventions</i> .....	4
1.2.2 <i>Operative Interventions</i> .....	5
1.3 Defining the Severity of Knee Osteoarthritis .....	6
1.3.1 <i>Clinical Assessment</i> .....	7
1.3.2 <i>Radiological Assessment</i> .....	9
1.3.3 <i>Symptom Severity and Eligibility for Joint Replacement</i> .....	10
1.4 Patient Reported Outcome Scores .....	13
1.5 Rationing and Threshold Criteria .....	15
1.6 Scope of Thesis.....	18
<b>CHAPTER TWO</b> .....	<b>21</b>
<b>2 The Profile of the Preoperative Oxford Knee Score – Defining Symptom Burden Prior to Knee Replacement</b> .....	<b>21</b>
2.1 Introduction .....	21

2.1.1	<i>Study Aims</i> .....	24
2.2	Methods .....	24
2.2.1	<i>Study Population</i> .....	24
2.2.2	<i>Database Preparation</i> .....	25
2.2.3	<i>Data Analysis</i> .....	27
2.3	Results .....	29
2.3.1	<i>Preoperative OKS distribution</i> .....	30
2.3.2	<i>Preoperative OKS profile by Surgeon</i> .....	36
2.4	Discussion.....	38
<b>CHAPTER THREE</b> .....		<b>44</b>
<b>3</b>	<b>Variation in the Preoperative Oxford Knee Scores – Analysis of the National Arthroplasty Database</b> .....	<b>44</b>
3.1	Introduction .....	44
3.1.1	<i>Study Aims</i> .....	45
3.2	Methods .....	46
3.2.1	<i>Study Population</i> .....	46
3.2.2	<i>Data Analyses</i> .....	47
3.2.3	<i>Multilevel Modelling</i> .....	50
3.3	Results .....	53
3.3.1	<i>Preoperative OKS by Centre</i> .....	57
3.3.2	<i>Preoperative OKS by Surgeon</i> .....	59
3.3.3	<i>Variation in the Oxford Knee Score</i> .....	62
3.3.4	<i>OKS Thresholds</i> .....	65
3.4	Discussion.....	66
<b>CHAPTER FOUR</b> .....		<b>71</b>

<b>4</b>	<b>Knee Replacement in the Presence of a High Preoperative OKS – An</b>	
	<b>Exploratory Pilot Study</b>	<b>71</b>
4.1	Introduction	71
	4.1.1 Study Aims	72
4.2	Methods	73
	4.2.1 Study population	73
4.3	Data Analysis	74
	4.3.1 Pain	75
	4.3.2 Function	75
	4.3.3 Consultation Process	76
	4.3.4 Final intervention	77
	4.3.5 Statistical Analyses	77
4.4	Results	78
	4.4.1 Patient Characteristics	78
	4.4.2 Pain	79
	4.4.3 Function	80
	4.4.4 Additional Factors	81
	4.4.5 Final Intervention	82
4.5	Discussion	83
	<b>CHAPTER FIVE</b>	<b>88</b>
<b>5</b>	<b>Preoperative Symptom Burden and TKR Outcomes– Less to gain with a higher</b>	
	<b>preoperative OKS?</b>	<b>88</b>
5.1	Introduction	88
	5.1.1 Study Aims	90
5.2	Methods	90

5.2.1	<i>Study Population</i> .....	90
5.2.2	<i>Data Analyses</i> .....	91
5.2.3	<i>Descriptive summary statistics</i> .....	94
5.3	Results .....	96
5.3.1	<i>Postoperative OKS and Change in OKS</i> .....	99
5.3.2	<i>Secondary postoperative outcomes</i> .....	103
5.4	Discussion.....	108
<b>CHAPTER SIX.....</b>		<b>113</b>
<b>6</b>	<b>General Discussion &amp; Conclusions.....</b>	<b>113</b>
6.1	Summary of the problem .....	113
6.2	Summary of Main Findings.....	114
6.3	Clinical relevance .....	116
6.3.1	<i>Understanding the normal ‘footprint’</i> .....	116
6.3.2	<i>Guiding Treatment Decisions for Patients</i> .....	119
6.3.3	<i>Preoperative Symptom Burden and TKR Thresholds</i> .....	120
6.3.4	<i>The Impact of OKS Thresholds on Postoperative Outcomes</i> .....	125
6.4	Future directions .....	128
6.5	Final Conclusions .....	130
<i>Appendix</i> .....		<i>132</i>
A.	Oxford Knee Score .....	132
B.	List of Abbreviations .....	134
<i>Bibliography</i> .....		<i>135</i>

## List of Figures

Figure 2-1	Representation of a positive skew to the preoperative OKS that might be expected in the presence of a significant preoperative symptom burden.....	23
Figure 2-2	Representation of possible variations in profiles of the preoperative OKS that may exist in different sub-populations a) similar preoperative OKS profiles; and b) different preoperative OKS profiles .....	23
Figure 2-3	Flowchart summarising the data flow for the datasets in each of the three centres. ....	27
Figure 2-4	Histogram of preoperative OKS in the three centres for patients undergoing a UKR .....	32
Figure 2-5	Histogram of preoperative OKS in the three centres for patients undergoing a TKR .....	32
Figure 2-6	Probability plots (P-P and Q-Q plots) examining normality of the preoperative OKS distribution for TKR in: a) Centre 1; b) Centre 2; and c) Centre 3.....	34
Figure 2-7	Probability plots (P-P and Q-Q plots) examining normality of the preoperative OKS distribution for UKR in: a) Centre 1; b) Centre 2; and c) Centre 3.....	35
Figure 2-8	Boxplot of preoperative OKS for UKR and TKR according to individual surgeons in Centre 1. Outlying data points reflected by $\circ > 1.5$ IQR. ....	37
Figure 2-9	Boxplot of preoperative OKS for UKR and TKR according to individual surgeons in Centre 2. Outlying data points reflected by $\circ > 1.5$ IQR ). ....	37
Figure 2-10	Boxplot of preoperative OKS for UKR and TKR procedures, according to individual surgeons in Centre 3. Outlying data points reflected by $\circ > 1.5$ IQR, * $> 3$ IQR.....	38
Figure 3-1	Flowchart summarising the data extraction and selection process for the study cohort .....	49
Figure 3-2	Histogram of the preoperative OKS distribution in the study cohort suggesting an overall normal distribution profile .....	54

Figure 3-3	Probability plots demonstrating normal distribution to the preoperative OKS. a) Q-Q plot, and b) P-P plots.....	55
Figure 3-4	Boxplots demonstrating the preoperative OKS distribution within individual orthopaedic centres. Circles represent outlier cases .....	58
Figure 3-5	Distribution of the preoperative OKS (interquartile range) according to individual surgeons. Solid lines represent the trend line for each respective value. ....	60
Figure 3-6	Distribution of the preoperative OKS (minimum to maximum OKS values) according to individual surgeons. Dashed lines represent the overall IQR for the combined cohort. ....	61
Figure 3-7	Proportion of TKR patients within individual centres presenting for TKR with preoperative OKS above proposed threshold values .....	65
Figure 5-1	Flowchart summarising the data extraction and selection process leading to the final study cohort.....	93
Figure 5-2	Histogram of preoperative OKS demonstrating the normal distribution profile.....	98
Figure 5-3	Histogram of postoperative OKS demonstrating a negative skew to the data, reflecting improvement in postoperative status .....	98
Figure 5-4	Histogram of postoperative change in OKS following TKR demonstrating a normal distribution. Dashed line represents no change in OKS, showing that most patients achieve some improvement postoperatively. ....	99
Figure 5-5	Boxplots demonstrating the distribution of the postoperative OKS in relation to preoperative OKS. Circles and stars represent outlying data points( $\diamond > 1.5$ IQR, $*$ $> 3$ IQR). Dashed line represents mean postoperative OKS (34.1).....	100
Figure 5-6	Boxplots demonstrating the distribution of postoperative change in OKS in relation to the preoperative OKS. Dashed line represents no change in OKS postoperatively. Circles	

represent outlying data points( $\circ > 1.5$  IQR,  $* > 3$  IQR). Dashed line represent postoperative OKS change=0..... 101

Figure 5-7 Histogram demonstrating the percentage of patients achieving improvement in OKS following TKR, according to their preoperative OKS. As preoperative OKS increases a greater proportion of patients achieve no postoperative improvement in their OKS. .... 102

Figure 5-8 Scatter plot demonstrating the relationship between preoperative OKS and postoperative change in OKS. Linear regression trend lines demonstrate the difference in the relationship for those patients with preoperative OKS  $< 16$  (blue) compared to those with preoperative OKS  $\geq 16$  (green). .... 104

Figure 5-9 Histogram demonstrating percentage of patients reporting a successful outcome in relation to the preoperative OKS. As preoperative OKS increases a greater proportion of patients report a successful outcome..... 105

Figure 5-10 Histogram demonstrating the percentage of patients reporting an improvement in knee symptoms following TKR according to the preoperative OKS. As preoperative OKS increases a greater proportion of patients report improvement in knee symptoms. .... 106

## List of Tables

Table 1-1	Common Clinical Signs of Osteoarthritis .....	8
Table 1-2	American College of Rheumatology criteria for classification of knee osteoarthritis [36] .....	8
Table 1-3	Summary of the Kellgren-Lawrence, and the Ahlbäck radiological classification systems of osteoarthritis .....	10
Table 2-1	Demographic characteristics of patients in each of the three centres according to procedure performed .....	31
Table 2-2	Summary descriptive statistics for the preoperative OKS in each of the three centres, according to procedure performed. ....	33
Table 3-1	Summary statistics for the preoperative OKS data from the national dataset .....	53
Table 3-2	Summary of the sociodemographic characteristics of the study cohort, and unadjusted regression coefficients from univariate analyses in relation to the preoperative OKS .....	56
Table 3-3	Summary of the multilevel model process leading to the final model. $\chi^2$ Statistic relates to comparison of each successive model to the previous model. ....	62
Table 3-4	Summary of the impact of sociodemographic characteristics on preoperative OKS in the adjusted final multilevel model. ....	64
Table 4-1	Summary of the patient characteristics in each of the three OKS groups .....	79
Table 4-2	Summary of the pain related characteristics according to the preoperative OKS .....	80
Table 4-3	Summary of the function related characteristics according to the preoperative OKS .....	81
Table 4-4	Summary of the preoperative consultation process according to OKS group .....	82
Table 4-5	Summary of the final intervention performed according to OKS group .....	83

Table 5-1	Summary of the sociodemographic characteristics of the study population .....	96
Table 5-2	Summary statistics for the preoperative OKS, postoperative change in OKS, and postoperative OKS in the study population .....	97
Table 5-3	Summary of results from the logistic regression model examining relationship between preoperative OKS and the whether the procedure was deemed successful. Results are presented for both unadjusted and adjusted models presented using Odds Ratios (OR) of achieving a successful outcome .....	107
Table 5-4	Summary of results from the logistic regression model examining relationship between preoperative OKS and reported improvement in knee symptoms postoperatively. Results presented for both unadjusted and adjusted models presented using Odds Ratios(OR) ....	108

# Chapter One

## 1 Introduction

### 1.1 The Burden of Osteoarthritis

Osteoarthritis (OA) represents one of the commonest musculoskeletal conditions resulting in disability worldwide, particularly in the older population. It is a condition characterised by progressive degenerative joint changes, ultimately leading to pain and functional impairment. True estimates of the incidence and prevalence of OA are difficult to achieve due to the variety of OA definitions reported in the literature [1]. It is estimated that in the UK alone a total of 8.75 million people, and a third of those aged 45 and over, have sought treatment for OA [2]. Globally, musculoskeletal conditions as a whole account for 21.3% of all years lived with disability (YLD), with OA contributing 17.1 million YLDs [3]. They also account for the 4<sup>th</sup> greatest impact on global health populations, considering death and disability, accounting for 6.8% of all disability-adjusted life years (DALY), just behind mental health disorders (7.4%), neoplasms (7.8%), and cardiovascular disease (11.4%) [4]. Between 1990-2010 disabilities secondary to musculoskeletal conditions as a whole increased by 45% worldwide and by 15.1% in the UK secondary to OA specifically [5].

Although OA commonly affects a range of different joints, the significant impact of pain and functional impairment on the larger weight bearing joints such as the hip and knee results in a considerably greater clinical impact [6]. The Global Burden of Disease 2010 study ranked hip and knee OA as the 11<sup>th</sup> highest contributor to global disability, and 38<sup>th</sup>

highest in terms of disability-adjusted life years [7]. Overall global knee OA prevalence was estimated at 3.8%. In the UK, it is estimated that 4.7 million people have sought treatment for knee OA [2]. Approximately half of patients seeking treatment of OA in the UK will have knee OA, and this is commoner in females. In the presence of an ageing population and increasing life expectancy, a significant growth in both incidence and prevalence of knee OA is anticipated in the next few decades [8]. This is further compounded by the additional impact of the increasing obesity and physical inactivity epidemic on both the incidence and management options in knee OA in particular [9]. The lifetime risk of symptomatic knee OA is estimated as high as 44.7%, and even higher in those with history of preceding knee injuries or associated obesity [10,11]. It is predicted that the number of people seeking treatment for knee OA in the UK is likely to reach 8.3 million by 2035 [2]. In addition to the general clinical impact, there is also a considerable economic burden associated with knee OA.

The economic burden can be considered according to both the direct and indirect costs [12]. Direct costs include what is classically considered the costs associated with healthcare resource utilisation in the treatment of OA. These include elements such as surgical costs, hospital resources, caregiver time, pharmacological & non-pharmacological treatment, cost of treatment side effects, and research [13]. Indirect costs of OA are also significant, reflecting the societal and economic burden of the disease as a result of factors including; loss of productivity, absenteeism, premature mortality, and disability benefits. Time lost from leisure, together with the dependence on unpaid informal caregivers has also been shown to account for significant indirect costs specifically in relation to knee OA [14,15].

Ruiz et al recently examined the impact of both direct and indirect cost in relation to the treatment of knee OA [12]. Although the lifetime direct costs of OA were \$20,635 per patient undergoing total knee replacement (TKR), the societal savings in lower indirect costs and functional improvement were \$39,565. This reflects a net benefit of \$18,930 per patient, which in the US reflects a \$12billion annual saving based on 2009 figures. Leardini et al found that indirect costs (€1236 per patient per year) accounted for 57% of total knee OA costs (€2170 per patient per year) [15]. Indirect costs increased linearly with worsening radiological knee OA, with increasing dependence on informal caregivers as the primary source. Gupta et al investigated the arthritis-attributable costs amongst patients with disabling hip and knee OA [14] Indirect costs were five times greater than direct costs (\$12,200 vs. \$2,200). These were more likely to be incurred amongst females, and those aged 75 or over. However, actual costs incurred were higher in those aged under 65, and when reported expenditure for men was 25% higher. There is also recognition that intangible costs such as pain and suffering, decreased QOL, potential depression/anxiety, which are generally unrecognised, may also be important, although not extensively examined [16]. Determining the true economic burden and economic cost effectiveness of interventions is therefore difficult.

There is a recognised lack of published data specifically relating to the UK population, [13] but the impact on the UK economy is thought to equate to approximately 1% of the gross domestic product. A recent review of available UK data suggested that pharmacological treatment and their potential adverse effects accounts for over £150million annually (all OA), whilst surgical intervention for knee OA totals over £426million [13]. Economic cost through lost working days as a result of OA was estimated at over £3.2billion, whilst disability living allowance payments totalled £2.4

billion. Knee-specific estimates are not available, but knee OA is likely to be a significant contributor to this total. Although accurate estimates of the true prevalence and the full costs of knee OA are difficult to ascertain there is little doubt that the overall impact and burden of OA is immense and rapidly growing, reflecting an important clinical condition worthy of ongoing dedicated research.

## **1.2 Management Options in Knee Osteoarthritis**

Several treatment modalities are described in the literature for knee OA, incorporating the whole spectrum of disease severity [17]. Many of those patients identified, with early stage OA will remain asymptomatic though, and may require very little intervention other than long term review of symptoms and reassurance. As the disease process progresses, with symptoms and functional impairment becoming more dominant, formal intervention is often eventually required.

### **1.2.1 Non-operative interventions**

The vast majority of knee OA is managed in the primary care setting, and in the UK the National Institute for Health and Care Excellence (NICE) have recently published updated, clear and structured guidelines for the primary care management of OA [18]. Non-operative intervention forms the mainstay of treatment for most patients. These guidelines include core treatments for all patients including; access to appropriate information regarding the condition, advice to encourage activity and exercise, and interventions to achieve weight loss if the patient is overweight. Additional non-pharmacological and pharmacological treatment is recommended according to individual

needs and preferences. These include manual therapy (e.g. physiotherapy), supports and braces, shock absorbing shoes or insoles, local heat and cold therapy, non-steroidal anti-inflammatory medication (topical or oral) or COX-2inhibitors (with a proton pump inhibitor), and opioid medication. For those patients with persistent pain and disability not responding to 3 months of evidence based non-surgical treatment it is recommended that secondary care referral is instigated for consideration of surgical intervention.

### **1.2.2 Operative Interventions**

A number of different surgical options have been described for the varying stages of knee OA. Knee arthroscopy has traditionally been utilised as a common diagnostic and bridging intervention to assess and aid with management of the symptoms of worsening knee OA. Recent evidence has however shown this to be ineffective in treating advanced OA, and its role is now limited to selected cases with mechanical symptoms [19]. In the younger high demand patient, realignment procedures such as a high tibial osteotomy may be an option. This currently though remains a low volume procedure, undertaken by a select group of surgeons, due to the technical demands and limited indications, which exclude the general knee OA population [20]. At present there is also only limited data relating to its outcomes and cost effectiveness [21].

Total knee replacement is an established and highly successful intervention for the pain and disability associated with end-stage OA [20,22]. Technological advancements, improved survival rates, and increasing demand have seen significant expansion in the indications for knee arthroplasty to include both younger active patients as well as frailer elderly patients [23]. Over 75,000 procedures were undertaken in the UK in 2011,

representing a growing trend [22,24]. In the United States, TKR growth is predicted to increase by 673% by 2030 [25].

A key question relating to knee arthroplasty is that of timing and indications [26,27]. Traditionally, knee arthroplasty was restricted to those with severe pain and disability, and radiological evidence of end-stage OA, with the primary aim of pain relief [23]. It is evident though that there is a lack of consensus on the exact indications and the defined disease severity that are associated with eligibility for surgery [27-31]. Attempts have been made to develop clear guidelines and criteria to define surgical eligibility [32]. Accurately measuring and defining the severity of knee OA in order to determine when TKR is indicated or not, is however extremely difficult.

### **1.3 Defining the Severity of Knee Osteoarthritis**

Osteoarthritis encompasses a broad and complex spectrum of symptoms, clinical signs, pathological changes, and radiological features all of which help define the disease process, with great variation between individuals [33]. Consequently, it is difficult to generate reliable and accurate classification criteria to establish the diagnosis, and more importantly define the severity of disease at an individual level. This is of particular interest for researchers interested in identifying the earliest presentation of OA and comparing treatment effects. However, these criteria are equally important in clinical practice in aiding clinicians to make decisions regarding progressive management options such as secondary care referral and possible surgical intervention.

Numerous classification systems have been described [34-42], but all incorporate key criteria which include i) clinical signs and symptoms; ii) radiological appearances; and iii) pathological characteristics, either in isolation or in combination. Even established classification systems have problems, and the accuracy and applicability of each method are not without limitations [43,44].

### **1.3.1 Clinical Assessment**

The commonest clinical symptoms associated with knee OA is localised joint pain, usually in conjunction with activity. As disease progresses pain levels often increase, transitioning from activity related pain through to increasing rest and night pain, together with subsequent loss of function. Clinically, numerous findings are described secondary to OA, and often used to aid diagnosis and define severity of disease (Table 1-1). There is however considerable variation in the presence of clinical signs amongst patients, and in isolation may not be specific to OA, with crossover between other disease processes. More formal diagnostic criteria have therefore been developed both in the US (Table 1-2) [36] and Europe [37]. Although clearly valuable in defining diagnostic criteria, particularly for research purposes, these criteria provide very little objective measurement of the severity of OA, and are not routinely utilised in the clinical setting.

**Table 1-1 Common Clinical Signs of Osteoarthritis**

---

Localised swelling / Effusion
Crepitus
Reduced range of motion
Joint Deformity
Pain with passive motion
Joint instability
Muscle atrophy
Altered gait

---

**Table 1-2 American College of Rheumatology criteria for classification of knee osteoarthritis [36]**

Clinical and Laboratory	Clinical and Radiographic	Clinical
<b>Knee Pain + at least 5 of 9:</b>	<b>Knee Pain + at least 1 of 3:</b>	<b>Knee Pain + at least 3 of 6:</b>
Age >50 years	Age >50 years	Age >50 years
Stiffness > 30 minutes	Stiffness > 30 minutes	Stiffness > 30 minutes
Crepitus	Crepitus	Crepitus
Bone tenderness	+ osteophytes	Bony tenderness
Bony enlargement		Bony enlargement
No palpable warmth		No palpable warmth
ESR <40mm/hr		
Rheumatoid Factor <1:40		
SF OA		

---

### **1.3.2 Radiological Assessment**

Clinical signs and symptoms are considered the key criteria in the clinical environment, when making decisions regarding management options. However perceived symptom severity does not always correlate with the clinical disease severity and this is an important consideration when considering more complex and invasive surgical treatment options [45]. Radiographic classifications have been utilised to further define the extent of OA within affected joints. Two commonly employed radiographic criteria include the Kellgren-Lawrence [35] and the Ahlbäck [34] classifications (Table 1-3). Both systems have withstood the test of time and remain in common use; however they are both subject to limitations [46]. Similar to clinical based classifications, radiographic classifications in isolation do not necessarily reflect the true subjective burden of disease at an individual level [47]. Technological advancements have also seen a growing trend towards the role of more modern modalities such as MRI using the BLOKS and WORMS scores [39,40]. Although these offer opportunities to study OA at much earlier stages, they currently offer very little input in the clinical decision-making process associated with end-stage OA.

**Table 1-3 Summary of the Kellgren-Lawrence, and the Ahlbäck radiological classification systems of osteoarthritis**

<b>Kellgren-Lawrence [35]</b>		<b>Ahlbäck [34]</b>	
<b>Grade 0</b>	No changes	<b>Stage 0</b>	No radiographic sign of arthritis
<b>Grade 1</b>	Doubtful narrowing of the joint space and possible osteophytic lipping	<b>Stage 1</b>	Narrowing of the joint space (with or without subchondral sclerosis).
<b>Grade 2</b>	Definite osteophytes and possible narrowing of the joint space	<b>Stage 2</b>	Obliteration of the joint space
<b>Grade 3</b>	Moderate multiple osteophytes, definite narrowing of the joint space, and some sclerosis, and possible deformity of the bone ends	<b>Stage 3</b>	Bone defect/loss <5 mm
<b>Grade 4</b>	Large osteophytes, marked narrowing of the joint space, severe sclerosis, and definite deformity of the bone ends	<b>Stage 4</b>	Bone defect/loss between 5 and 10 mm
-	-	<b>Stage 5</b>	Bone defect/loss >10 mm, often with subluxation and arthritis of the other compartment

### 1.3.3 Symptom Severity and Eligibility for Joint Replacement

A key problem in relation to knee OA, is to determine the point in the evolution of the disease, at which the symptom burden is sufficient to justify knee replacement surgery.

In recent years TKR procedures are increasingly being performed at an earlier stage in the disease process, and amongst a changing population group [48]. There has traditionally been a significant unmet need for TKR due to limited access, which restricted surgery to those with severe disease [49]. Although demand still outstrips supply, improved access to surgery is likely to result in earlier surgery and a lesser preoperative symptom burden. Ongoing technological advancements together with improved outcomes have also resulted in expansion of the indications for TKR [23]. Finally, improved patient and primary care education about the management options for knee OA, and application of decision-making tools may result in greater demand and expectations together with earlier referral into secondary care [50].

In most cases the patient's own subjective assessment of symptom severity and personal thresholds will lead to the initial presentation to their general practitioner (GP) [51,52]. From this point forward the clinician's subjective evaluation of the overall symptom burden and assessment of objective features such as radiological changes will determine much of the subsequent management pathway including referral for joint replacement surgery.

The decision making process leading to TKR has evolved significantly over the years. Early TKR procedures were limited to older patients with end-stage radiological OA, in the presence of severe pain and functional restrictions. The decision to operate was based on subjective assessment of pain and disability together with an objective radiological analysis. As the indications for TKR have changed over the years so has the decision making process. There is significant lack of consensus recognised amongst both health professionals and patients as to the specific indications that determine the need for knee

replacement [27,29]. Although this decision making process is often complemented by more formal objective clinical and radiological assessment as described above, there remains poor correlation between symptom severity, radiological changes, and perceived eligibility for joint replacement [45,53,54]. Various TKR priority criteria scores have been developed in an attempt to standardise the decision making-process, although these have not been universally successful. Additional factors such as the patient-clinician interaction can also be an important determinant of progression to TKR, and the importance of the patient's involvement in a shared decision-making process is increasingly recognised [55].

These discrepancies have led to concerns regarding variability in selection criteria, and the potential for variation in access to arthroplasty services amongst those suffering with knee OA [56-58]. Riddle et al suggested that a third of patients undergoing knee replacement were deemed inappropriate [59]. A study of 7557 individuals diagnosed with hip or knee OA in Spain, found that only 16.0% of patients with knee OA were deemed appropriate for consideration of TKR [60]. There is also further evidence to suggest that appropriateness criteria for TKR are directly linked to postoperative outcomes, raising further concern about variation in the selection process [61,62]. Evaluation of preoperative status in patients awaiting hip replacement has confirmed significant variation in the symptom burden [63,64]. Detailed analysis of the preoperative status of patients proceeding to knee replacement has not yet been formally reported in the literature.

Determining the true burden of OA in individual patients is therefore not an easy task. For the vast majority of patients with OA in the presence of mild or no symptoms, this is

of little consequence. However, as disease burden progresses and decisions regarding need for more invasive interventions, including surgery come into play, accurate estimation of the true disease burden becomes increasingly important to achieve appropriate decision making for surgical intervention, and defining thresholds in the treatment ladder. As concerns grow about variation in healthcare provision there is a growing urgency to further standardise this selection process [23].

#### **1.4 Patient Reported Outcome Scores**

Traditionally, knee arthroplasty outcomes have been measured according to implant survival rates [65]. In recent years it has increasingly been recognised that implant survival alone does not necessarily reflect the true outcome from the patient's perspective [66]. Consequently there has been a growing trend towards the utilisation of alternative outcome measures, and in particular Patient Reported Outcome Measures (PROMs) to better define the outcome of arthroplasty procedures [67]. As the name suggests, the key element of PROMs is that they are completed by the patient themselves. These instruments usually measure the severity, or the particular impact of a condition as viewed by the individual. As such they are considered a better reflection of the patients own opinions regarding the success of treatment compared to objective clinician completed measures. PROMs are now well established in healthcare practice, and exist in two main types; generic health assessments and disease specific assessments.

Several knee specific scoring systems have been described, validated and now routinely reported in the arthroplasty literature [42]. One of the more commonly utilised measures is the Oxford Knee Score (OKS) [41]. This was first described in 1998, as a short and

simple disease-specific patient-administered questionnaire to assess the outcome of TKR in clinical trials. Although originally developed for use in clinical trials, its short and simple format has seen it popularised and accepted as an alternative PROM to monitor TKR outcomes in the general arthroplasty population. Since its introduction it has been extensively validated, translated to numerous languages and is now considered a fully validated and established measure of TKR outcome, which is routinely used worldwide [68-70]. The scoring system comprises a 12-item questionnaire, evaluating the two key domains of pain and function. Each item in the questionnaire contains 5 possible responses grading the relative impact of pain and disability in relation to various normal daily scenarios [Appendix 1]. Further modifications have been reported since its original description, with the current version based on a total score ranging from 0-48 [71]. Lower scores reflect worse outcome, and higher scores reflect better outcome. The OKS has also recently undergone further validation examining the individual pain and function subscales, as well as defining the clinically important difference and minimally significant change in OKS, further reinforcing its clinical utility [72,73].

Since 2009, the OKS has been adopted by the Department of Health as the primary outcome measure incorporated into the UK's national arthroplasty database. Data submission is mandatory for all knee arthroplasty procedures performed within the National Health Service (NHS) in England and Wales. As a result, the OKS is now an established and accepted component of the knee arthroplasty process in the UK.

Ease of administration, and wide acceptance has made outcome scores such as the OKS, particularly appealing for adoption into alternative roles outside of their original remit as measures of knee arthroplasty outcomes. Alternative roles for the OKS currently reported

in the literature include outcome monitoring following pharmacological interventions, rehabilitation interventions, osteotomies as well as fracture treatments [74-77]. There is also emerging validation data to support its role in monitoring response to non-operative intervention for knee OA, however it remains to be formally validated for most of the other reported applications [78,79]. Although it seems feasible that it may remain valid in monitoring treatment response across different clinical scenarios, there have been recent proposals encouraging its application in the pre-operative setting as a measure of disease severity and in turn a criterion to define eligibility for TKR.

## **1.5 Rationing and Threshold Criteria**

The combination of increasing demand and general technological advancements in health care treatment in the face of major financial restrictions, has led to significant pressures on health care resources. Savings of £20billion are required from the current UK health budget of £110billion [80]. The growing cost and demand of modern arthroplasty prostheses is a particular issue in the management of knee OA. Estimated UK expenditure on knee replacement in 2010 was approximately £426 million, from a total musculoskeletal budget of £10billion [13,81]. Consequently, publicly funded healthcare systems such as the NHS have seen increasing pressure to reduce expenditure and maximise cost-effectiveness, while maintaining an accessible, equitable and safe health system [81]. Rationing or more subtly stated, priority-setting criteria, have consequently become increasingly evident in planning healthcare resource allocation [82]. In a recent survey of NHS trusts, it was found that 49% of trusts had seen increasingly restrictive criteria being applied to arthroplasty referrals [83].

Rationing and prioritisation of elective surgical procedures such as knee replacement is not an entirely new concept. Discrepancies between supply and demand have plagued publicly funded healthcare systems worldwide, resulting in problems with long waiting times for surgery [84-86]. As a result a number of different rationing approaches have been described [87]. The first approach is described as rationing by exclusion, with low-priority interventions completely excluded from health service provision. A second approach is to ration by guidelines, whereby nationally agreed guidelines or criteria are developed for specific conditions and interventions. This ensures that treatment is targeted towards patients most likely to gain benefit, thus providing the greatest level of cost effectiveness. This approach has already been adopted by similar healthcare systems in Canada, Spain, and New Zealand [88-90]. These systems utilise scoring criteria developed specifically to define prioritisation and eligibility for TKR using recognised methods such as the Delphi process. Despite extensive time spent developing specific criteria these have not been problem free, and highlight the difficulties in formally applying eligibility criteria for surgery [91,92].

Although recent NICE guidelines have sought to offer guidance as to the pathway and appropriate management of knee OA, these remain open to variation in subjective assessments [18]. In the UK, no nationally agreed threshold or priority scoring system have been developed or adopted for routine use, and such decisions have generally been undertaken regionally under the guidance of clinical commissioning groups (CCG). This has been described as a third approach to rationing – rationing by muddling through [87]. Neither the lack of an accepted threshold scoring system or the accepted clinical and cost effectiveness of knee arthroplasty, has however prevented primary care trusts (PCT), and

more recently CCGs from pursuing proposals to ration certain interventions, often using erroneous data and methods [93].

As a validated measure of pain and function, which is easily administered, and already widely accepted in the orthopaedic community, the OKS has found itself a prime candidate for possible application as a threshold criterion to further define surgical eligibility [94-100]. A recent review of commissioning policies undertaken by the Royal College of Surgeons of England found that 44% (23 of 52) of CCGs had already instituted arbitrary referral criteria for hip replacement [93]. Furthermore, 30% (16 of 52) of CCGs were routinely applying the similarly structured Oxford Hip Score (OHS) as part of their commissioning policy for THR. This review did not include analysis of knee replacement policies, but it is likely that most CCGs will have similar criteria in place for both hip and knee replacements. Unfortunately, there is no evidence to support or validate the OKS (or the OHS) in this role. Likewise, there is no evidence to define a specific OKS value that could be considered appropriate as a threshold cut-point. As a result, there is considerable variation in the range of OKS values that have been proposed for use, ranging from 16-32, all of which are completely arbitrary in nature. Concerns about commissioning policies have been raised, but in the absence of any evidence to challenge this role it is likely to continue to be utilised [101,102]. In fact NICE guidelines state that in considering joint replacement clinicians should "...base decisions on referral thresholds on discussions between patient representatives, referring clinicians and surgeons, rather than using scoring tools for prioritisation" [18] .

## **1.6 Scope of Thesis**

It is evident from the current literature that knee OA is an extremely common and important condition. Although extensively studied there remains a number of highly significant issues to be explored, of relevance to patients, clinicians, as well as healthcare managers and policymakers. Accurately defining the symptom burden of knee OA is difficult, and although pain and functional loss are considered important factors, the true symptom profile amongst those undergoing knee replacements is poorly described in the current literature. There is also no clear consensus as to the indications that define eligibility and relative need for knee replacement. In the longstanding cost-constrained health economy it seems inevitable however that rationing will be remain prominent in some form or another and that defined criteria will be utilised to determine access to arthroplasty procedures. A greater understanding of the preoperative symptom burden amongst the knee replacement population and how this relates to the pre-operative OKS is therefore essential, particularly if prioritisation of patients becomes widely adopted.

Review of the literature has identified that our current understanding of the preoperative symptom burden in the knee replacement population remains poor. Furthermore, there is very little evidence defining how the preoperative OKS relates to preoperative symptom burden, or what constitutes typical arthroplasty practice. The scope of this thesis is to explore, and gain a greater understanding of how the preoperative OKS reflects current arthroplasty practice in the UK. It will aim to gain further insight into the preoperative symptom burden of patients proceeding to knee replacement and to evaluate whether variation in preoperative status, as measured by the OKS, exist in the knee replacement population. The thesis will also examine how variation in preoperative OKS may impact

on subsequent postoperative outcomes. With the increasing application of the preoperative OKS as a potential indicator of surgical eligibility this information, it will explore the impact of these findings in relation to proposed preoperative OKS thresholds. This will provide greater appreciation of how OKS based thresholds may impact on current arthroplasty practice, and identify further possible research in relation to this topic.

Chapter 2 will initially examine the preoperative OKS of patients undergoing knee replacement across three independent NHS orthopaedic centres. The aim is to define the distribution of the preoperative OKS in the population, to characterise the preoperative symptom burden, and gain an insight into how this relates to current arthroplasty practice in the UK.

Chapter 3 will explore the variability in preoperative OKS amongst the knee replacement population across the NHS, using data from the national PROMs and HES databases. The aim is to determine the extent of variation in practice across different healthcare providers, and highlight the potential impact of proposed threshold criteria in these populations.

Chapter 4 describes an exploratory pilot analysis of a single centre patient cohort, evaluating the possible reasons why TKR is undertaken in the presence of a high preoperative OKS, and to evaluate whether the decision to pursue TKR is a true reflection of the clinical presentation in relation to pain and function, or reflects a possible breakdown in the consultation process.

Chapter 5 will build on the work of the earlier chapters to examine the detailed relationship between preoperative OKS and subsequent postoperative outcome, with specific focus on those patients undergoing surgery with high preoperative OKS, further evaluating the relationship between preoperative symptom burden and subsequent postoperative outcomes.

In Chapter 6 the main findings of the thesis will be reviewed. The clinical relevance of the findings from the previous chapters will be explored, together with identification and suggestions for further research that may be required to supplement the findings presented in this thesis.

## **Chapter Two**

# **2 The Profile of the Preoperative Oxford Knee Score – Defining Symptom Burden Prior to Knee Replacement**

## **2.1 Introduction**

Despite a lack of consensus regarding the indications for knee replacement, pain and disability which has failed to respond to non-operative intervention, are generally accepted as key indications [23]. However it remains unclear as to the point at which this pain and disability becomes severe enough to justify surgical intervention [45]. Although no formal threshold scoring criteria have been developed in the UK, a number of clinical commissioning groups have developed local referral criteria, some of which have incorporated the preoperative OKS as a key component [95-100,103-105]. Such criteria are based on an assumption that a lower OKS equates to a greater symptom burden, and thus a higher priority for surgery, and in turn greater potential to benefit.

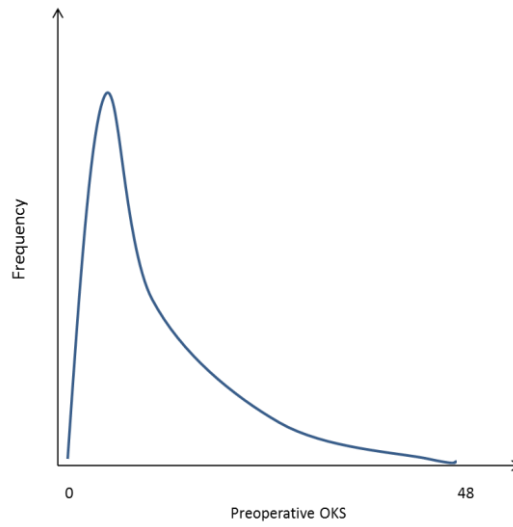
Despite the widespread use of the OKS in published outcome studies, and its general utilisation in monitoring arthroplasty outcomes, there remains a paucity of published data specifically relating to the preoperative OKS within the knee replacement population [106-108]. In addition, the preoperative symptom burden of patients undergoing knee replacement has not been explored objectively in any great detail, and thus we have very little understanding of what constitutes ‘typical’ arthroplasty practice in the UK.

Applying threshold criteria without an appreciation of how they relate to current practice, risks causing more problems than are solved.

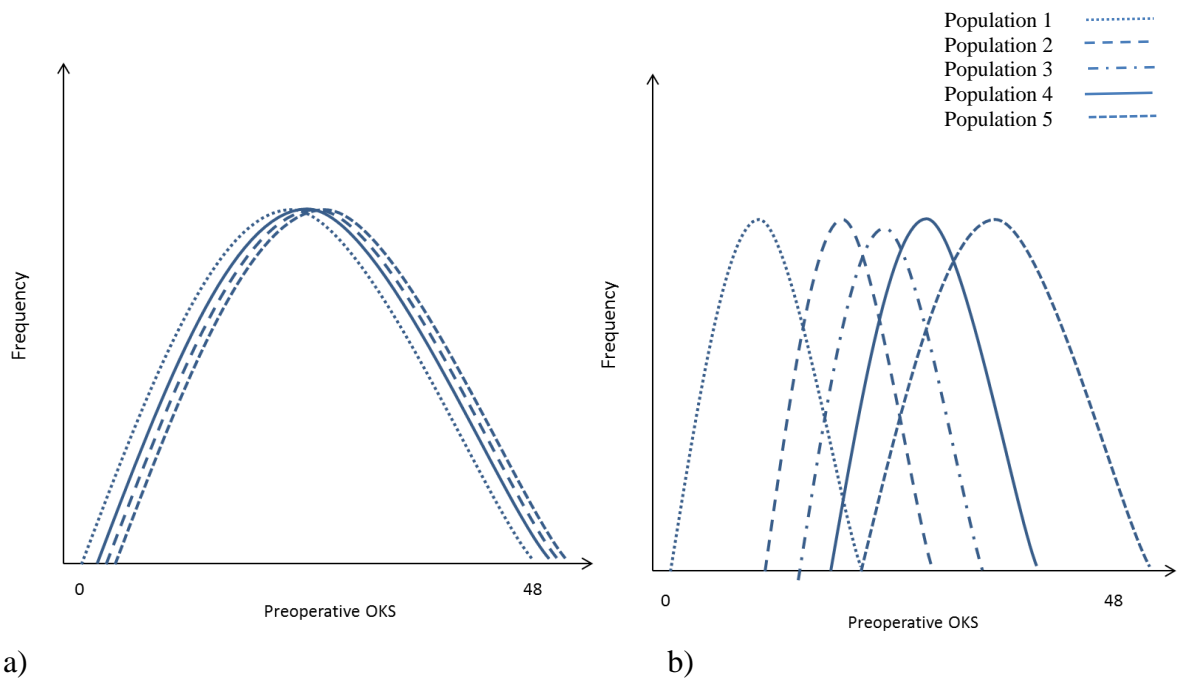
If it is accepted that pain and disability are key indications, then we might expect this to be reflected in the preoperative OKS, by a preponderance of cases with lower OKS values, and a highly skewed distribution (Figure 2-1). Limited data from the initial UK PROMs report however, demonstrated that the preoperative OKS appeared to conform to a normal ‘Gaussian’ distribution profile [109]. Thus the majority of patients appear to have a lower symptom burden than might have been predicted. It also suggests that there may be a sizeable subgroup of patients, represented by a higher preoperative OKS, proceeding to knee replacement with very minimal symptom burden. The data presented in this report was however restricted to a single simple graphical representation, with no further detailed analysis of the data reported. As an overall representation of the national data, it remains unclear whether this is an accurate representation of the preoperative symptom burden across different providers, or whether it may reflect the amalgamation of a number of different preoperative OKS profiles which would be indicative of possible variation in selection criteria (Figure 2-1 and Figure 2-2).

With the on-going growth in arthroplasty demand, an increasing focus on reducing healthcare variation, and the move towards healthcare rationing there is an urgent need to gain a greater understanding of the preoperative symptom burden in patients proceeding to knee arthroplasty. Specifically, in the UK, it is also essential to determine how this relates to the preoperative OKS, and before implementing broad changes in access to joint replacement it seems rational to gain a greater understanding of current practice.

**Figure 2-1 Representation of a positive skew to the preoperative OKS that might be expected in the presence of a significant preoperative symptom burden**



**Figure 2-2 Representation of possible variations in profiles of the preoperative OKS that may exist in different sub-populations a) similar preoperative OKS profiles; and b) different preoperative OKS profiles**



### **2.1.1 Study Aims**

The aims of this study were to examine the distribution profile of the preoperative OKS in knee replacement patients in three separate orthopaedic centres within the UK in order to:

1. Explore and characterise the preoperative symptom burden of patients undergoing knee replacement; and
2. Examine whether the preoperative symptom burden of patients was similar across different providers, including:
  - Centres; and
  - Surgeons
  - Arthroplasty procedure (UKR and TKR)

## **2.2 Methods**

### **2.2.1 Study Population**

This study was based on three separate patient cohorts that had undergone knee replacement surgery at one of three tertiary elective orthopaedic centres. Centres included in the study were i) Nuffield Orthopaedic Centre, Oxford (Centre 1); ii) The Elective Orthopaedic Centre, Epsom (Centre 2); and iii) Musgrave Park Hospital, Belfast (Centre 3). All three centres represented NHS hospitals undertaking pure elective orthopaedic procedures within the UK and Northern Ireland. Within each unit there was

an established multi-surgeon, high-volume primary joint arthroplasty practice providing a representative population of patients for analysis.

Additionally, all units had independently established prospectively collected arthroplasty outcome programmes, which included routine measurement of the preoperative OKS. This data was collected prior to surgery, during the preoperative assessment process within respective centres. Although the OKS was routinely collected preoperatively, the decision to proceed to knee replacement was based on individual surgeons' clinical assessment on a case by case basis, and no formal policies existed at any of the centres that required the preoperative OKS to be utilised to guide treatment decisions at the time of the study.

Due to the independent nature of the three databases, the data collected within the respective arthroplasty outcome programme was not homogenous, and thus there were restrictions to the amount of data that could be combined for analyses across all three datasets. Each database had also been established for different intervals, and at the time of this study the OKS data available for analysis comprised the following time periods; Centre 1 – November 2008 to August 2011; Centre 2 – January 2004 to June 2011; Centre 3 – January 2004 to January 2011.

### **2.2.2 Database Preparation**

Each database was interrogated individually to identify the relevant study population. Full validation of each database to determine the compliance and accuracy of the dataset in relation to TKR practice was not possible. Data was extracted from each database and

subjected to initial data cleaning process to identify database errors, followed by application of inclusion and exclusion criteria. Inclusion criteria for this study included:

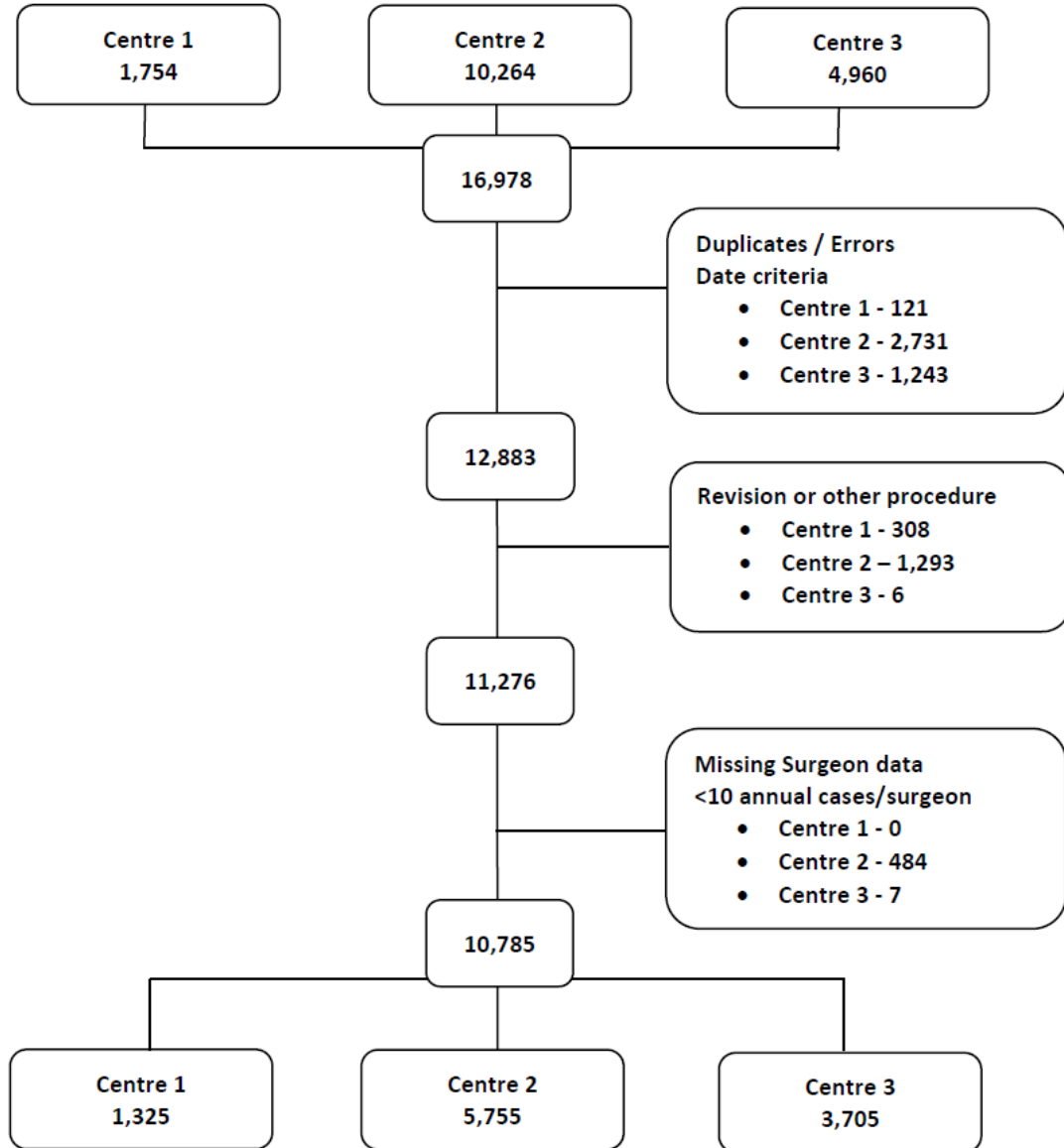
- Primary knee replacement procedure – TKR or unicompartmental knee replacement (UKR)
- Completed preoperative OKS recorded in the database

Exclusion criteria included:

- Revision knee replacement or non-arthroplasty procedure
- Cases with missing or incomplete OKS data
- Cases with missing surgeon identifier
- Surgeons with less than 10 cases/year
- Missing operative data, preventing confirmation that a procedure was performed

Where duplicate entries were identified, the most recent entry was retained, and duplicates excluded. Only age, gender and procedure type (TKR or UKR) were available and complete across all three databases and included in further analyses. Where missing or erroneous data was encountered, cases were cross-referenced with individual electronic records when possible to minimise data loss. Following completion of initial data cleaning, and application of the exclusion criteria, a total 10,785 patients remained available for inclusion in the study (Centre 1 = 1325 cases; Centre 2 = 5755 cases; Centre 3 = 3705 cases).

Figure 2-3 Flowchart summarising the data flow for the datasets in each of the three centres.



### 2.2.3 Data Analysis

Statistical analyses were performed using SPSSv18 (IBM, Armonk, New York). The key aim of this study was to describe the profile and distribution of the preoperative OKS within the respective population. Description of a set of data points within a defined

population can be undertaken using a combination of individual summary statistics and graphical representation (histograms, boxplots). Summary statistics include; measures of central tendency (Mean, Mode, Median); measures of dispersion (Standard deviation (SD), variance, percentiles); and measures of shape (skew, kurtosis). As well as providing a description of the data and the distribution around the central values, they enable estimation of whether the data conforms to a normal distribution (bell-shaped) curve. Simple statistics such as the SD provide some information about dispersion of data in isolation. When comparing dispersion in different populations it is only meaningful in the context of the mean value, and can therefore be misleading. The coefficient of variation (CV), expressed as a ratio of the SD to the mean, provides a measure of the relative dispersion, enabling direct comparison of data dispersion between populations.

Supplemental information can also be obtained utilising histograms, boxplots, and probability plots (Q-Q and P-P plots) to further identify deviations in the data from a normal distribution [110]. These methods incorporate key components reflecting the distribution properties of the data, which in combination create a visual description that defines the data, and permits an understanding of variations in different populations. In the case of the P-P plot, the cumulative probability of the variable (preoperative OKS) is plotted against the cumulative probability of a normal distribution. Similarly, for the Q-Q plots, the values seen in the dataset (observed values) are plotted against the normal distribution (expected values) after creating quantiles. Both methods generate plots that help determine whether the data is normally distributed or not. Deviation from normality can subsequently be further evaluated depending on the particular distribution of data observed. Although formal statistical tests of normality are available such as the Shapiro Wilk, and the Kolmogorov Smirnov tests, these were not utilised in this study due to the

large sample size of the study cohorts, and the recognised problems with minor deviations from normality in such samples [110]. In order to gain a comprehensive understanding of the preoperative OKS profile, the statistical summaries described above were utilised to establish the distribution of the preoperative OKS within each of the three centres.

Due to the high level of UKR practice anticipated within the Oxford cohort, separate analyses were undertaken within each centre reflecting UKR and TKR practice. Data was examined for differences in patient characteristics between centres using ANOVA for continuous variables (age) and the  $\chi^2$  test for categorical data (gender). Preoperative OKS was compared both between centres, and between surgeons within the same centres using simple univariate analyses, followed by incorporation of age and gender into a multivariate model as potential explanatory variables. A p-value of <0.05 was considered significant throughout.

## **2.3 Results**

Mean preoperative OKS in the combined patient cohort was 17.9 (95% confidence interval (CI) 17.8-18.1, Range 0 to 48). The overall mean preoperative OKS was significantly greater for patients in the UKR compared to the TKR group (TKR 17.6, 95% CI 17.5-17.8; UKR 20.3, 95% CI 19.9-20.8;  $p < 0.001$ ). Further sub-analysis within the individual centres demonstrated that the mean preoperative OKS was also higher in the UKR compared to TKR cohorts in all three centres ( $p < 0.001$  in all centres).

No significant differences in patient characteristics (age, gender) were observed between centres with regards to UKR procedures. However significant differences were evident in age and gender distribution for the TKR group (Table 2-1)

### **2.3.1 Preoperative OKS distribution**

The overall distribution of the preoperative OKS in each of the three centres is depicted for UKR (Figure 2-4) and TKR (Figure 2-5). These demonstrate a wide range of preoperative OKS values amongst patients proceeding to knee replacement in all three centres, with a similar general profile evident for both UKR and TKR cases. The overall profile appeared to conform to a normal distribution curve in all three centres. In centre 3, there was evidence of a complete shift of the curve to the left, reflecting overall lower OKS compared to Centres 1 and 2.

Exploration of the summary descriptive statistics (Table 2-2) showed that Centres 1 & 2 exhibited a similar preoperative OKS profile to each other, for both UKR and TKR procedures. Centre 3 demonstrated a significantly lower overall mean OKS compared to the other two centres (Centre 1; 19.9 (95%CI 19.5-20.3) & Centre 2; 19.9(95%CI 19.7-20.1) vs. Centre 3; 14.3(95%CI 14.0-14.5),  $p=0.001$ ). A difference of up to 6.1 points was observed in mean OKS for the UKR group, and up to 5.6 points for the TKR group between centres (ANOVA  $p<0.001$ ). The mean range of OKS values was 47 points. The overall profile was similar for both UKR and TKR groups, although the UKR group exhibited slightly less variability in preoperative OKS values.

The summary statistics suggested that both Centre 1 and 2 conformed to a normal distribution curve. Data from Centre 3 showed greater skew values suggesting a trend towards a positive skew, although measures of central tendency were similar. Further analysis of the probability plots confirmed the normal distribution in Centre 1 & 2, whereas overall profile in Centre 3 showed a suggestion of a positive skew. Deviation from normality appeared to be influenced by the presence of a small number of high value outliers. A total of 28 (0.008%) cases in Centre 3 exhibited OKS values greater than 3 SD from the mean, representing outlying cases.

**Table 2-1 Demographic characteristics of patients in each of the three centres according to procedure performed**

		Centre 1	Centre 2	Centre 3	p-value
<b><i>UKR</i></b>					
<b>Number of cases (%)</b>		605 (45.6)	445 (7.7)	266 (7.2)	-
<b>Age (Years)</b>	Mean	65.9	65.5	66.8	0.256
	(95% CI)	(65.1-66.7)	(64.6-66.4)	(65.6-68.0)	
	Range	23.7 - 87.2	41.5 - 89.5	38.0 - 89.0	-
<b>Gender n (%)</b>	Male	281 (46.4)	220 (49.4)	116 (43.6)	0.307
	Female	324 (53.6)	225 (50.6)	150 (56.4)	
<b><i>TKR</i></b>					
<b>Number of cases (%)</b>		720 (54.4)	5310 (92.3)	3439 (92.8)	-
<b>Age (Years)</b>	Mean	70.6	71.6	69.5	<0.001
	(95% CI)	(69.9-71.3)	(71.4-71.9)	(69.2 - 69.8)	
	Range	18.1 - 93.1	18.0 - 96.1	35.0 - 97.0	-
<b>Gender (%)</b>	Male	310 (43.1)	2006 (37.8)	1272 (37.0)	0.009
	Female	410 (56.9)	3304 (62.2)	2167 (63.0)	

Figure 2-4 Histogram of preoperative OKS in the three centres for patients undergoing a UKR

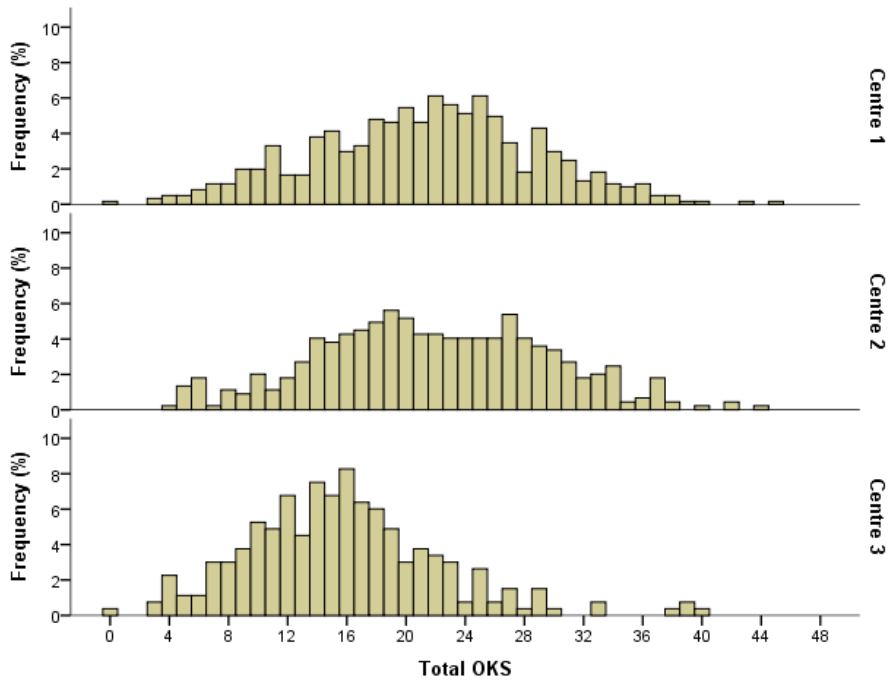
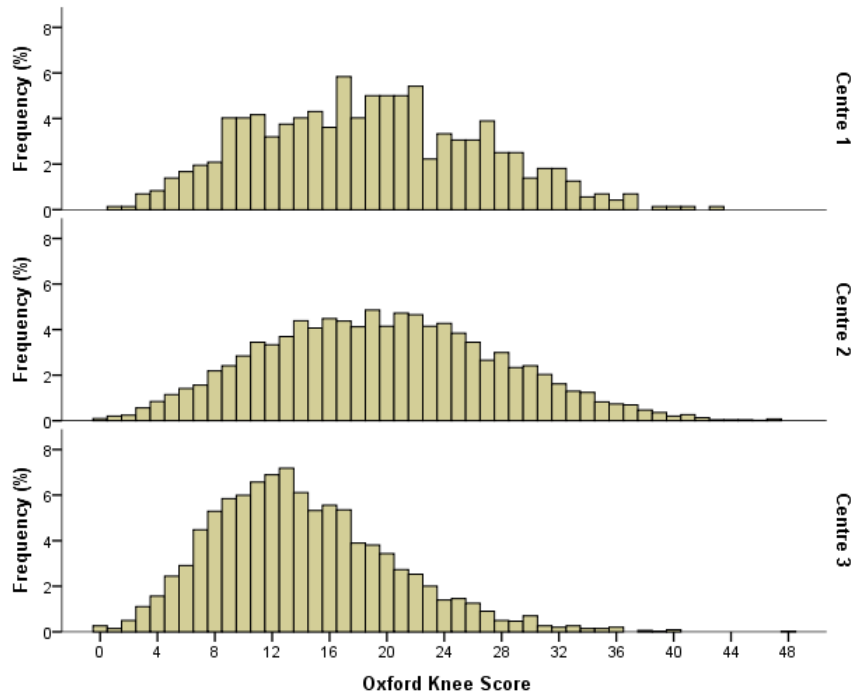


Figure 2-5 Histogram of preoperative OKS in the three centres for patients undergoing a TKR

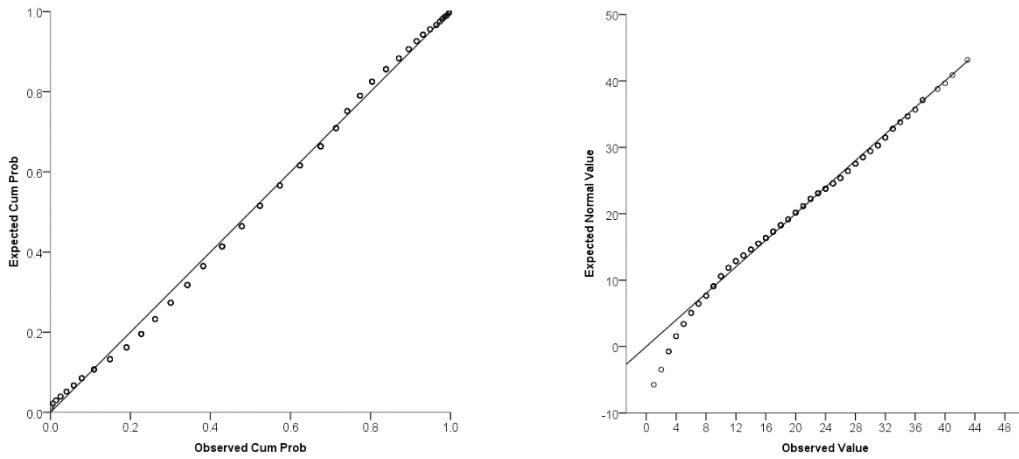


**Table 2-2 Summary descriptive statistics for the preoperative OKS in each of the three centres, according to procedure performed.**

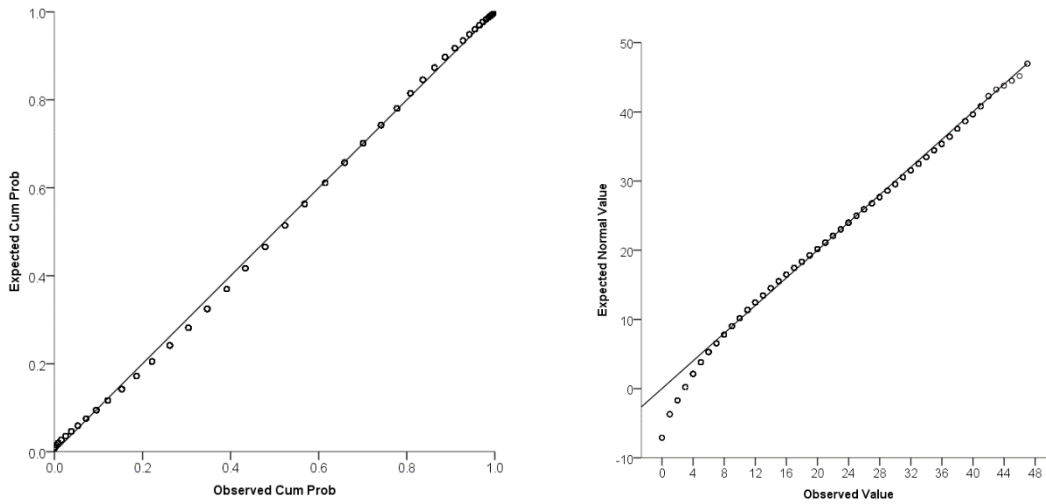
	Centre 1	Centre 2	Centre 3	p-value
<b>Unicompartmental replacement</b>				
Mean OKS (95% CI)	21.3 (20.7-21.9)	21.8 (21.1-22.5)	15.7 (14.9-16.5)	<0.001
Median	22	22	15	
Mode	22	19	16	
SD	7.48	7.70	6.58	
Min	0	4	0	
Max	45	44	40	
Range	45	40	40	
Interquartile range	16,26	16,27	11,19	
Skewness	-0.04	0.04	0.75	
Kurtosis	-0.21	-0.36	1.42	
<b>Total Knee Replacement</b>				
Mean OKS (95% CI)	18.7 (18.1-19.3)	19.7 (19.5-19.9)	14.1 (13.9-14.4)	<0.001
Median	19	19	13	
Mode	17	19	13	
SD	7.81	8.13	6.33	
Min	1	0	0	
Max	43	47	48	
Range	42	47	48	
Interquartile Range	13,24	14,25	10,18	
Skewness	0.22	0.20	0.69	
Kurtosis	-0.48	-0.35	0.72	

**Figure 2-6** Probability plots (P-P and Q-Q plots) examining normality of the preoperative OKS distribution for TKR in: a) Centre 1; b) Centre 2; and c) Centre 3.

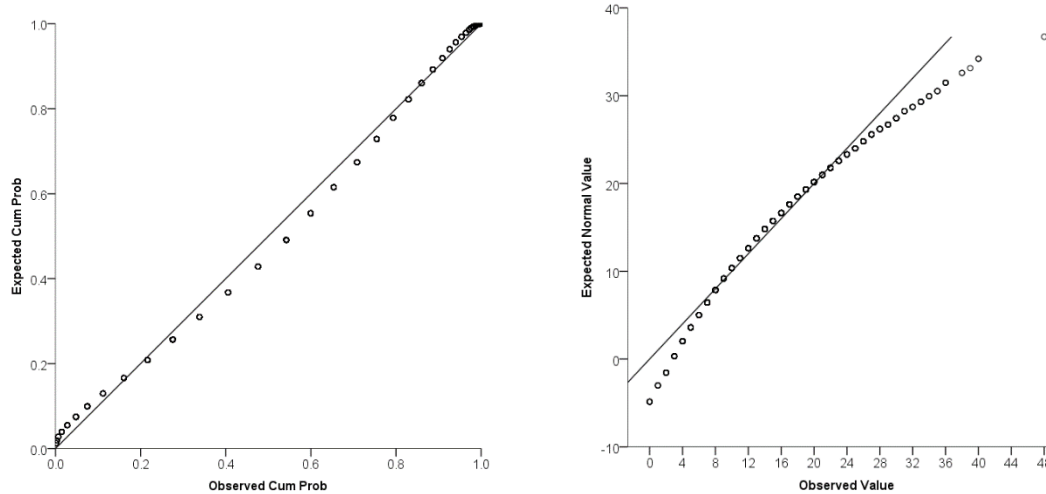
a)



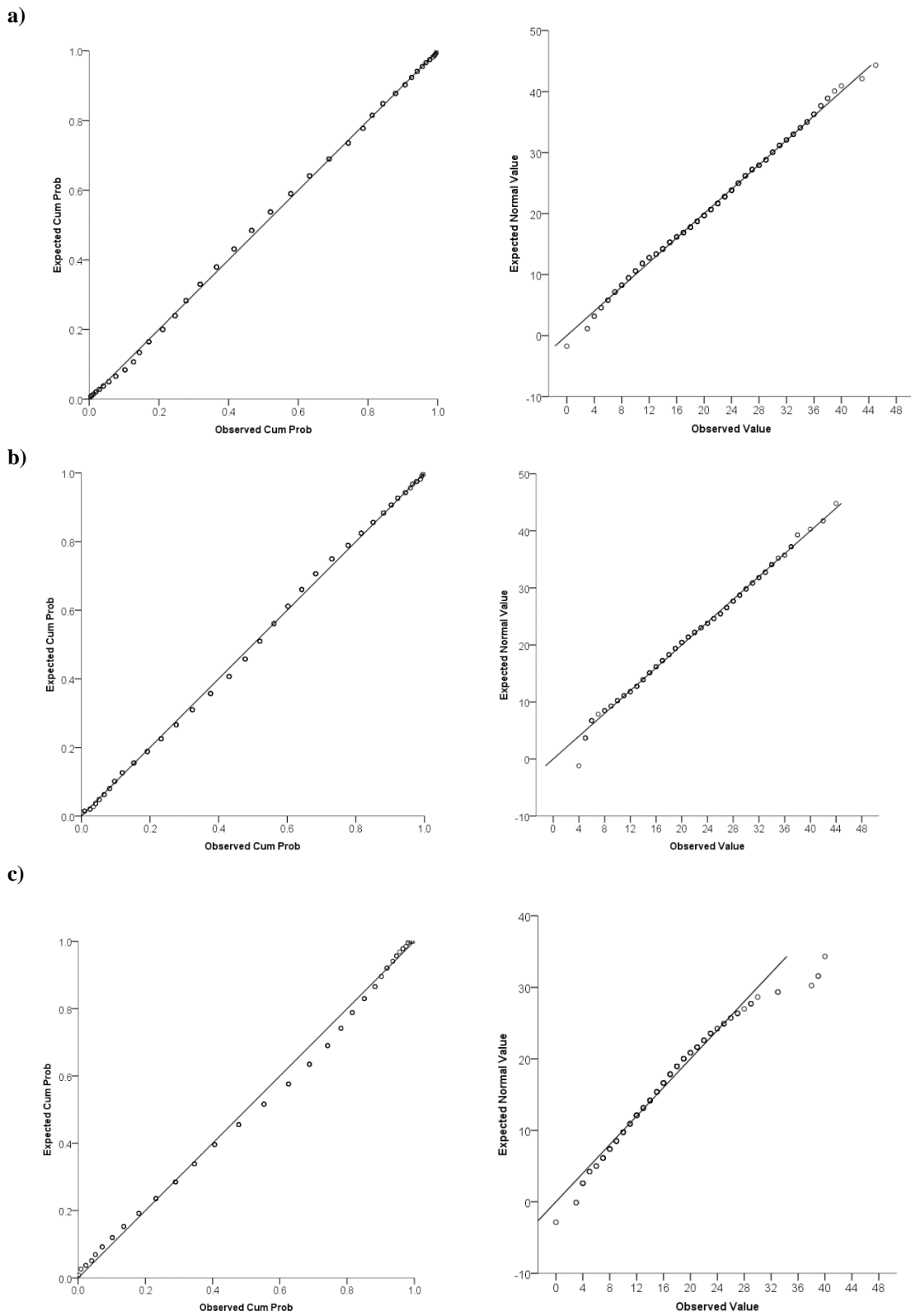
b)



c)



**Figure 2-7** Probability plots (P-P and Q-Q plots) examining normality of the preoperative OKS distribution for UKR in: a) Centre 1; b) Centre 2; and c) Centre 3.

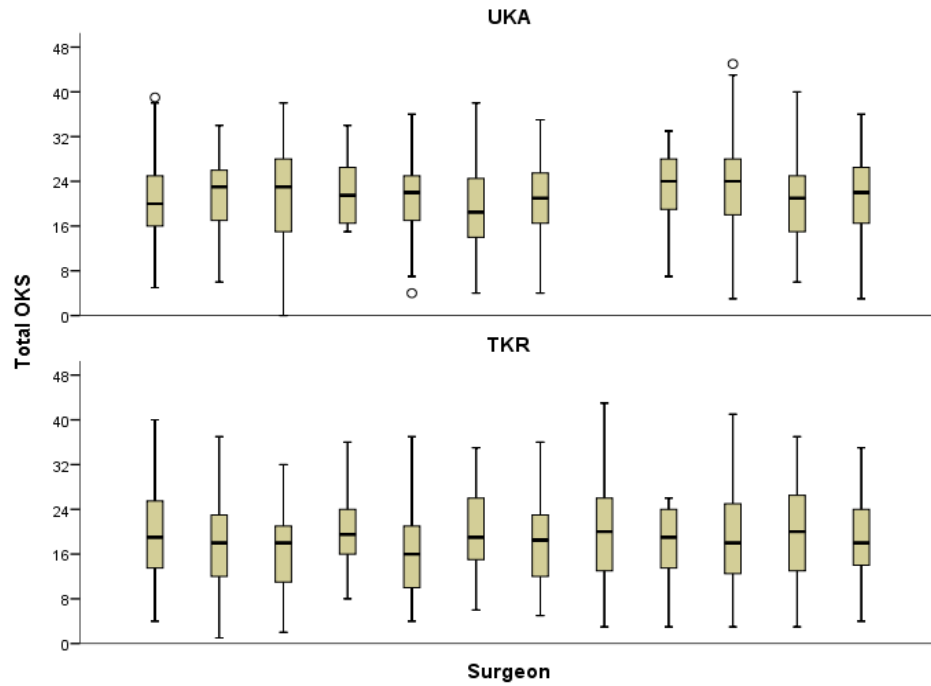


### **2.3.2 Preoperative OKS profile by Surgeon**

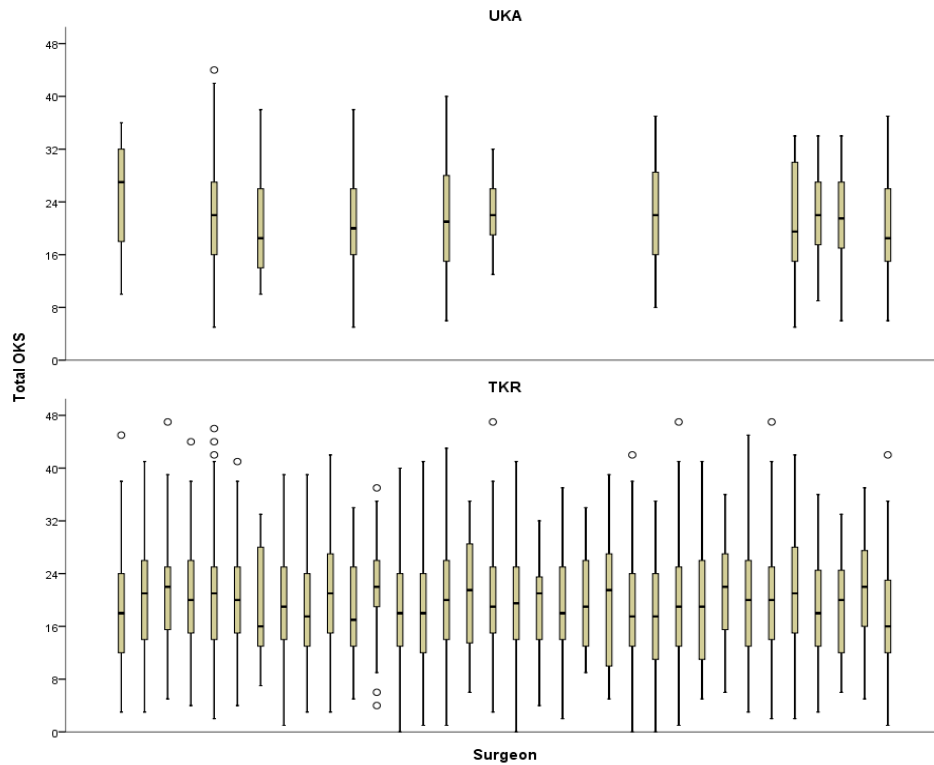
A total of 61 surgeons were included in the analyses (Centre 1 = 12, Centre 2 = 33, and Centre 3 = 16). All surgeons (61) performed TKR procedures, whilst UKR was only performed by 29 (47.5%) surgeons. Comparison of preoperative OKS amongst surgeons within the same centre demonstrated mixed results. In the case of UKR, no significant differences were evident in any of the three centres (Centre 1,  $p=0.442$ ; Centre 2,  $p=0.728$ ; Centre 3,  $p=0.518$ ). Similarly, in Centre 1 no difference in preoperative OKS was evident between surgeons with regards to TKR ( $p=0.821$ ). However, both Centre 2 and 3 demonstrated significant differences in preoperative OKS between surgeons in relation to TKR cases ( $p<0.001$ ). This difference persisted in both centres after adjusting for age and gender in a multivariate model ( $p<0.001$ ).

Wide variability in the preoperative OKS values was observed amongst individual surgeons, with the coefficient of variation ranging from 31-61%. Probability plots demonstrated that the distribution of preoperative OKS closely approximated a normal distribution for surgeons across all three centres. The distribution of OKS values for individual surgeons is illustrated in Figure 2-8 to Figure 2-10.

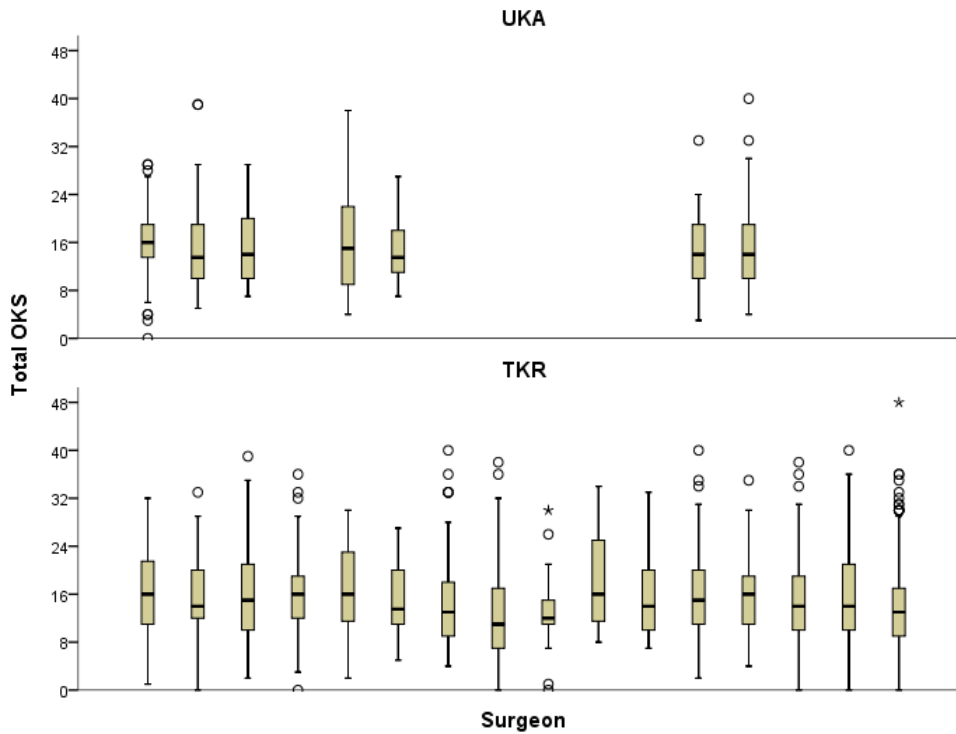
**Figure 2-8** Boxplot of preoperative OKS for UKR and TKR according to individual surgeons in Centre 1. Outlying data points reflected by  $>1.5$  IQR.



**Figure 2-9** Boxplot of preoperative OKS for UKR and TKR according to individual surgeons in Centre 2. Outlying data points reflected by  $>1.5$  IQR ).



**Figure 2-10** Boxplot of preoperative OKS for UKR and TKR procedures, according to individual surgeons in Centre 3. Outlying data points reflected by  $\circ >1.5$  IQR, \*  $>3$  IQR.



## 2.4 Discussion

This initial study represents the largest reported series to date, exploring the detailed profile of the preoperative OKS amongst 10,785 patients proceeding to knee replacement. It provides a valuable insight into the preoperative symptom burden within this specific population. More importantly it also provides an indication of what constitutes current arthroplasty practice in the UK. The main observation of this study was the normal distribution profile to the preoperative OKS evident across all three centres. In the postoperative setting, a large negative skew to the OKS data is generally observed, reflecting the fact that most patients achieve significant pain relief, and restoration of function following knee replacement [106]. Conversely we might expect

most patients awaiting knee replacement to have considerable pain and loss of function [23]. However, a corresponding positive skew in the preoperative OKS was not seen in this study, echoing limited data from previous studies [106-108]. Interestingly, the range of preoperative OKS values encompassed the whole possible spectrum (0 to 48) in all three centres, implying that some patients proceeded to surgery with minimal pain and/or disability. Even amongst those patients in centre 3, where an overall shift in the distribution towards lower preoperative OKS values were observed, less than 25% of patients underwent knee surgery with a preoperative OKS<10. Despite the perceived importance of severe pain and disability as criteria for knee replacement eligibility these observations suggest that the level of pain and disability tolerated by patients before proceeding to a knee replacement may be lower than we might assume. The wide range of preoperative OKS values may also reflect a broader range of clinical indications where pain and function are not the primary features. These might include surgery performed for indications such as severe deformity. As a primary measure of pain and function, we should consider whether a single OKS value accurately reflects the full complement of clinical indications leading to knee replacement.

The OKS is established and widely reported in the literature in the context of TKR outcome studies. Basic data pertaining to the preoperative OKS is often included in such studies, primarily as a baseline comparison. In most reports however this data usually consists of mean OKS values, together with reporting of standard deviations or confidence intervals. Although these values may provide a useful summary and comparative baseline marker for the purpose of outcome studies, they provide very limited use for application in routine preoperative clinical practice, and may not be easily interpretable to most individuals examining the data. At present there are very few

published studies specifically exploring the detailed profile of the preoperative OKS in patients proceeding to knee replacement [106-108].

The initial national PROMs report highlighted the OKS distribution across the whole UK cohort but no further in-depth exploration was undertaken in this report to explore the profile within individual centres [106]. Pynsent et al reported on the distribution of preoperative OKS in a cohort of 1739 TKR patients within a single centre [107]. Significant variability in preoperative OKS was also seen in their study, encompassing the entire range of possible values, but with an overall trend towards lower scores (Median 15, IQR 10-21). This more closely resembled the profile observed in Centre 3 (Median 13, IQR 10, 18) in the current study. Jenny et al also described a general trend towards worse scores, in their OKS validation study of 200 patients (Mean 16.3, Range 4-39) [108]. There have also been some explorative studies of preoperative status based on alternative outcome measures. A Canadian study examining waiting times for hip and knee arthroplasty found that the majority of patients on an arthroplasty waiting list had only mild or moderate pain, and functional limitations, as defined by the Western Ontario and McMaster Osteoarthritis Index (WOMAC) [111]. Approximately one-third reported severe functional restriction, and only a quarter reported severe pain. Similarly for hip arthroplasty, Dieppe et al found wide variability in preoperative symptom burden, and concluded that simple scores of pain and function did not reflect the complexities of the decision process leading to THR [64]. Although there is only limited data published in relation to the preoperative symptom burden of knee replacement patients, similar findings with other PROM tools, and for THR suggest that variability in preoperative symptom burden seen in this study is a real finding rather than simply a feature of the OKS.

In addition to the general variability in preoperative OKS observed, there was also evidence to suggest the presence of both inter-hospital and inter-surgeon variation in symptom burden. This represented a clinically important difference, with up to 6.1-point difference seen between the three centres' mean preoperative OKS, and up to 9-point difference between surgeons [112]. This would seem to support some of the concerns raised in earlier studies regarding variation in access and provision of arthroplasty services across different populations. Variation between surgeons, particularly where patients are undergoing knee replacement with minimal pain and disability (high OKS) may reflect application of inappropriate selection criteria. This would support the arguments of threshold criteria proponents, whereby OKS based thresholds standardise the selection process, and limit inappropriate access to surgery. A broad range of preoperative OKS values was however evident for most surgeons' practice, demonstrating that surgery in the presence of a high preoperative OKS appears to be common practice, and not isolated to a few individual surgeons. The limited data available in the present study prevented formal detailed analyses on inter-surgeon variation, but there was also some overall homogeneity evident for surgeons within the same centres. Although we might expect some variation in subjective and objective selection criteria amongst individual surgeons, this would suggest that regional factors, such as sociodemographic variables, and access to services may be more important in explaining these differences.

Although the primary focus of this study was to examine total knee replacement, the mixed arthroplasty workload within the centres provided additional information relating to both UKR and TKR. Previous outcome studies have identified that the mean preoperative OKS is generally slightly higher for UKR compared to TKR [113]. This is

generally assumed to reflect slightly less severe disease state due to the involvement of a single compartment. However, similar to the TKR population, there is no detailed profiling of the preoperative status of UKR patients published in the literature. As has been noted in previous studies, the mean preoperative OKS was up to 2.6 point higher for UKR patients compared to TKR [113]. Although this reached statistical significance in all three centres, this does not necessarily represent a clinically important difference [112]. The overall profile of the UKR cohort was also very similar to the TKR group, with wide variability in the preoperative OKS evident, and a normal distribution profile in each centre. The slight rightward shift in the OKS data for UKR may be an important consideration in the context of the UK threshold proposals, as a single threshold value for all arthroplasty procedures would risk excluding a greater proportion of patients deemed suitable for UKR. Nonetheless the overall preoperative symptom burden of arthroplasty patients seems similar regardless of whether they have unicompartmental or tricompartmental disease. Application of OKS based thresholds as eligibility criteria may therefore be inappropriate for some patients, and in the context of very low threshold levels, could potentially submit patients to additional unnecessary and inappropriate suffering.

The strengths of this study include the large study population, assessed utilising a prospectively collected, and validated measure of pain and function. With the lack of any other detailed analysis of the preoperative symptom burden, this therefore represents an important and clinically relevant exploratory analysis of this subject. There are however limitations which should be considered. Despite the large sample size, only three individual centres were available for analysis, and the limitations of individual arthroplasty datasets prevented full validation. It is therefore not possible to accurately

determine how representative the study population may be, and the results observed may not be a true reflection of the general population. Similarly, all three centres represented dedicated elective orthopaedic centres, participating in high-volume primary and revision arthroplasty practice, and the findings may not be directly applicable to smaller orthopaedic units within district general hospitals. OKS data was not available for the complete arthroplasty population, and the potential impact of the missing data on the overall distribution of the OKS should be considered. The current study was also limited by the lack of demographic data readily available across all three databases. Such factors have been identified as important contributors to the variation in access to arthroplasty [57,114], and failure to fully account for these variables will potentially bias the results.

In summary, this preliminary study provides a detailed description of the preoperative OKS profile in the knee replacement population. It suggests substantial variation in the preoperative symptom burden of patients proceeding to knee replacement, with some patients proceeding to surgery with seemingly minimal symptom burden. Furthermore, there appears to be possible variation in the preoperative status of patients across different centres and surgeons which warrant further exploration.

## **Chapter Three**

### **3 Variation in the Preoperative Oxford Knee Scores – Analysis of the National Arthroplasty Database**

#### **3.1 Introduction**

Variation in healthcare represents a global concern that has been regularly reported in the literature over the last few decades [26,58,64,115]. Studies have highlighted evidence of variation in health care performance, access and utilisation of services, the influence of demand and supply, and professional decision making [115]. In the current climate of limited healthcare resources, together with a drive to improve healthcare provision and outcomes, there has been renewed interest in identifying and abolishing so called ‘unwarranted variations’ in healthcare within the UK [115,116]. Variation in relation to joint arthroplasty procedures has been specifically targeted in recent reports [58,116]. The Atlas of Variation report suggested a twofold variation in severity of knee symptoms and disability in patients undergoing knee replacement in the UK [58]. Although several factors have been highlighted as having an important impact on healthcare variation, the subjective assessments and decision-making role of individual clinicians is considered a major component of unwarranted variation [116].

The previous chapter (Chapter 2) examined the preoperative symptom burden prior to knee replacement as measured by the preoperative OKS. The work highlighted that in addition to the overall variability in preoperative OKS amongst these patients there may

also be variation between different centres and surgeons. Detailed analysis could not be undertaken using this data due to the limited number of centres examined and an inability to adjust for sociodemographic differences. Having already identified potential variation between individual surgeons and different centres (in the previous chapter), further exploration and analysis of the preoperative OKS from a larger and more representative sample is essential. Following on from the development of the initial study in Chapter 2, data from the UK national arthroplasty database became freely available, providing an opportunity to further expand and explore the profile of the preoperative OKS using much a larger and representative dataset of current national practice.

If true variation is indeed evident, standardisation of knee replacement referral processes and application of threshold criteria may be appropriate. However, there is currently no published report exploring the potential variation in preoperative OKS. Without fully understanding the potential variations in preoperative OKS, we cannot fully comprehend how OKS thresholds will impact on different populations.

### **3.1.1 Study Aims**

The aim of this study was to examine in greater depth, the apparent variation (as identified in Chapter 2), in the preoperative OKS prior to knee replacement across different population groups. In addition, further detailed aims include determination of;

1. The amount of variation in preoperative OKS attributable to differences between individual centres and surgeons;
2. The impact of sociodemographic characteristics on preoperative OKS variation; and

3. The impact of proposed OKS threshold limits on different knee replacement populations

## **3.2 Methods**

### **3.2.1 Study Population**

Basic demographic data pertaining to all admissions, outpatient appointments, and A&E attendances at NHS hospitals in England have been routinely collected as part of the national Hospital Episode Statistics (HES) database since 1987. In addition, a separate national arthroplasty PROMs database was established in 2009, to collect, collate and report outcomes following joint replacement [109]. Since its introduction, collection and submission of preoperative and postoperative PROMs data has been mandatory for all hip and knee replacement procedures, which in turn has led to a large established dataset of outcome data relating to hip and knee replacements. In combination with the HES database, this provides an opportunity for detailed analyses of the PROMs data whilst incorporating the potential influence of sociodemographic factors. Approval was therefore requested to access the PROMS and corresponding HES data collected between April 2009 to August 2011 from the relevant authorities.

All patients that had undergone a primary TKR procedure under the care of a NHS provider were included in the study. In order to ensure that the preoperative OKS was representative of the immediate preoperative status of patients undergoing knee replacement, only cases where the preoperative questionnaires were completed 90 days or less, prior to surgery were included. Unicompartamental knee replacement procedures

were excluded from analysis in this study in order to focus primarily on total knee replacement procedures. Furthermore, the accuracy of data coding for unicompartmental procedures within the HES database was not considered robust enough for further analysis, with the number of UKR procedures not correlating with expected values based on data available from the national joint registry. This raised concerns about the potential data quality for unicompartmental procedures, further supporting the decision to exclude from further analyses. Revision procedures were also excluded from the analyses. Centres and surgeons with less than 10 cases recorded in the database were excluded in order to minimise the potential impact of spurious results from non-representative samples.

Data from each database was explored and subjected to initial data cleaning to prepare the data for this study. The respective datasets were subsequently merged to combine the PROMs data with its respective HES counterpart. Data was matched utilising two unique patient identifiers (EPIKEY and HES Identifier No.), present in both databases. A further cleaning process was performed post-merger to deal with unmatched cases, and to apply the full inclusion and exclusion criteria. The final study cohort included 70,016 patients for further analysis (Figure 3-1).

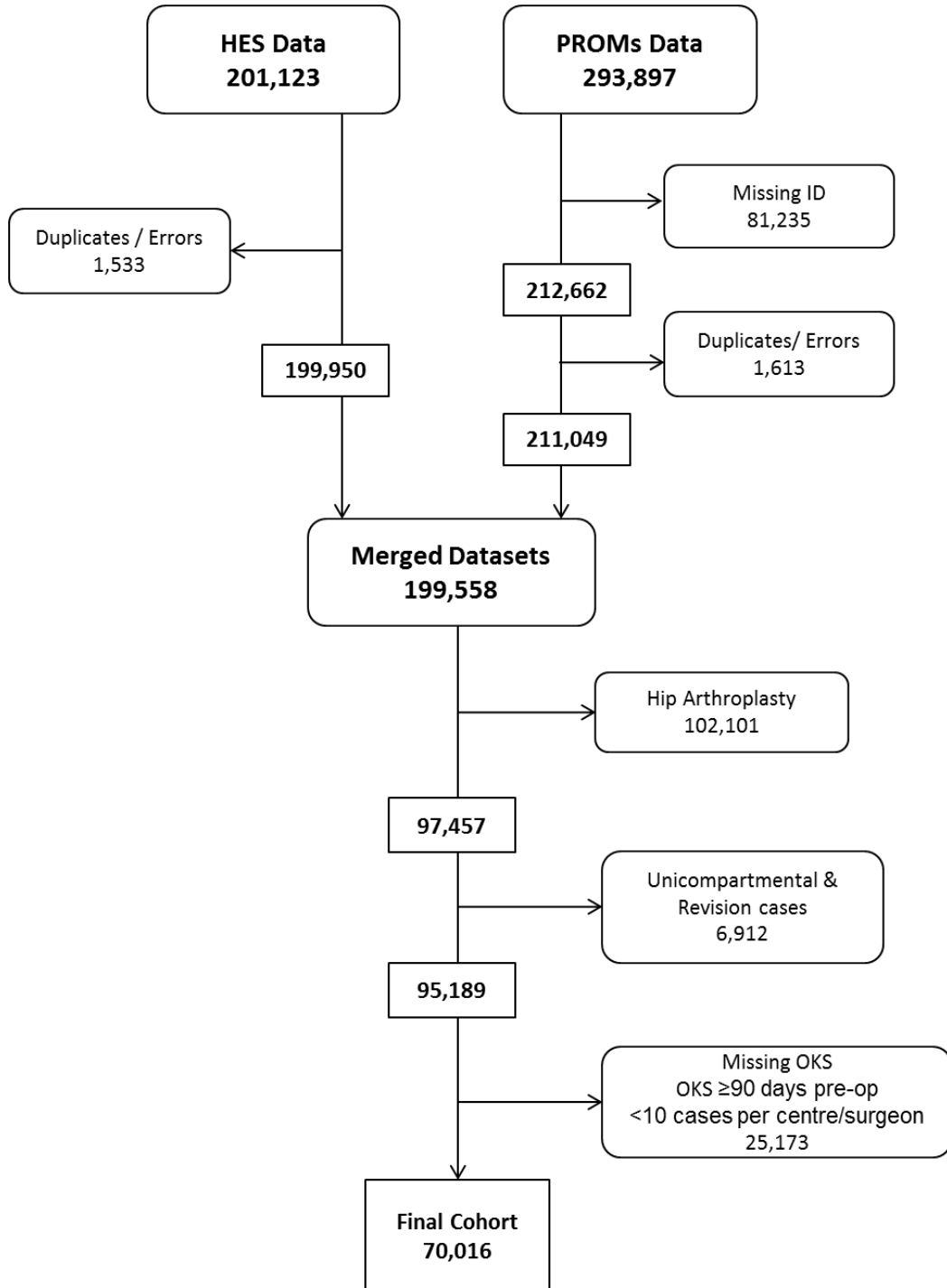
### **3.2.2 Data Analyses**

Basic data exploration was repeated for this larger dataset, utilising a combination of summary descriptive statistics, and graphical plots as previously described in Chapter 2. Due to the large number of individual centres and surgeons included in this larger dataset, formal detailed analysis of data distribution and assessment of normality could

not be undertaken for each individual centre and surgeon. The OKS distribution was therefore assessed and presented using a combination of boxplots (Centre) and scatterplots (Surgeon) incorporating median, IQR, minimum, and maximum values.

Patient characteristics available in the dataset for incorporation into the analyses included age, gender, diagnosis, Index of Multiple Deprivation (IMD), ethnicity, self-reported comorbidities, and symptom chronicity. Within the original databases this data was sub categorised to several levels. For ease of interpretation and analyses the categorisation was simplified accordingly; IMD (quintiles, most deprived (1) through to least deprived (5)); ethnicity (White, Black, Asian, Mixed, other); self-reported co-morbidities (none, one, two, three, four or more); symptom chronicity (<1yr, 1-5yrs, 6-10yrs, >10yrs). Initial univariate regression analyses were undertaken to examine the relative impact of the sociodemographic factors on the preoperative OKS.

**Figure 3-1** Flowchart summarising the data extraction and selection process for the study cohort



### 3.2.3 Multilevel Modelling

When comparing data between groups using simple regression modelling there is usually an assumption that all data is independent. In examining variation between different centres and different surgeons, this assumption of independence cannot be met, as patients associated with the same centre or same surgeon are more likely to be similar to each other, and a hierarchical structure is therefore considered to exist in the data. Failure to fully account for these potential biases risks overestimating the true differences between the groups [117].

In order to account for the hierarchical nature of the data, where individual patients are clustered by surgeon and their respective centre, whilst also accounting for other explanatory variables, a multilevel model was utilised. The hierarchy of the cohorts analysed in this study represented a 3-level model. Individual centres and surgeons were incorporated into the model as random factors. A small number of surgeons performed procedures across more than one centre, and therefore a cross-classified model approach was used. The additional sociodemographic factors described previously were incorporated into the model as fixed factors. A stepwise approach was utilised to develop the final model, and best-model estimation undertaken using maximum likelihood estimation, to allow comparison of models in response to both fixed and random factors. Successive models were compared using the Log 2 likelihood value, and likelihood ratio test based on the  $\chi^2$  distribution statistic.

The primary aim of this study was to determine the impact of individual centre and surgeons on variation in the preoperative OKS. The relative contribution of individual

centres and surgeons to the overall variability in preoperative OKS was therefore established by calculating the Variance Partition coefficient (VPC), and intraclass correlation coefficients (ICC) [118] . For simple 2-level models, the ICC and VPC are represented by the same calculation, and can be defined as:

$$ICC / VPC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$$

Where  $\sigma_u^2$  represents the level-2 variance and  $\sigma_e^2$  represents the level-1 variance. This study however represents a more complex model where ‘centre’ represents a level-3 variable, ‘surgeon’ a level-2 variable and the individual patient as the level-1 variable. For these more complex 3-level models, individual calculations of ICC and VPC values permit additional analyses to further determine the source of variability in the data being examined; where  $\sigma_v^2$  represent level-3 variance,  $\sigma_u^2$  represents level-2 variance, and  $\sigma_e^2$  represents the level-1 variance. VPC calculation allows determination of the amount of variation attributable to each respective level.

Centre level (Level-3) VPC is calculated as the centre variance in relation to the total variance:

$$VPC = \frac{\sigma_v^2}{\sigma_v^2 + \sigma_u^2 + \sigma_e^2}$$

Surgeon level (Level-2) VPC is calculated as the ratio of surgeon variance to the total variance:

$$VPC = \frac{\sigma_u^2}{\sigma_v^2 + \sigma_u^2 + \sigma_e^2}$$

Patient level (Level-1) VPC is calculated as the ratio of the patient variance to the total variance:

$$VPC = \frac{\sigma_e^2}{\sigma_v^2 + \sigma_u^2 + \sigma_e^2}$$

Calculation of the ICC provides a measurement of the degree of similarity within a given cluster level. Correlation between patients treated by the same surgeon and by the same centre is defined by:

$$ICC = \frac{\sigma_v^2 + \sigma_u^2}{\sigma_v^2 + \sigma_u^2 + \sigma_e^2}$$

Finally, correlation between different surgeons within the same centre can be estimated by:

$$ICC = \frac{\sigma_v^2}{\sigma_v^2 + \sigma_u^2}$$

All statistical analyses were performed using SPSSv20 (IBM, Armonk, NY), and a p-value <0.05 was considered significant throughout.

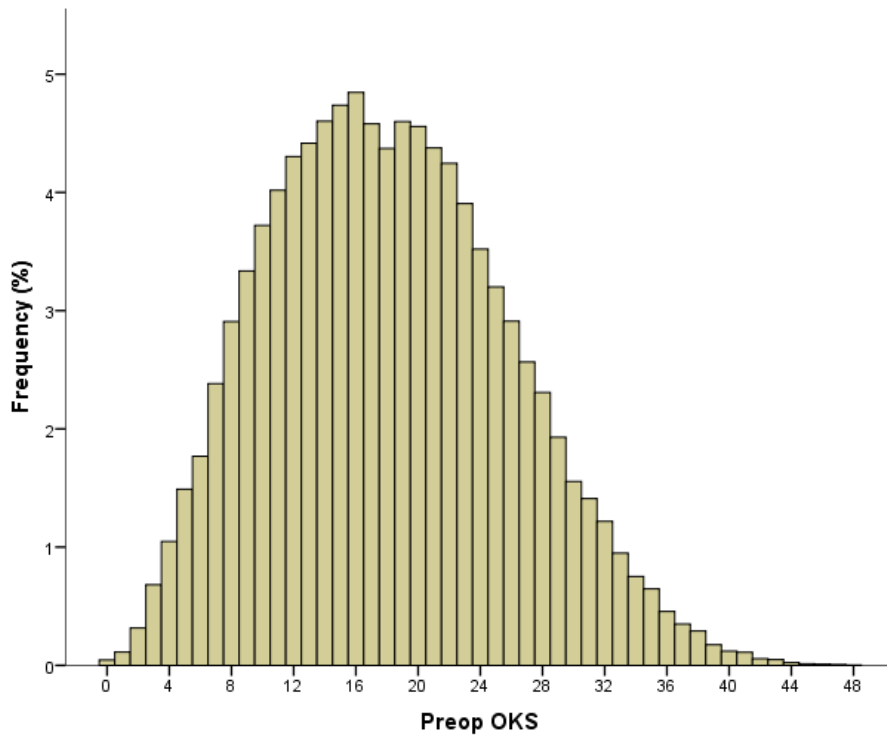
### 3.3 Results

Summary statistics and graphical representation of the distribution of the preoperative OKS within the complete study cohort are presented in Table 3-1 and Figure 3-2. These demonstrate that the overall profile approximates a normal distribution profile as observed in Chapter 2, which is further confirmed by examination of the probability plots relating to this data (Figure 3-3). A similar ‘tail’, reflecting patients presenting with high preoperative OKS values is also evident.

**Table 3-1 Summary statistics for the preoperative OKS data from the national dataset**

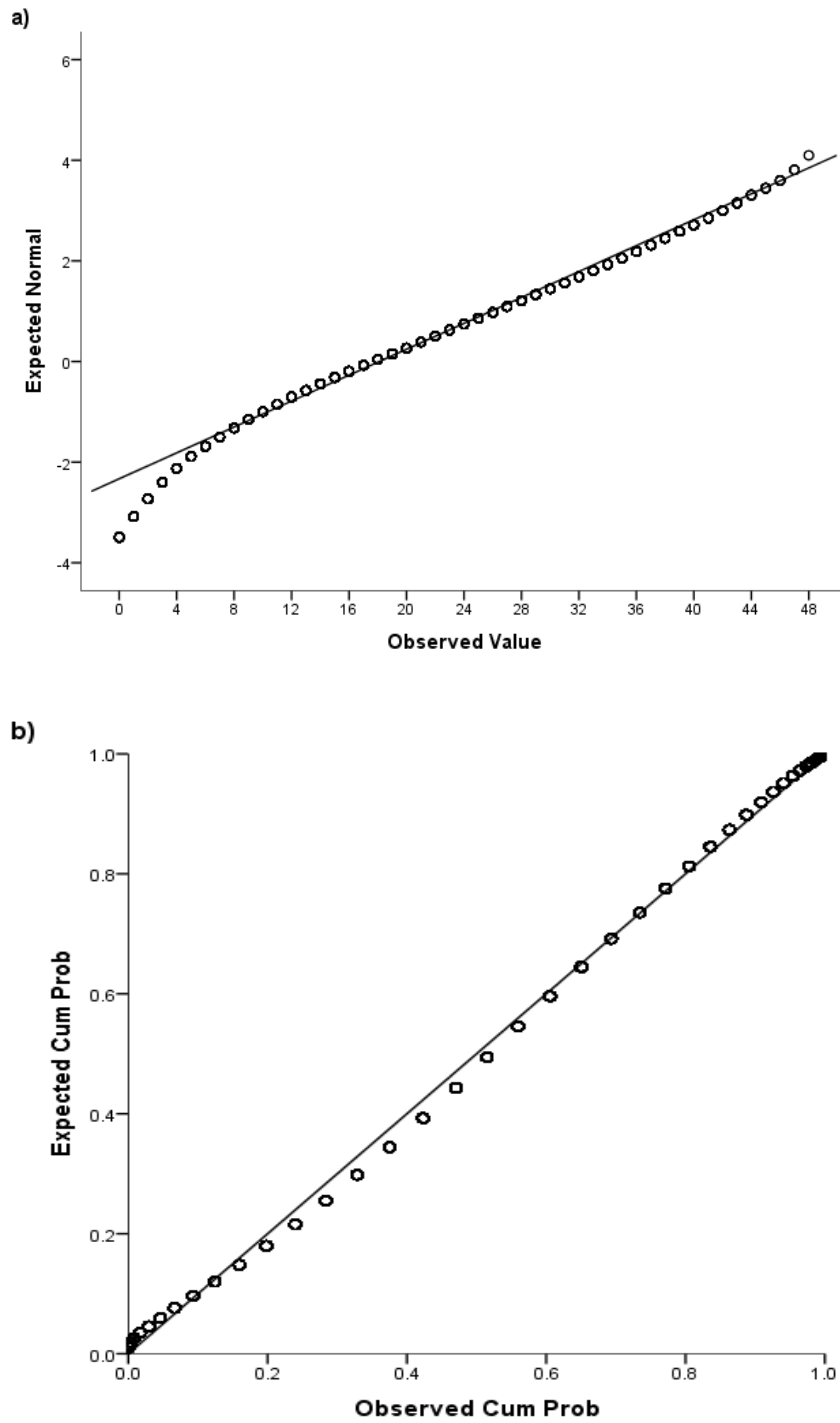
	Value
Mean	18.1
<i>95% Confidence Interval for Mean</i>	18.1-18.2
5% Trimmed Mean	17.9
Median	18.0
Variance	60.3
Std. Deviation	7.8
Minimum	0
Maximum	48
Range	48
Interquartile Range	11
Skewness	0.31
Kurtosis	-0.33

**Figure 3-2** Histogram of the preoperative OKS distribution in the study cohort suggesting an overall normal distribution profile



The sociodemographic characteristics of the study cohort are summarised in Table 3-2, and demonstrate the presence of a broad population mix to the study cohort. Univariate analyses examining the individual influence of these sociodemographic factors on preoperative OKS demonstrated all variables to have a significant effect. All variables were therefore included in the subsequent statistical models.

**Figure 3-3** Probability plots demonstrating normal distribution to the preoperative OKS.  
a) Q-Q plot, and b) P-P plots



**Table 3-2 Summary of the sociodemographic characteristics of the study cohort, and unadjusted regression coefficients from univariate analyses in relation to the preoperative OKS. Unadjusted estimates reflect the difference in preoperative OKS observed in relation to the reference group**

	Number of patients (%)	Unadjusted estimates (95% CI)	p-value
<b>Age</b>	70,016 (100)	0.07 (0.06, 0.07)	<0.001
<b>Gender</b>			
Male	29,736 (42.5)	3.48 (3.36, 3.59)	<0.001
Female	40,280 (57.5)	Reference	-
<b>IMD quintile (low to high)</b>			
1	11,340 (16.2)	-4.32 (-4.51, -4.14)	<0.001
2	13,179 (18.8)	-2.77 (-2.95, -2.59)	<0.001
3	15,155 (21.6)	-1.83 (-2.00, -1.65)	<0.001
4	15,369 (22.0)	-1.12 (-1.30, -0.95)	<0.001
5	14,120 (20.2)	Reference	-
Missing data	853 (1.2)	-	-
<b>Self-reported comorbidities:</b>			
None	39,727 (56.7)	5.72 (5.07, 6.36)	<0.001
1	20,888 (29.8)	4.08 (3.44, 4.73)	<0.001
2	6,991 (10.0)	2.51 (1.85, 3.18)	<0.001
3	1,857 (2.7)	1.14 (0.41, 1.87)	0.002
4+	553 (0.8)	Reference	-
<b>Diagnosis:</b>			
Primary OA	66,461 (94.9)	4.03 (3.73, 4.34)	<0.001
Secondary OA	528 (0.8)	3.73 (3.33, 4.78)	<0.001
Inflammatory	2,509 (3.6)	Reference	-
Missing data	518 (0.7)	-	-
<b>Ethnicity</b>			
Mixed	158 (0.2)	-2.22 (-3.42, -1.02)	<0.001
Black	827 (1.2)	-4.38 (-4.71, -4.05)	<0.001
Asian	2,182 (3.1)	-3.66 (-4.19, -3.13)	<0.001
Chinese/other	364 (0.5)	-1.38 (-2.17, -0.58)	0.001
White	61,688 (88.1)	Reference	-
Missing data	4,797 (6.9)	-	-
<b>Duration of symptoms:</b>			
<1yr	3,229 (4.6)	0.92 (0.63, 1.22)	<0.001
1-5yrs	35,929 (51.3)	0.88 (0.73, 1.03)	<0.001
6-10yrs	15,461 (22.1)	0.14 (-0.35, 0.31)	0.118
>10yrs	15,009 (21.4)	Reference	-
Missing data	388 (0.6)	-	-

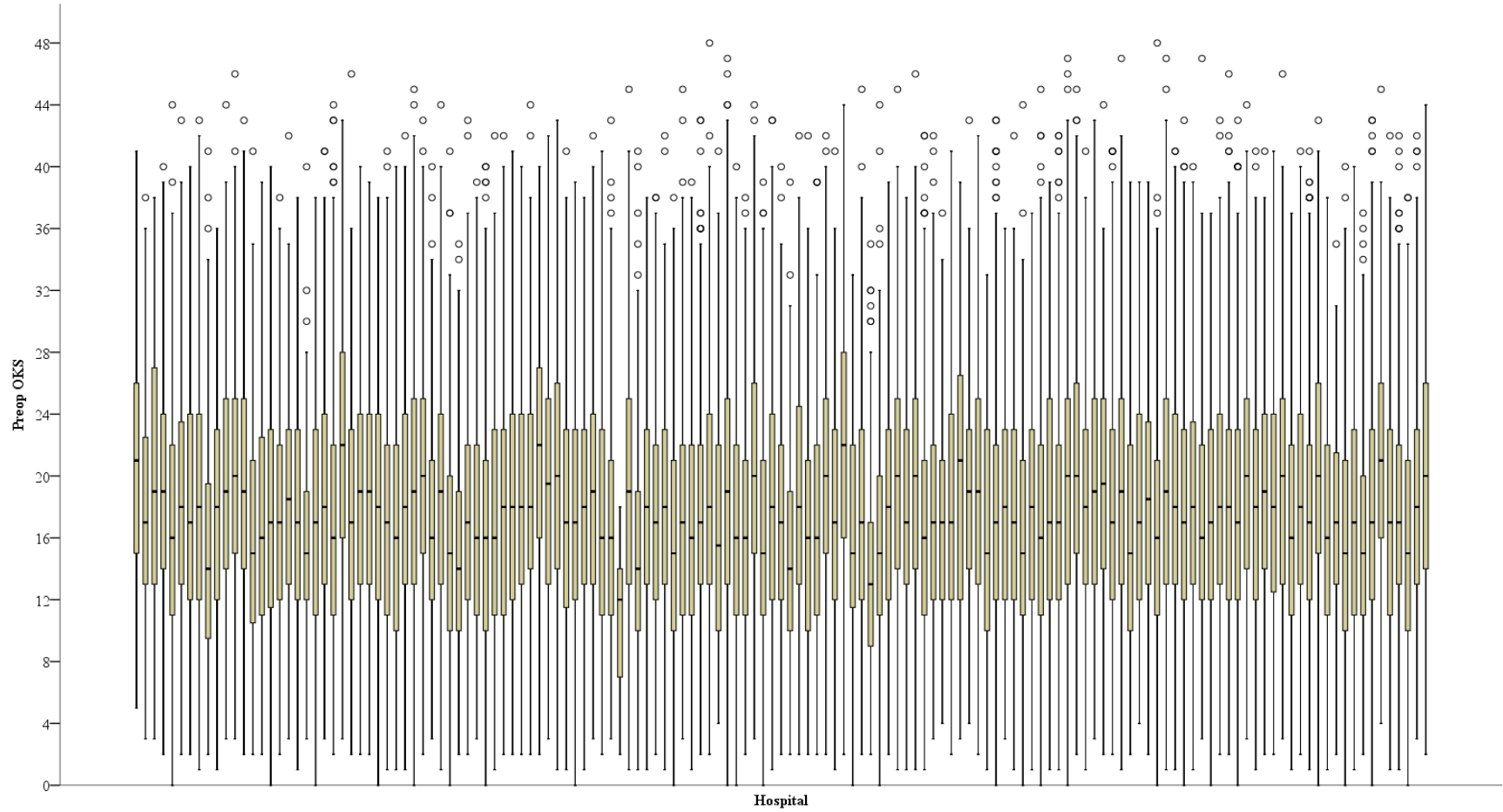
### **3.3.1 Preoperative OKS by Centre**

The study cohort was distributed across 145 individual NHS hospitals. The average number of cases performed in each centre during the study period was 717.2 (Range 35-2025). The distribution of the preoperative OKS within individual centres is shown in Figure 3-4. Wide variability in the preoperative OKS was observed across all centres, although the majority of patients fell within a similar narrower zone of OKS values as reflected by the IQR. The greatest variation is observed at the upper end of the OKS spectrum rather than the lower values. The minimum OKS values across centre ranged between 0-5 (5 points), compared to the range of maximum values from 33-48 (15 points). It can also be seen that all centres, except one, exhibit preoperative OKS values above 30. Mean OKS values in individual centres ranged from 14.0 to 22.1. Univariate analysis suggests that the overall difference in mean OKS between centres was statistically significant ( $p=0.033$ ).

Examination of the summary statistics and skew values from each centre showed that the trend was towards a slight positive skew, although only 23 (15.8%) cases had skew values  $>0.5$ , and all centre fell within skew values of -1.0 to 1.0, suggesting general conformity to a normal distribution.

Variation in the preoperative OKS

Figure 3-4 Boxplots demonstrating the preoperative OKS distribution within individual orthopaedic centres. Circles represent outlier cases



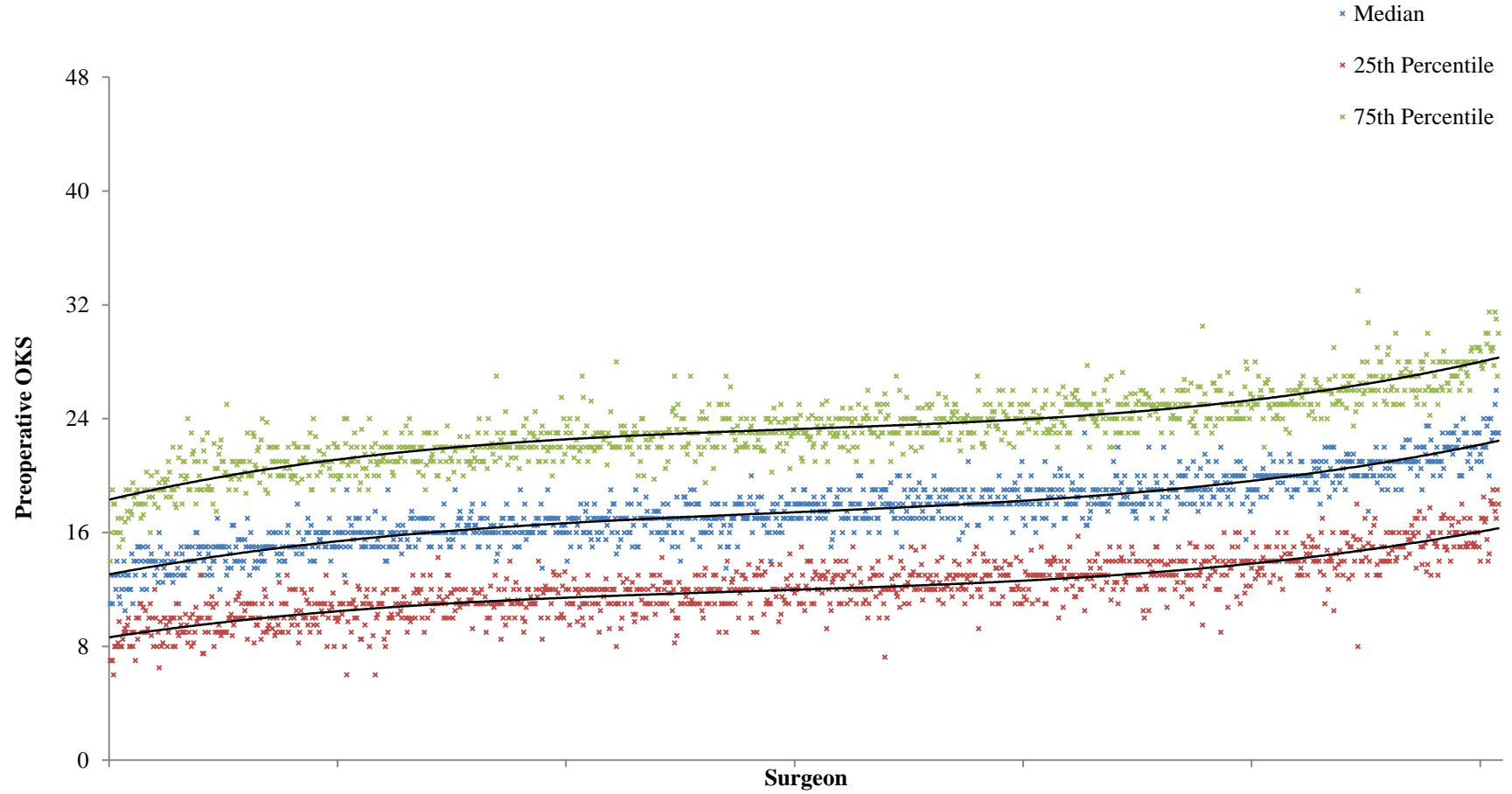
### 3.3.2 Preoperative OKS by Surgeon

The study cohort was distributed across a total of 1,216 surgeons. Due to the large number of surgeons, data distribution could not be easily summarised or evaluated using standard summary statistics, and data distribution is summarised in Figure 3-5 and Figure 3-6. Variability in the preoperative OKS distribution was apparent between surgeons, but it can be seen that all surgeons undertook TKR procedures across a broad spectrum of OKS values. Mean preoperative OKS values across surgeons ranged from 10.4 to 25.1, reflecting a large and clinically important difference in preoperative status as measured by the OKS. The IQR ranged from 5-25, although the overall trend for the IQR was similar across the cohort. As mean OKS increased the IQR demonstrated an overall upward shift as well, reflecting the overall normal distribution of data.

Greater variability was evident in the upper range of the preoperative OKS; maximum values ranged from 18-48 (30 points), whilst minimum OKS values ranged from only 0-16 (16 points). The lowest possible OKS value (0) was recorded for only 32 (2.6%) surgeons, whilst the maximum possible OKS (48) was seen for 2 (0.2%) surgeons. A total of 1186 (97.5%) of surgeons performed a TKR in the presence of an OKS  $\geq 26$ , whilst 963 (79.5%) performed TKR with an OKS  $\geq 32$ .

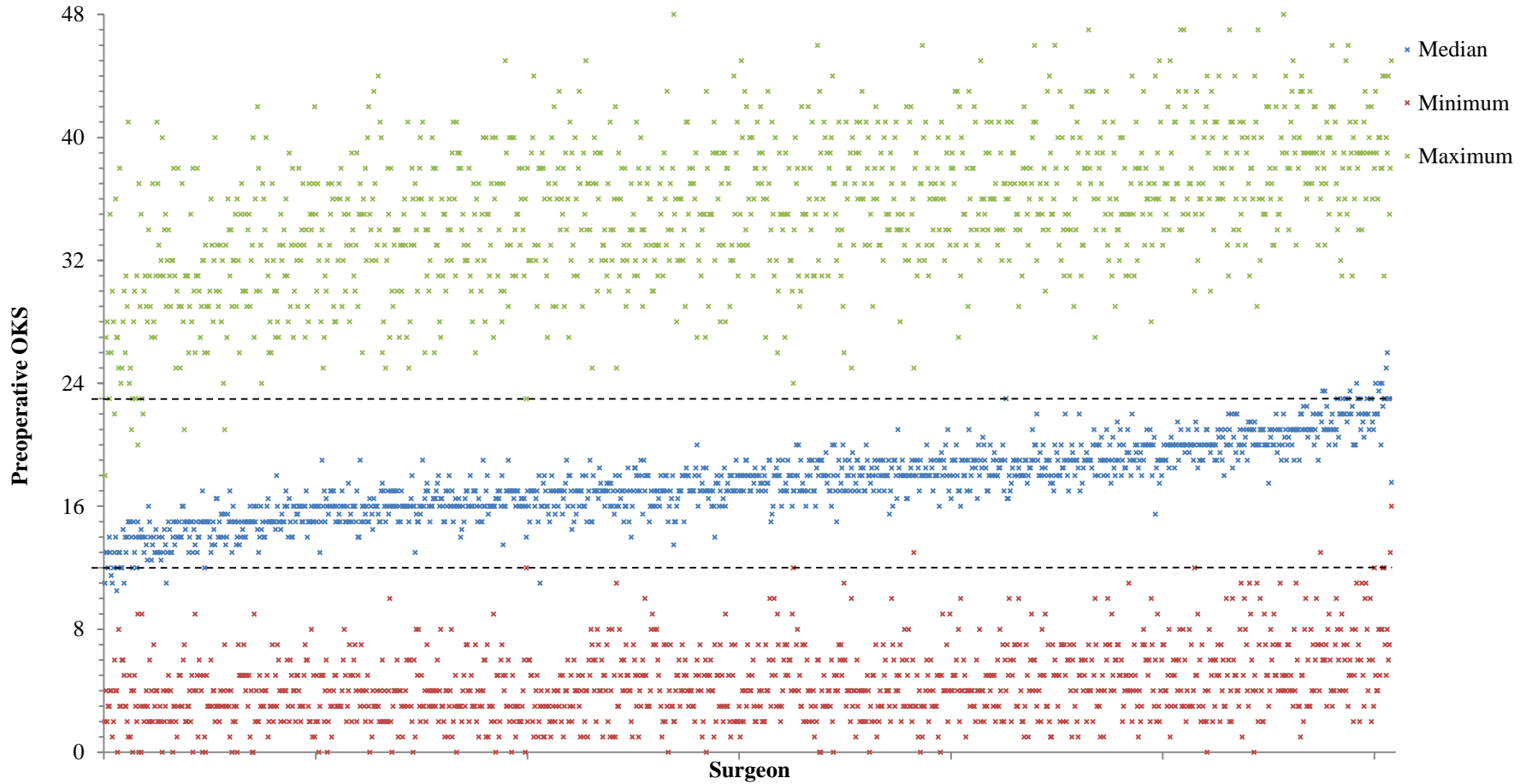
Variation in the preoperative OKS

Figure 3-5 Distribution of the preoperative OKS (interquartile range) according to individual surgeons. Solid lines represent the trend line for each respective value.



Variation in the preoperative OKS

Figure 3-6 Distribution of the preoperative OKS (minimum to maximum OKS values) according to individual surgeons. Dashed lines represent the overall IQR for the combined cohort.



### 3.3.3 Variation in the Oxford Knee Score

A summary of the process leading to selection of the final best-fit model, including incorporation of the additional sociodemographic characteristics, is detailed in Table 3-3. All potential explanatory variables entered into the model significantly improved the final model fit, and were therefore incorporated into the final model (Model 9). Comparison of a model with the random factors (surgeon, centre) excluded (model 10) with the full model (model 9), demonstrated that both centre and surgeon remained important factors even after adjusting for the other explanatory variables.

**Table 3-3 Summary of the multilevel model process leading to the final model.  $\chi^2$  Statistic relates to comparison of each successive model to the previous model.**

	Model variables	-2 RLL	Parameters	$\chi^2$ Statistic
Model 1	Random only Centre	483,875	3	-
Model 2	Random only Centre & surgeon	483,721	4	<0.001
Model 3	Random & fixed Model 2 & age	483,347	5	<0.001
Model 4	Random & fixed Model 3 & gender	479,693	6	<0.001
Model 5	Random & fixed Model 4 & comorbidities	477,467	10	<0.001
Model 6	Random & fixed Model 5 & IMD	470,457	14	<0.001
Model 7	Random & fixed Model 6 & diagnosis	466,498	16	<0.001
Model 8	Random & fixed Model 7 & ethnicity	433,803	20	<0.001
Model 9	Random & fixed Model 8 & chronicity	431,258	23	<0.001
Model 10	Fixed only Model 9 without random factors (centre, surgeon)	432,186*	21	<0.001

\* -2 RLL indicates better fit for model 9 over model 10

All explanatory variables demonstrated a significant influence on the preoperative OKS. Factors associated with a higher preoperative OKS were increasing age, male gender,

lesser deprivation, OA, less comorbidities, white ethnic background, and lesser symptom chronicity.

Significant differences were present in preoperative OKS between both individual centres and individual surgeons, as demonstrated by the model statistics. Although significant differences were evident between surgeons and centre, the VPC calculations showed that surgeons accounted for only 0.7% of total variability ( $VPC = 0.007$ ), whilst differences between centres accounted for only 1.8% of the total variability ( $VPC = 0.018$ ). The remaining variability (97.5%) in preoperative OKS was accounted for by patient-level factors. Centre-cohort ICC again demonstrated minimal correlation between preoperative OKS for patients within the same centre-cohort ( $ICC = 0.024$ ). Patient-cohorts within the same centre but under the care of different surgeons demonstrated similar variability ( $ICC = 0.728$ ). In simple terms, in those centres where the preoperative OKS tends to be higher for a single surgeon patient cohort, other surgeons at the same centre will also tend to exhibit higher preoperative OKS for their patient cohorts as well.

The majority of variability in preoperative OKS existed at the patient-level. Inclusion of the explanatory sociodemographic variables into the final model reduced the amount of variance at centre-level by 52% and the surgeon level variance by 42%. However, these additional variables still accounted for only 13.2% of the total variance in preoperative OKS.

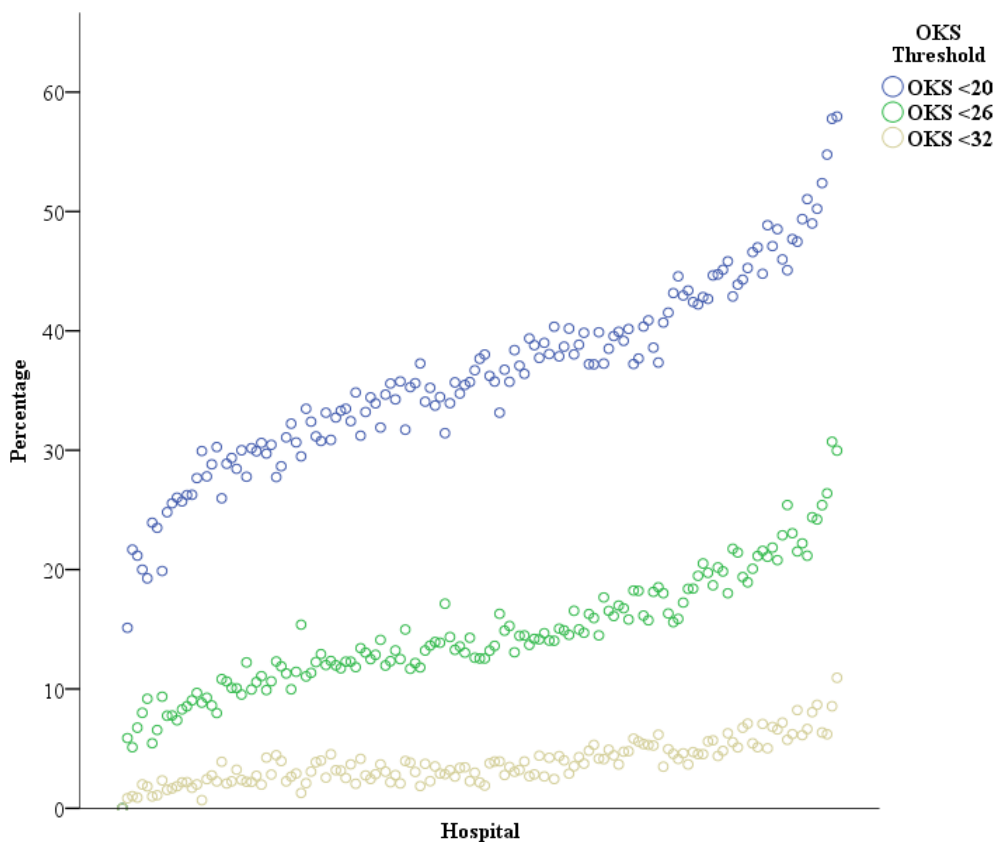
**Table 3-4 Summary of the impact of sociodemographic characteristics on preoperative OKS in the adjusted final multilevel model.**

Variable		Regression coefficients	
		(95% CI)	p-value
<b>Age</b>		0.05 (0.05, 0.06)	<0.001
<b>Gender</b>	Male	3.42 (3.31, 3.54)	<0.001
	Female	<i>Reference</i>	-
<b>Comorbidities</b>	None	5.39 (4.77, 6.01)	<0.001
	1	3.81 (3.19, 4.44)	<0.001
	2	2.38 (1.74, 3.02)	<0.00
	3	1.12 (0.42, 1.82)	0.002
	4+	<i>Reference</i>	-
<b>IMD quintile (Low to high)</b>	1	-3.03 (-3.23, -2.83)	<0.001
	2	-2.04 (-2.22, -1.85)	<0.001
	3	-1.50 (-1.68, -1.32)	<0.001
	4	-0.99 (-1.16, -0.81)	<0.001
	5	<i>Reference</i>	-
<b>Diagnosis</b>	Primary OA	3.03 (2.73, 3.33)	<0.001
	Secondary OA	2.49 (1.77, 3.21)	<0.001
	Inflammatory	<i>Reference</i>	-
<b>Ethnicity</b>	Mixed	-1.51 (-2.66, -0.35)	0.042
	Asian	-2.80 (-3.13, -2.48)	<0.001
	Black	-1.91 (-2.42, -1.39)	<0.001
	Chinese/other	-0.54 (-1.29, 0.21)	0.173
	White	<i>Reference</i>	-
<b>Symptom Chronicity</b>	<1 year	0.75 (0.46, 1.04)	<0.001
	1-5years	0.97 (0.83, 1.12)	<0.001
	6-10 years	0.32 (0.15, 0.49)	<0.001
	>10 years	<i>Reference</i>	-

### 3.3.4 OKS Thresholds

In considering the diverse OKS thresholds currently proposed for utilisation by CCGs, there is a marked difference in the potential impact of these thresholds on different population groups. Amongst the combined cohort, the highest proposed OKS threshold value (OKS<32) would exclude 2796 (4%) patients from the study cohort from proceeding to a TKR. An OKS threshold <26 would exclude 10,494 (15%) patients, whilst the lowest proposed threshold value (OKS<20) would exclude 26,013 (37%) patients. Figure 3-7 demonstrates that the impact of the proposed OKS threshold cut-points would result in major differences across the different centres.

**Figure 3-7 Proportion of TKR patients within individual centres presenting for TKR with preoperative OKS above proposed threshold values**



### 3.4 Discussion

This analysis of preoperative OKS in a cohort of 70,016 patients identified from the national PROMs database provides new and interesting insight into the preoperative symptom burden of patients proceeding to knee replacement in the UK, addressing some of the limitations of the study reported in the Chapter 2. Similarly, wide variability was seen in the overall distribution of preoperative OKS amongst this wider study population. Also, a normal distribution profile was observed, with a distinct lack of a major positive skew that might be expected with a preponderance of patients with high levels of pain and disability. As a national arthroplasty database based analysis, this provides a vital indication of current arthroplasty practice in the UK. Although general variability in preoperative symptom burden appears to exist, a greater concern relates to potential variation between individual centres and surgeons, as this would imply inequity in access depending on provider.

Wide variation was observed in preoperative OKS profiles within all 145 centres examined in this study. Mean preoperative OKS however ranged from only 14.0-22.1, a range which only just sits within a clinically important difference [112]. Although formal analyses demonstrated the presence of statistically significant variation between both centres and individual surgeons, this accounted for only 1.8% and 0.7% respectively, of total variability in the preoperative OKS. Thus these are not considered substantial factors influencing the variability in preoperative symptom burden. As a surrogate marker of preoperative symptom burden, these observations on the preoperative OKS provide further evidence confirming that current arthroplasty practice encompasses a broad spectrum of symptom severity.

Formal studies exploring variation in preoperative symptom burden of TKR patients are limited. Previous studies have primarily focused on the WOMAC score, and generally examined cohorts across different countries and healthcare systems rather than within the same system. This introduces additional biases that may influence some of the variation observed in these studies. The original Atlas of Variation 2010 report aimed to highlight variation across a number of key clinical conditions in the UK [58]. Amongst patients undergoing knee replacement in the UK it reported a twofold variation in the severity of preoperative knee symptoms and disability across different centres. This however, provided only a limited snapshot analysis of the apparent variation, and was based on the Eq5D [119] which is a generic-health PROM. Ackerman et al evaluated patients from a total of 16 European and Australian centres, evaluating variation in pain and function using the WOMAC score [63]. Their study identified significant variability in both pain and physical function prior to TKR, with greater than 8-point difference in WOMAC scores observed. As part of their analyses on TKR outcomes, Lingard et al also identified preoperative status to be significantly different amongst patients in centres from the UK, US and Australia [120]. Interestingly, patients in the UK demonstrated the worst WOMAC scores for both pain and function, as well as SF-36 physical function scores. Fortin et al compared a single US and Canadian centre, and found that the Canadian patients had significantly worse pain, function and QOL [121]. In an analysis of 11 Swedish hospitals, Löfvendahl et al found that although most patients proceeding to TKR were deemed to have severe OA, some variation in subjective classification of disease severity was evident between the centres investigated [28]. Patient self-reported assessment using the Eq5D however was found to be similar across all 11 centres. Contrasting with other studies, Cobos et al suggested that there was very low overall variability in preoperative status, measured by the WOMAC and SF-12 scores, when

evaluated across 15 Spanish hospitals [122]. However, sub-analysis of the subjective and objective elements of these scores revealed that variability was still high for certain signs. These included pain at rest, ambulating capacity, and need for walking aids. Despite low overall variability, the rate of inappropriate surgery reported varied between 7.3-41.7% across the centres. It was suggested that these observations reflected physician adopting different clinical criteria for surgical intervention.

Studies examining variation in preoperative symptom burden between surgeons are extremely uncommon. Herickhoff et al examined preoperative outcome scores in 1896 TKR patients across four different surgeons within the same centre [30]. They found no significant difference in either the preoperative WOMAC function and stiffness domains, or the SF-36, but a small difference was evident in the WOMAC pain domains. At present there are no published studies examining the preoperative status variation as measured specifically by the OKS, or variation between centres within the same healthcare system. Although it is important to appreciate differences between different countries, understanding potential variation within the same healthcare system is of much greater importance and clinical relevance.

The influence of sociodemographic factors on healthcare access and utilisation are commonly reported and are also essential considerations when discussing variation in access to TKR [26,57]. Higher deprivation, younger age, female gender, greater comorbidities, ethnic minorities, greater chronicity and inflammatory arthritis were all associated with a worse preoperative status, although these still accounted for only 13% of total OKS variation. It should also be noted that within these individual sociodemographic groups there was no clinically important difference in the preoperative

OKS observed. Inclusion of sociodemographic factors into the multilevel analyses also further reduced centre and surgeon variation by around 50%. This suggests that although some variation exists across different population groups, individual centre and/or surgeons are not the primary source. Additional patient level factors which cannot be fully assessed may be much more important. This is an important consideration in relation to application of threshold criteria.

Threshold criteria in the UK have not received any central approval or recommendations, and the most recent NICE guidelines advised against this [18]. Consequently, proposed threshold values centred on the OKS are entirely arbitrary, and considerable variation exists across different healthcare providers [94,95,97-100]. When combined with the general variation in the preoperative status of different populations, the impact of arbitrary threshold values are immense. When the lowest proposed OKS threshold (<20) was applied to the individual centre data, over 40% of cases would have been rejected in around 25% of centres. A recent study examining patient with hip or knee OA in New Zealand over a 12-month period, demonstrated that 36% of patients deemed clinically appropriate and who wished to pursue surgery were rejected on the basis of local threshold criteria [123]. Application of arbitrary threshold criteria without understanding the local population group and the role of additional covariates, may therefore in itself create risk of inequity in access to treatment.

The normal distribution profile observed in both this and the previous chapter suggest that the level of pain and disability associated with need for TKR may be considerably lower than anticipated. Pain and disability alone may therefore not fully define the need for TKR, and additional factors may play an important role in the decision making

process. The preoperative OKS as a threshold criterion in isolation may therefore exclude patients with a lower level of pain and disability, but who would otherwise remain eligible candidates for surgery. The findings of the recent OARSI/OMERACT report highlight this potential, in which they failed to identify a specific cut-point to define need for knee replacement [45]. Without a clear understanding of the true indications and the symptom burden that correlates with eligibility for TKR, such threshold criteria must be approached with caution

This study represents the largest published detailed analysis of the preoperative symptom burden in the UK knee replacement population, as measured by the OKS. Unlike previous studies it examines variation within a single health system using data sourced from the national arthroplasty database. The results are therefore representative of current arthroplasty practice in the UK, and therefore highly important and relevant. Despite the large sample size and wealth of data available in the national database, it does not however allow for further interpretation and assessment of the decision making process that ultimately led to the knee replacement procedure. Although proceeding to TKR in the presence of a high OKS implies there must valid reasons for undergoing surgery, no further comments can be made in the current study as to the appropriateness of this decision. Further work is therefore required to gain a greater understanding of this particular population sub-group, in order to define whether TKR may be appropriate despite the apparent milder symptom burden.

## **Chapter Four**

### **4 Knee Replacement in the Presence of a High Preoperative OKS – An Exploratory Pilot Study**

#### **4.1 Introduction**

It is evident from the previous two chapters that within the knee replacement population there is small cohort of patients that undergo knee replacement in the presence of a high preoperative OKS. With current concerns regarding variation in access to surgery, limited healthcare resources, and a drive to optimise postoperative outcomes this particular cohort of patients represent an important group for further evaluation. If it is accepted that significant pain and disability are essential requisites to define eligibility for knee replacement, then these patients seemingly could be considered inappropriate candidates for surgery, and it must be questioned as to why they have proceeded down the surgical pathway?

Furthermore, this observation is not isolated to a few individual surgeons or centres, suggesting that this is indeed common practice, rather than being reflective of a few uncharacteristic or maverick cases. In light of proposals to implement thresholds based on the preoperative OKS, these patients would be excluded from surgical intervention. As a measure of pain and function, one would naturally assume that these patients with high preoperative OKS have a lower symptom burden, and it may be questionable as to whether surgery is appropriate or not. One possibility is that these patients have managed

to bypass normal selection criteria, reflecting problems in the preoperative pathway. The OKS however only measures pain and function across 12 separate domains, and it should therefore also be considered whether in this group of patients, the OKS has failed to fully capture their true symptom burden. These are important considerations to explore. If the latter is true, then implementation of threshold criterion may inappropriately exclude patients with an ‘apparent’ low level of symptoms. Conversely if surgeons are failing to implement satisfactory preoperative consultations and assessment pathways then threshold criteria may be justified as a means to standardise the selection process.

Although the OKS is a validated measure of pain and function [41], it is possible that the limited focus on pain and function may be failing to capture the true impact of OA within these individuals, and thus inaccurately measuring the symptom burden. Alternatively, it may be a true reflection of inappropriate patient selection, which in turn could be due to individual deviant practice or more general weaknesses in the selection process allowing patients to inappropriately slip through the net.

#### **4.1.1 Study Aims**

The aim of this study was to undertake an analysis of patients undergoing knee replacement in the presence of a high preoperative OKS, at an UK elective orthopaedic centre, to determine whether

1. Preoperative OKS reflected the true clinical presentation in relation to pain and function; and

2. Secondly to evaluate whether there was evidence of a breakdown in the consultation process accounting for the surgical decision in these patients.

## **4.2 Methods**

### **4.2.1 Study population**

This study was based on a cohort of patients identified from a prospectively collected database of preoperative OKS collected at the Nuffield Orthopaedic Centre, Oxford. All patients placed onto the waiting list for knee replacement surgery, routinely completed the OKS as part of their preoperative assessment, which was then recorded locally prior to onward submission to the national arthroplasty database. Data was extracted from this local database encompassing the period November 2008 to April 2010, and included a total sample of 1058 cases.

Patients were stratified into three separate groups reflecting their preoperative OKS. An OKS value that is indicative of appropriateness or eligibility for knee arthroplasty has not been defined in the current literature. The mean postoperative OKS in the national PROMS database is 33.7. [106] For the purpose of this study it was therefore decided to define the main study group ('high' OKS) based on selection of those patients with a preoperative OKS  $\geq 34$  (n=50, 4.7%). A second group was defined based on a preoperative OKS 26-33 (n=214, 20.2%), which encompassed the commonly suggested threshold values reported by CCGs. The final group incorporated all remaining patients, with preoperative OKS <26 (n=794, 75.1%) which would reflect those with the highest perceived symptom burden. As this represented a pilot study an equal number of cases

were selected into each group (n=50, total N=150), with cases selected chronologically from the database for inclusion in the study. Although the focus of this study primarily reflected the preoperative symptom burden of TKR procedures, this dataset did not include data on the final operative outcome. All patients had been formally listed for a knee replacement procedure, and it was therefore decided that analyses should incorporate the whole group, in order to fully evaluate the group with a high preoperative OKS. As a result, the analyses include cases where final outcome incorporate both TKR and UKR procedures, as well as some non-arthroplasty and postponed cases. Exclusion criteria included cases originally listed for revision procedures, non-arthroplasty procedures, and inflammatory arthritis. After application of these additional criteria a total of 131 patients were available for further analyses.

### **4.3 Data Analysis**

Retrospective case-note analysis was undertaken, using a combination of electronic patient records and paper-based case-notes. Data was extracted from case-notes using GP referral letters, and orthopaedic clinic letters as the primary source of clinical information. Where relevant data was not included here, additional documentation was examined including nursing, physiotherapy and occupational therapy assessments in relation to the relevant treatment episode. Detailed content analysis of the medical records was undertaken, primarily aiming to evaluate pain and function levels, based on the documented pre-operative clinical presentation. Where conflicting data was identified within the case notes in relation to pain and function, the worst documented category was selected. Basic demographic data was extracted to gain an insight into the characteristics of these different groups to include; age, gender, body mass index (BMI), employment

status, ASA grade, previous treatment and consultations, and surgical details. All clinical data was categorised for analytical purposes as detailed below:

### **4.3.1 Pain**

Pain was evaluated and categorised according to; chronicity (<2yrs / 2-5yrs / 5-10yrs / >10yrs); pain events (night pain / daytime rest pain / only on mobilisation / none); and analgesic group use (none / one / two / three or more). Analgesia use was evaluated to estimate compliance with non-operative intervention, as well as a surrogate marker of pain. Analgesia use was grouped according to simple analgesia (paracetamol), non-steroidal anti-inflammatories (NSAID), and opiate based analgesics. In cases of documented analgesic intolerances, contraindications, or failed response, patients were considered to be compliant with analgesic therapy. The original study plan included an estimation of both pain severity and pain frequency, however due to the quality of data recorded in the records it became apparent that this could not be accurately categorised across the cohort. These categories were therefore not included in the final analyses. Subjective comments such as 'significant pain' were however commonly encountered in the case-notes amongst the whole cohort.

### **4.3.2 Function**

Function was similarly categorised according to; functional limitations (activities of daily living (ADLs) or work activity / walking or standing < 30mins or < 1 mile / walking or standing >30mins or >1 mile / leisure activities only); use of walking aids; and usual activity levels (low / medium / high). Usual activity level was defined according to the

patient's normal (or desired) reported lifestyle and activities; low': ADLs only, with minimal leisure activity; 'medium': still in employment and/or active leisure lifestyle (e.g. walking, bowls, swimming), 'high': heavy manual employment and/or high activity levels (e.g. tennis, skiing).

Data relating to maximal knee flexion was extracted as an objective assessment of knee function. Coronal plane deformity (varus or valgus) was initially considered as an added objective assessment, but this was poorly documented preventing formal assessment.

### **4.3.3 Consultation Process**

As a major surgical intervention, it is imperative that the decision to proceed to TKR is not taken lightly or inappropriately. Although there are no published guidelines that formally define the gold-standard consultation pathway leading to joint replacement, such a structured pathway may help minimise the number of clinically inappropriate cases proceeding to surgery. In order to evaluate whether possible breakdown in the preoperative assessment process could account for cases proceeding to TKR with a high OKS, we also evaluated the consultation process leading to knee arthroplasty. The key criteria considered important, and which were examined in this study included;

1. Formal referral for assessment of the knee problems from GP or other medical professional;
2. Formal review, and listing decision reached in an appropriate arthroplasty clinic;
3. Formal review and clinical assessment in a separate pre-assessment clinic prior to surgery;

4. Documented clinical review by an arthroplasty consultant prior to admission for surgery; and
5. Decision to proceed to TKR documented in the notes (whether appropriate or not).

This information was collated and summary statistics presented to determine whether there was evidence of possible breakdown in the preoperative consultation process. Additionally, the seniority of the medical professional responsible for the decision to proceed to TKR was noted. Where discussion with the responsible consultant was clearly documented at time of listing, this was considered to reflect a consultant decision.

#### **4.3.4 Final intervention**

As this study cohort was based on the preoperative OKS database, and not a postoperative database, it could not be assumed that all patients ultimately proceeded with the proposed surgery. Data was therefore reassessed after initial case-note review to determine the final treatment outcome. Where surgery was postponed or an alternative procedure was performed, notes were examined to identify the primary reason for the change in management.

#### **4.3.5 Statistical Analyses**

All data was extracted and collated for further analyses using SPSSv18 (IBM, Armonk, New York). Categorical data was examined using Fisher's exact test, due to the small

number of cases, and continuous data examined using ANOVA. A p-value <0.05 was considered significant throughout.

## **4.4 Results**

### **4.4.1 Patient Characteristics**

The mean preoperative OKS in the combined cohort (n=1058) was 19.6 (Range 1-43). A higher preoperative OKS was associated with male gender and a lower BMI (Table 1). No significant difference was observed in age or diagnosis across the three OKS groups. There was a suggestion of a trend towards better health (lower ASA grade) and greater level of employment in association with a higher OKS, although these did not reach statistical significance.

**Table 4-1 Summary of the patient characteristics in each of the three OKS groups.**

		Preoperative OKS Group, n (%)			p-value
		<26	26-33	≥34	
Sample size	Study (main cohort)	43 (794)	47 (214)	41 (50)	-
Age	Mean (95% CI)	69.9 (67.1-72.8)	67.8(64.4-71.1)	69.4 (66.5-72.4)	0.566
BMI		31.4 (29.8-33.1)	29.4 (27.6-31.2)	28.2 (26.8-29.5)	0.021
Gender	Male	17 (39.5)	29 (61.7)	27 (65.9)	0.034
	Female	26 (60.5)	18 (38.3)	14 (34.1)	
ASA Grade	1	6 (14.0)	5 (10.6)	9 (22.0)	0.218
	2	26 (60.5)	35 (74.5)	28 (68.3)	
	3	11 (25.6)	6 (12.8)	4 (9.8)	
	4	0 (0.0)	1 (2.1)	0 (0.0)	

#### 4.4.2 Pain

Most patients reported symptom chronicity of at least 2 years duration, with no significant difference observed between the groups. Patients in the lowest preoperative OKS group were more likely to experience night pain, and have greater reliance on analgesia. Although night pain was also present in 37% of those in the highest OKS group, the majority complained of activity-related pain only. One patient in the OKS $\geq$ 34 group reported no pain, but had severe chronic deformity causing significant disability.

**Table 4-2 Summary of the pain related characteristics according to the preoperative OKS**

		Preoperative OKS Group, n (%)			p-value
		<26	26-33	≥34	
Chronicity	<2years	4 (9.3)	4 (8.5)	6 (14.6)	0.100
	2-5 years	19 (44.2)	18 (38.3)	24 (58.5)	
	5-10 years	10 (23.3)	13 (27.7)	8 (19.5)	
	>10 years	9 (20.9)	9 (19.1)	1 (2.4)	
	Not Documented	1 (2.3)	3 (6.4)	2 (4.9)	
Pain Activity	Night	33 (76.7)	26 (55.3)	15 (36.6)	0.001
	Rest pain	1 (2.3)	1 (2.1)	1 (2.4)	
	Mobilisation	7 (16.3)	18 (38.3)	24 (58.5)	
	None	0 (0.0)	0 (0.0)	1 (2.4)	
	Not Documented	2 (4.7)	2 (4.3)	1 (2.4)	
Analgesia	None	3 (7.0)	10 (21.3)	10 (24.4)	0.057
	1	8 (18.6)	14 (29.8)	13 (31.7)	
	2	20 (46.5)	15 (31.9)	8 (19.5)	
	3	12 (27.9)	8 (17.0)	10 (24.4)	

### 4.4.3 Function

Patients with the highest preoperative OKS had less functional restriction ( $p=0.014$ ), although almost half of the  $OKS \geq 34$  cohort had a walking time of less than 30 minutes/1-mile. Similarly, those with the highest preoperative OKS also reported a higher normal preoperative activity level ( $p<0.001$ ). They were also less likely to be using walking aids than those with the lowest OKS ( $p=0.002$ ). Preoperative knee flexion was incompletely recorded ( $n=91$ ) in the notes, but where available no significant difference was observed between the groups;  $OKS \geq 34$ ,  $118.0^\circ$ ;  $OKS 26-33$ ,  $114.8^\circ$ ;  $OKS < 26$ ,  $113.5^\circ$  ( $p=0.534$ ).

**Table 4-3 Summary of the function related characteristics according to the preoperative OKS**

		Preoperative OKS Group, n (%)			p-value
		<26	26-33	≥34	
<b>Functional Restriction</b>	ADL / Work	9 (20.9)	6 (12.8)	9 (22.0)	0.014
	Walking<30min	23 (53.5)	25 (53.2)	10 (24.4)	
	Walking>30 min	2 (4.7)	3 (6.4)	8 (19.5)	
	Leisure activities	0 (0.0)	3 (6.4)	6 (14.6)	
	Not Documented	9 (20.9)	10 (21.3)	8 (19.5)	
<b>Usual Activity</b>	Low	23 (53.5)	4 (8.5)	5 (12.2)	<0.001
	Medium	18 (41.9)	40 (85.1)	25 (61.0)	
	High	1 (2.3)	2 (4.3)	10 (24.4)	
	Not Documented	1 (2.3)	1 (2.1)	1 (2.4)	
<b>Walking Aids</b>	Yes	22 (51.2)	16 (34.0)	6 (14.6)	0.002
	No	21 (48.8)	31 (66.0)	35 (85.4)	
<b>Knee Flexion</b>		113.5°	114.8°	118.0°	0.534

#### 4.4.4 Additional Factors

Assessment of the preoperative consultation pathway demonstrated no significant differences between the three groups in any of the criteria evaluated. No difference was seen between OKS group when considering whether patients had previous consultations for the same problem (p=0.133), previous arthroscopy on the same knee (p=0.462), or previous joint replacement (hip or knee, p=0.479).

**Table 4-4 Summary of the preoperative consultation process according to OKS group**

		Preoperative OKS Group, n (%)			p-value
		<26	26-33	≥34	
<b>Previous Consultation</b>	Yes	35 (81.4)	29 (61.7)	29 (70.7)	0.133
	No	8 (18.6)	18 (38.3)	12 (29.3)	
<b>Previous Arthroplasty</b>	Yes	16 (37.2)	12(25.5)	12 (29.3)	0.479
	No	27 (62.8)	35 (74.5)	29 (70.7)	
<b>Previous Arthroscopy</b>	Yes	21 (48.8)	17 (36.2)	16 (39.0)	0.462
	No	22 (51.2)	30 (63.8)	25 (61.0)	
<b>Arthroplasty Clinic</b>	Yes	35 (81.4)	43 (91.5)	39 (95.1)	0.135
	No	5 (11.6)	1 (2.1)	2 (4.9)	
	Listed Directly	3 (7.0)	3 (6.4)	-	
<b>Pre-op Assessment</b>	Yes	41 (95.3)	45 (95.7)	39 (95.1)	1.000
	No	2 (4.7)	2 (4.3)	2 (4.9)	
<b>Consultant Review</b>	Yes	34 (79.1)	39 (83.0)	35 (85.4)	0.808
	No	9 (20.9)	8 (17.0)	6 (14.6)	
<b>Listing Decision</b>	Consultant	23 (53.5)	31 (66.0)	24 (58.5)	0.774
	Registrar/Fellow	15 (34.9)	13 (27.7)	15 (36.6)	
	Extended role physiotherapist	4 (9.3)	3 (6.4)	2 (4.9)	
	Unknown	1 (2.3)	-	-	

#### 4.4.5 Final Intervention

A total of 115 (87.8%) patients ultimately proceeded to a knee replacement after being originally listed for the procedure. In three cases, an arthroscopy was undertaken which demonstrated less severe degenerative changes, and that knee replacement was not currently indicated. There was a trend that partial knee replacement was commoner in the higher OKS group. More cancellations were observed in the highest OKS group, although this did not reach statistical significance (OKS<26 =2(4.5%); OKS 26-33=4 (8.5%); OKS≥34 = 7 (17.1%), p=0.093). Improvement in symptoms was the primary reason for cancellation in 5 (71.4%) cases with OKS≥34. One patient was offered an

arthroscopy at the pre-assessment clinic which was declined. A further patient declined to proceed at the pre-assessment clinic, although the reason for this was unclear, and no documented reason was given for a further patient. Symptom improvement was also observed in the other groups (OKS<26 = 1 (33.3%); OKS 26-33 = 3 (75%)), although concerns regarding medical co-morbidities also contributed within all but one of these cases.

**Table 4-5 Summary of the final intervention performed according to OKS group.**

	Preoperative OKS Group, n (%)		
	<26	26-33	≥34
TKA	20 (46.5)	18 (38.3)	13 (31.7)
UKR	20 (46.5)	23 (48.9)	18 (43.9)
PFR	0 (0.0)	2 (4.3)	1 (2.4)
Arthroscopy	1 (2.3)	0 (0.0)	2 (4.9)
Cancelled	2 (4.7)	4 (8.5)	7 (17.1)

All interventions, p=0.361; Arthroplasty Vs. arthroscopy/cancellation, p=0.093.

## 4.5 Discussion

A key assumption in utilising the preoperative OKS as a threshold criterion is that this single measure accurately reflects the preoperative symptom burden. Although there is little doubt about the validity of the OKS as a measure of pain and function, [71] this study raises some questions about its role, when used alone, as an indicator of surgical eligibility. Patients in the higher OKS group had evidence of less severe pain and better overall function, suggesting that an upper OKS threshold could indeed be a useful indicator of surgical need. However, even though this group was represented by very

high OKS values ( $\geq 34$ ) there was also evidence to suggest that many of these patients remained clinically appropriate for surgery. Almost 40% had rest or night pain, and almost 50% were limited to less than 30 minutes of activity, both of which might be considered to reflect a significant symptom burden [88]. Furthermore, there was no evidence of any major discrepancy in the consultation process between the three groups, suggesting that the decision to proceed to surgery was based on sound and valid clinical reasons, and not purely a reflection of flaws in the selection process. This implies that there may be additional factors, not captured by the OKS alone, which influenced the decision to proceed to knee replacement. This is a concern if simple measures such as the OKS are to be utilised as the only measure to define surgical eligibility.

Even if we accept that additional factors may be important in the decision to pursue TKR, this study raises important questions regarding this small group of patients with high preoperative OKS. The clinical presentation recorded in the notes suggests that overall these patients had a significant clinical symptom burden, but this was not reflected in the OKS. One possibility is that these patients with a high OKS have an altered perception of their symptom burden [124]. In a study evaluating pain sensitivity in patients with symptomatic knee OA, King et al demonstrated that the individuals classed as having high symptom levels were more sensitive to all experimental pain stimuli [125]. Patients with a high OKS may still experience severe and regular pain, but due to a different symptom perception may behave differently. In tolerating their symptoms, they may opt to not utilise analgesia, and persevere to maintain as much function as is feasibly possible thus achieving higher preoperative OKS. Differences in coping strategies have also been shown to be directly related to the pain and disability levels in knee OA [126]. Patients that demonstrated activity avoidance coping techniques

were found to have greater disability and pain. Interestingly though, those patients that initially utilised pain distraction strategies to maintain high activity levels also demonstrated higher future pain levels. The only group not associated with higher pain and disability were those that lowered the intensity of their activity level in response to pain. Thus patients that persevere to maintain higher activity level may in fact have greater clinical need than we may presume. Attitudes to the disease and management options may also be influential. Smith et al concluded in their systematic review that there is generally a negative approach towards non-operative management, particularly in lower limb OA, with attitudes of inevitability towards surgical intervention [52]. It is possible that this higher OKS group may simply reflect patients that have a lower personal threshold for intervention.

If these patients are coping with their symptoms, it could be argued that these patients should have lower priority for surgery. However, such decision cannot be considered in isolation, and additional factors such as potential outcome from surgery needs to be taken into account. If delaying these patients with a slightly different symptom burden profile, until they are considerably worse, subsequently leads to worse outcome then they may achieve worse outcomes

A slightly higher cancellation rate was observed in the high OKS group, and greater care may therefore be required when reaching a decision to pursue TKR in this particular group to ensure that the individual symptom burden is significant and sustained over a period of time. McHugh et al encountered an overall cancellation rate of 23% in their study of 105 patients listed for TKR, which is slightly higher than observed in the current study [127]. Self-delay/cancellation, and conversion to knee arthroscopy accounted for a

third of these cancellations. However, they found no statistical difference in preoperative OKS between those that did and did not proceed to surgery. Variability in OKS over time could potentially explain some of the discrepancy between clinical symptoms and OKS observed. It is possible that at the time of the original listing decision the symptom burden may have been higher, and that this subsequently subsided resulting in a higher OKS at the pre-assessment visit. Even though most patients underwent further review by a surgeon at the pre-assessment visit, it may be that the decision process was too advanced for patients to be swayed away from surgical intervention.

This study was conceived as an initial exploratory analysis to gain insight into the apparent lower symptom burden in those with a high preoperative OKS. A further in-depth prospective based study had been considered as a further development of this analysis, but could not be undertaken within the confines of this current thesis. It should be acknowledged that there are therefore significant limitations to the current study. As a small, single-centre retrospective case-note analysis, this study cannot fully explore the individual indications and reasons behind the decision to pursue surgery in each of these cases, due to the limitations of data documentation in the clinical case notes. Additionally, a separate sub-analysis of the individual pain and function domains of the OKS in relation to the clinical presentation could not be performed as this data was unavailable in the local records. However, as an initial exploratory study it highlights some important issues in relation to the preoperative OKS as a possible threshold criterion in the UK, which have not been previously reported. Further prospective work should be focussed on gaining a greater understanding of the decision process, and reasons for pursuing knee replacement in the presence of a high OKS. Furthermore, this study did not examine the subsequent outcome following knee replacement in this

specific group of patients, and this is an additional important factor that should be further examined before considering implementing preoperative OKS as a criterion defining eligibility for surgery. Despite the limitations discussed, it does highlight the need for further prospective analysis to gain a better understanding of the group of patients undergoing TKR with a high preoperative OKS.

In summary, this study demonstrated that most patients undergoing knee replacement, even in the presence of a high preoperative OKS, had undergone an appropriate consultation process prior to surgery. A significant preoperative symptom burden was also reflected in the clinical presentation, suggesting that surgery remained clinically justified. This raises concerns regarding the current trend towards implementation of arbitrary threshold criteria for knee replacements, which may not fully capture the true symptom burden.

## **Chapter Five**

### **5 Preoperative Symptom Burden and TKR Outcomes— Less to gain with a higher preoperative OKS?**

#### **5.1 Introduction**

Total knee replacement is a well-established and highly successful procedure in the treatment of end stage knee OA, restoring function and relieving pain, with published satisfaction rates generally in excess of 80% [128,129]. Concerns remain however that up to 20% of patients are still dissatisfied with surgery despite a large positive treatment effect in many patients [130]. In recent years there has been increasing focus on identifying the cause(s) of this observation, aiming to further optimise postoperative outcome following TKR [131,132]. Although dissatisfaction and poor outcome reflects a complex multifactorial problem, there is evidence in the literature demonstrating a direct relationship between preoperative status and postoperative outcomes [120,133,134].

Those with the worst preoperative status are commonly considered to have the greatest potential to benefit from surgery, and this is supported by a number of studies evaluating postoperative change in outcome scores [135,136]. Conversely, studies focussing on absolute postoperative scores, demonstrate better outcomes in those with better preoperative status [134,137]. Judge et al found that those with postoperative change in  $OKS > 14$ , or total postoperative  $OKS > 30$  had highest levels of satisfaction [138]. Patients with the highest preoperative OKS however, required a smaller change in OKS

to achieve satisfaction compared to those with a lower preoperative OKS yet still achieving a high final postoperative OKS. Defining a successful outcome is therefore not straightforward, and is highly dependent on the metric utilised. A key question is whether absolute postoperative scores or the actual change in score best represent a successful outcome. Although there is no clear consensus, the change in score is generally considered to be the best indicator of the gains achieved from surgery [139]. In the context of preoperative OKS thresholds for TKR, this would seemingly support their role, limiting intervention to those patients that have the greatest potential improvement in OKS.

The previous chapters have highlighted considerable variability in the preoperative status of patients currently proceeding to TKR in the UK. More importantly they have also highlighted that a small cohort of patients proceed to surgery in the presence of a 'high' preoperative OKS, possibly reflecting a lesser symptom burden. In light of the relationship evident between pre and postoperative outcomes, this wide variability in preoperative status should certainly raise concerns regarding the potential impact on subsequent outcomes. If there is very little potential gain for those patients with a lower preoperative symptom burden, then it is questionable whether surgery is the correct option in this subgroup of patients.

The group of patients proceeding to TKR in the presence of a high preoperative OKS, in fact represent only a small proportion of the TKR population, and thus can be difficult to fully interpret. Earlier studies have generally been too small to provide detailed analysis of outcome at the extreme ends of the preoperative OKS scale, instead reporting general trends [137], or have simply focussed attention on defining specific cut-points associated

with success or satisfaction [130,138]. It therefore remains unclear as to the expected outcome in those patients with a seemingly lesser preoperative symptom burden. Having identified in the previous chapters that it is common practice across the UK for TKR to be performed in the presence of a high OKS, it is imperative that we gain a greater understanding of the implications of this practice on postoperative outcomes. If outcome is indeed compromised within this group of patients, this would support the possible implementation of OKS based thresholds.

### **5.1.1 Study Aims**

The aim of this study was to examine the relationship between preoperative OKS and postoperative PROMs in the national dataset, with a specific focus to determine whether TKR in the presence of a high OKS is justified based on the postoperative outcome.

## **5.2 Methods**

### **5.2.1 Study Population**

This study utilised data from the combined HES and PROMs datasets, which include data pertaining to sociodemographic data as well as pre- and postoperative PROMs outcome data. Approval was requested to access the PROMS and corresponding HES data collected between April 2009 to August 2011 from the relevant authorities as previously detailed in Chapter 3. The data cleaning and merger process is summarised in Figure 5-1. Initial data cleaning was performed on each database separately. Following initial exclusions, the datasets were merged using two independent and unique identifiers

(EPIKEY & HES Identifier No.) to combine the PROMs data with its respective HES counterpart. Inclusion criteria were all TKRs performed, where the preoperative questionnaires were completed 90 days or less, prior to surgery. Unicompartamental and revision procedures were again excluded from the analyses. Due to the specific interest in examining the postoperative outcomes, additional criteria were applied to this study pertaining to the postoperative outcomes. Only cases where an additional postoperative OKS was collected between 180-365 days postoperatively were included, in line with the standard PROMS data collection process, and to ensure that all cases analysed were representative of postoperative outcome. Cases where the date of surgery or PROMs collection dates were missing, or pre and/or postoperative OKS were incomplete were excluded from further analysis. A further cleaning process was performed post-merger to deal with unmatched cases, and to apply the full inclusion and exclusion criteria as detailed previously. The final study cohort included 55,886 patients for further analysis.

### **5.2.2 Data Analyses**

The primary outcome measure in this study was the OKS. Postoperative outcome was assessed using the postoperative OKS, together with calculation of the postoperative *change in OKS* to determine the relative improvement in pain and function from surgery. As part of the national arthroplasty outcome programme, additional data is collected in the form of two Likert-based questions:

- 1) How would you describe the results of your operation?
  - Excellent, Very good, Good, Fair, or Poor; and
- 2) Overall, how are your knee problems now, compared to before your operation?

- Much better, A little better, About the same, A little worse, or Much worse.

These additional anchoring questions formed the basis for the secondary outcomes analysed in this study, to determine the patient overall perceived success following TKR. As the main aim was to establish success using a dichotomous model i.e. whether the procedure was successful or not, and to ease interpretation the Likert-based responses were dichotomised as follows:

1) How would you describe the results of your operation?

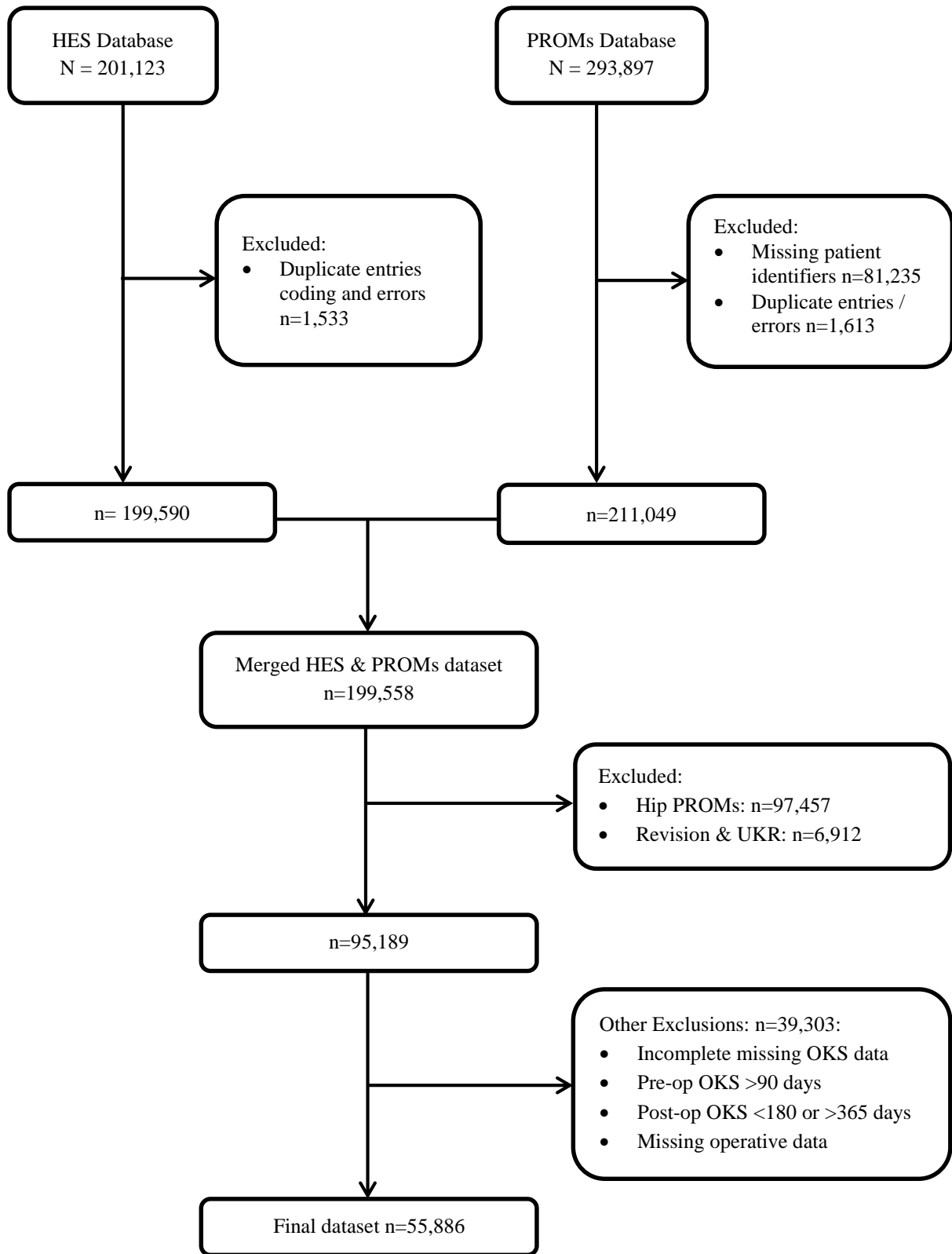
- *Successful* (excellent, very good, good) and *Unsuccessful* (fair, poor); and

2) Overall, how are your knee problems now, compared to before your operation?

- *Better* (much better, a little better), and *Same/worse* (about the same, a little worse, much worse)

Both the preoperative and postoperative OKS data (primary) were complete for all cases reflecting the initial selection criteria (n=55,886). Secondary outcome data was however incomplete in a small number of cases for both Question 1 (missing=422, 0.8%) and Question 2 (missing=256, 0.5%).

**Figure 5-1** Flowchart summarising the data extraction and selection process leading to the final study cohort



### **5.2.3 Descriptive summary statistics**

Statistical analyses were undertaken using SPSSv20 (IBM, Armonk, New York). Descriptive summary statistics and histograms were initially reviewed to evaluate the distribution of preoperative OKS, postoperative OKS, as well as the postoperative change in OKS. Probability plots were analysed to determine any deviation from normality. Simple correlations between preoperative OKS and the postoperative outcome measures were plotted graphically using histograms and scatter plots to determine any general trends in data prior to further formal statistical analyses

The primary outcome measures under investigation were the postoperative change in OKS, and the postoperative OKS. These were analysed using linear regression modelling. Simple regression modelling was initially performed utilising preoperative OKS as the only independent variable in the model in order to determine the overall trend in the data. Previous work in Chapter 3 based on the same parent databases demonstrated a relationship between preoperative OKS and a number of sociodemographic covariates (age, gender, comorbidities, deprivation and ethnicity). Further regression models were therefore generated incorporating these additional covariates sequentially into the model to account for the potential confounding effect of these factors on the results. Categorical predictors were dummy-coded to allow incorporation into the models.

Analyses of the initial summary statistics and plots suggested that the trend between preoperative OKS and postoperative OKS / OKS change may not be entirely linear in nature. There was a suggestion that a different linear relationship existed for different ranges of the preoperative OKS. In addition to a simple linear regression model, a further

model was therefore generated where preoperative OKS was represented by both a quadratic and non-quadratic term and compared to determine the best model fit. Further analyses of the apparent change in relationship between preoperative OKS and postoperative change in OKS, were also performed using a piecewise regression model. This represents a modification of the simple linear regression model whereby the independent variable is split at defined breakpoints or splines, creating two or more groups. A separate regression models is subsequently generated for each group, and the resulting regression models can be compared to determine the change in linear trend across the different ranges of the independent variable. Based on analyses of the initial summary statistics and plots, the piecewise regression model was generated utilising a single breakpoint, creating two groups; preoperative  $OKS < 16$  and preoperative  $OKS \geq 16$ .

Secondary outcomes were analysed based on the dichotomous ratings detailed above, utilising binary logistic regression modelling. Due to the small number of cases within the preoperative OKS at each end of the spectrum, the independent variable (preoperative OKS) was grouped in blocks of 4 to create 12 separate preoperative OKS groups encompassing the full range from 0-48. The likelihood of success (Q1) and likelihood of achieving improvement in knee symptoms (Q2) was determined through generation of odds ratios (OR), with the lowest preoperative OKS group (0-3) acting as the reference group. As with the linear regression model, analyses were initially performed using simple unadjusted models, together with additional adjusted models incorporating covariates; age, gender, comorbidities, deprivation and ethnicity. A p-value  $< 0.05$  was considered significant throughout.

### 5.3 Results

The sociodemographic characteristics of the study cohort are summarised in Table 5-1. The mean age of the cohort was 69.4 years (SD 9.0, Range 20-97), and there were more women (31,592; 56.5%) compared to men (24,294; 43.5%).

**Table 5-1 Summary of the sociodemographic characteristics of the study population**

	Number of patients (%)
<b>Gender</b>	
Male	24,294 (43.5)
Female	31,592 (56.5)
<b>IMD quintile (low to high)</b>	
1	7,663 (13.7)
2	10,203 (18.3)
3	12,394 (22.2)
4	12,888 (23.1)
5	12,109 (21.7)
Missing data	629 (1.1)
<b>Self-reported comorbidities:</b>	
None	33,300 (59.6)
1	16,027 (28.7)
2	5,015 (9.0)
3	1,225 (2.2)
4+	319 (0.6)
<b>Ethnicity</b>	
Mixed	108 (0.2)
Black	502 (0.9)
Asian	1,133 (2.0)
Other	239 (0.4)
White	48,727 (87.2)
Missing data	5,177 (9.3)

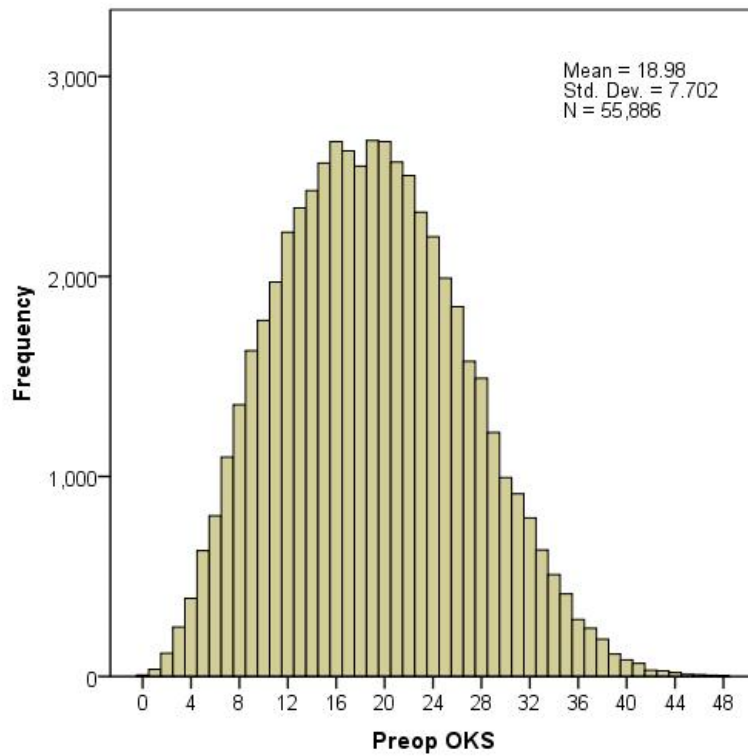
Distributions of the preoperative, postoperative, and postoperative change in OKS are demonstrated in Figure 5-2, Figure 5-3, and Figure 5-4 respectively. Both preoperative OKS and the postoperative change in OKS demonstrate a normal distribution curve, whereas postoperative OKS is negatively skewed, with scores clustered at the higher end

of the scale, reflecting improvement following surgery. Only 1585 (2.6%) patients achieved the highest possible postoperative score (OKS=48), indicating the absence of a significant ceiling effect. Although postoperative OKS showed deviation from normality, summary statistics (Table 5-2) showed that this was not a severe deviation. On this basis, and in line with previous recommendations relating to statistical analyses of the postoperative OKS, this was therefore considered suitable for standard parametric analyses for the remaining study analyses [71].

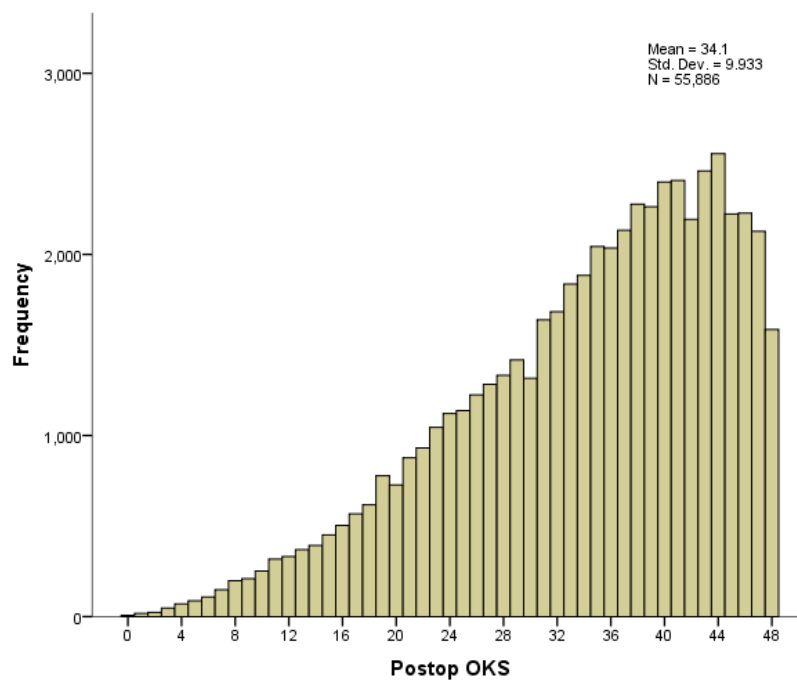
**Table 5-2 Summary statistics for the preoperative OKS, postoperative change in OKS, and postoperative OKS in the study population**

	<b>Preoperative OKS</b>	<b>Postoperative OKS change</b>	<b>Postoperative OKS</b>
Mean	19.0	15.1	34.1
95% Confidence Interval for Mean	18.9-19.0	15.0-15.2	34.0-34.2
5% Trimmed Mean	18.8	15.3	34.7
Median	19.0	15.0	36.0
Variance	59.3	98.0	98.7
Std. Deviation	7.7	9.9	9.9
Minimum	0	-29	0
Maximum	48	46	48
Range	48	75	48
Interquartile Range	11	13	14
Skewness	0.26	-0.21	-0.72
Kurtosis	-0.35	-0.10	-0.15

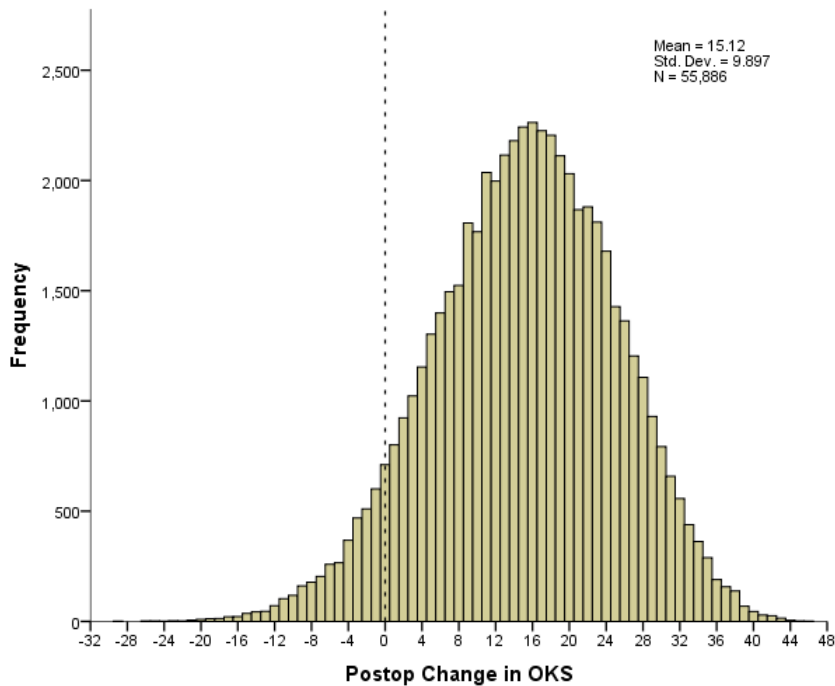
**Figure 5-2 Histogram of preoperative OKS demonstrating the normal distribution profile**



**Figure 5-3 Histogram of postoperative OKS demonstrating a negative skew to the data, reflecting improvement in postoperative status**



**Figure 5-4** Histogram of postoperative change in OKS following TKR demonstrating a normal distribution. Dashed line represents no change in OKS, showing that most patients achieve some improvement postoperatively.

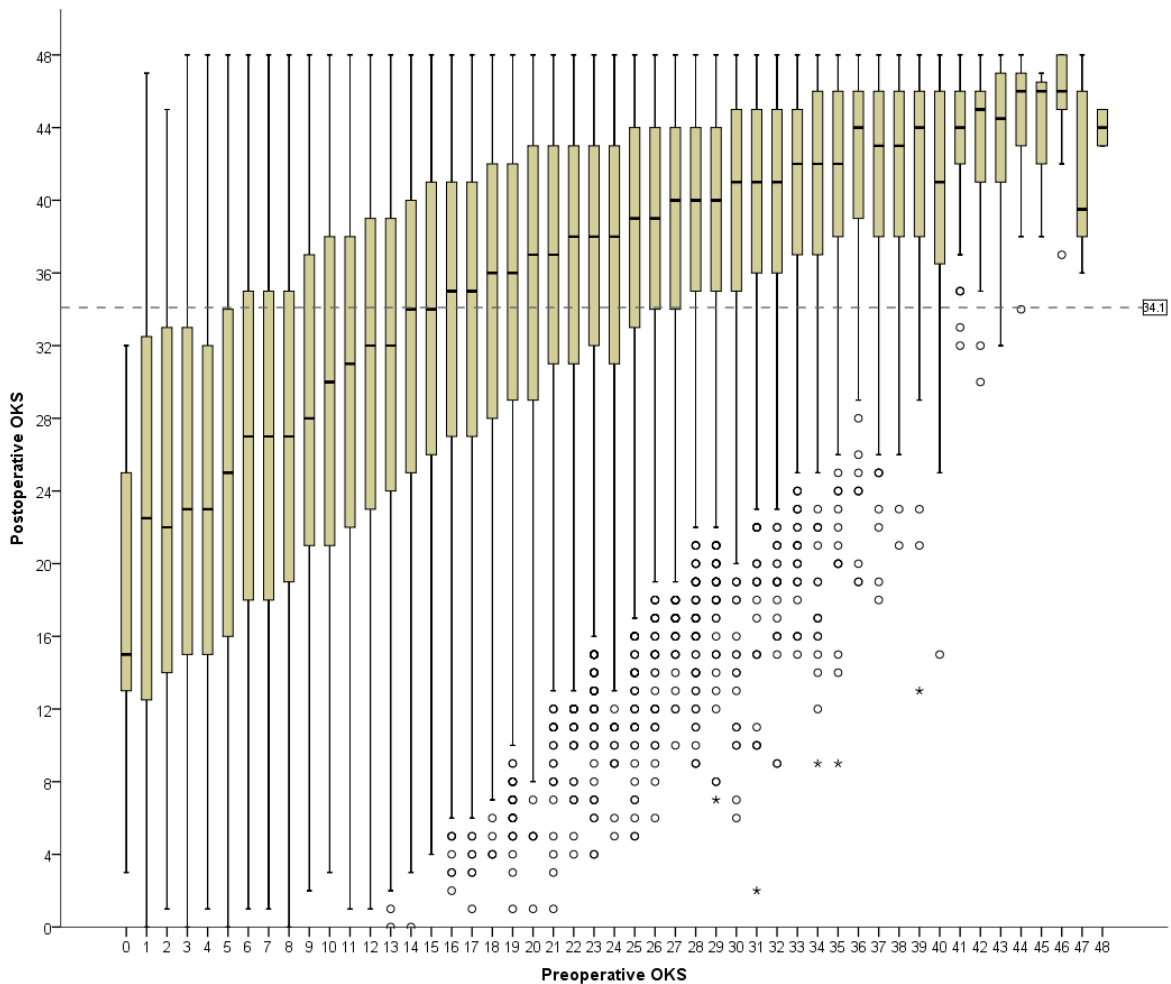


### 5.3.1 Postoperative OKS and Change in OKS

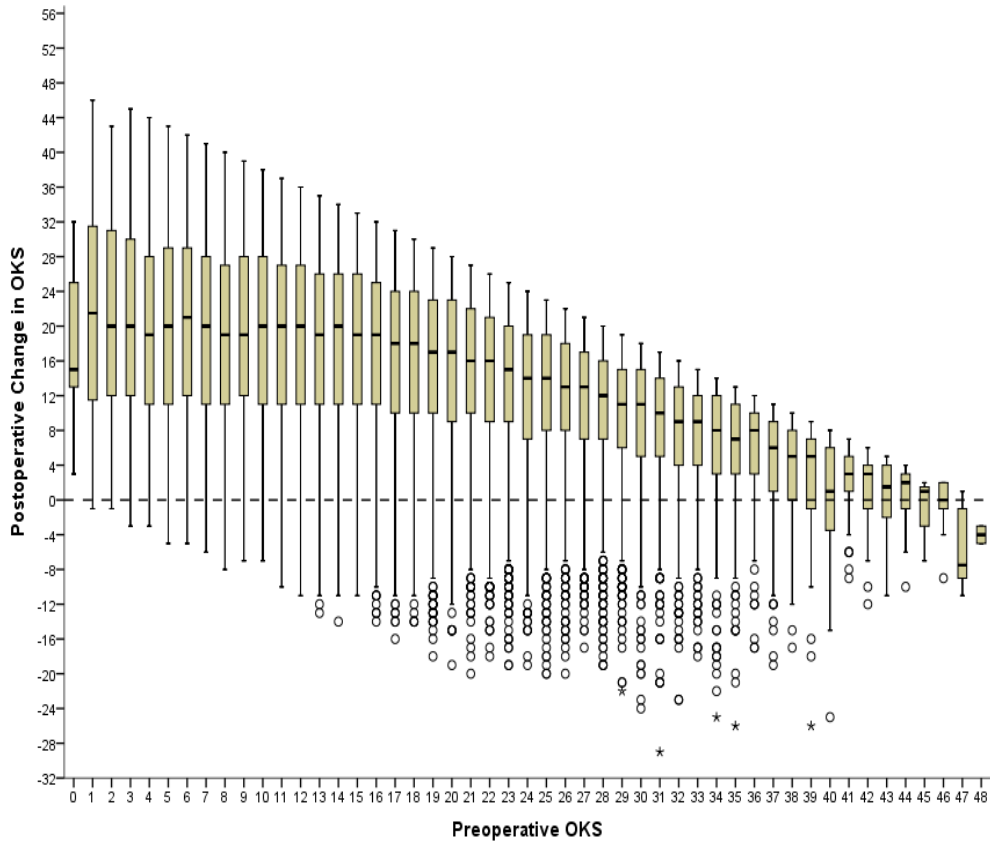
The overall relationship between preoperative OKS and subsequent postoperative outcomes is demonstrated in Figure 5-5 and Figure 5-6. As preoperative OKS increased the final postoperative OKS also increased showing that the final postoperative OKS was greatest for patients with best preoperative status. Conversely, the postoperative change in OKS was greatest for those with the lowest preoperative OKS. Closer inspection of the change in OKS (Figure 5-6) shows that the overall trend does not appear to be completely linear in nature. It can be seen that for preoperative OKS values  $\geq 16$  there appears to be a trend of increased postoperative change in OKS as the preoperative OKS decreases. For preoperative OKS values  $< 16$  however, this trend appeared to become more static. Similarly, for the postoperative OKS, there was a suggestion of two specific

trends (Figure 5-5). For preoperative OKS values  $\geq 16$ , the postoperative OKS gradually decreased for each unit decrease in preoperative OKS. For preoperative OKS values  $< 16$  however there appeared to be a greater decrease in postoperative OKS for each unit decrease in preoperative OKS.

**Figure 5-5** Boxplots demonstrating the distribution of the postoperative OKS in relation to preoperative OKS. Circles and stars represent outlying data points ( $\circ > 1.5$  IQR,  $* > 3$  IQR). Dashed line represents mean postoperative OKS (34.1)

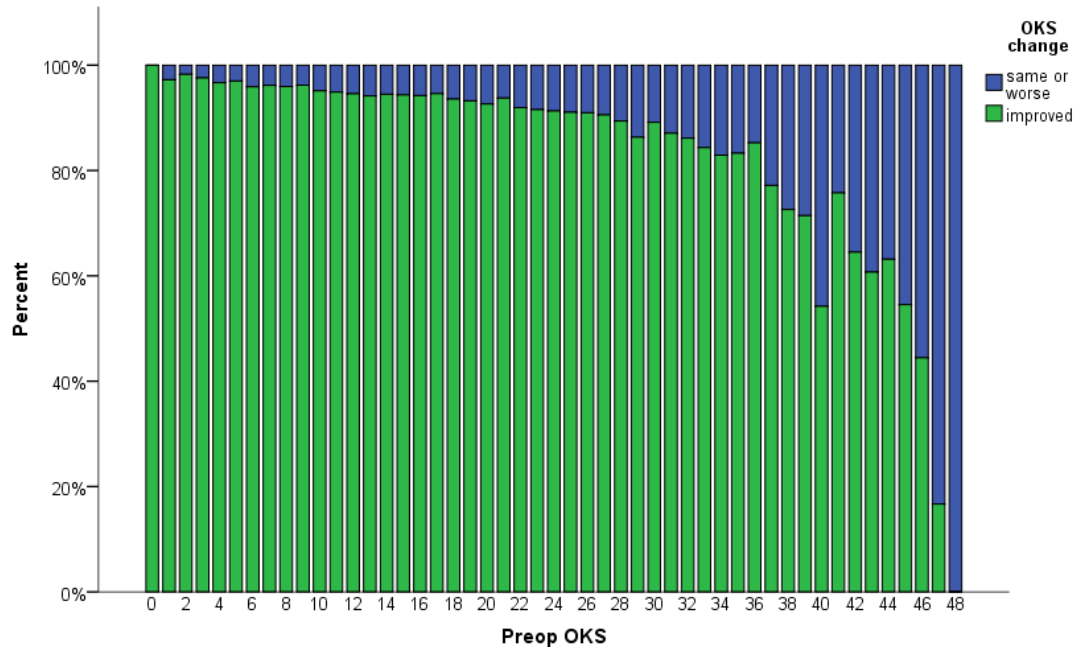


**Figure 5-6** Boxplots demonstrating the distribution of postoperative change in OKS in relation to the preoperative OKS. Dashed line represents no change in OKS postoperatively. Circles represent outlying data points ( $>1.5$  IQR, \*  $>3$  IQR). Dashed line represent postoperative OKS change=0.



Although those with the highest preoperative OKS gained higher overall postoperative OKS, Figure 5-7 demonstrates that they were more likely to see no change or even deterioration in OKS following surgery. For preoperative OKS values  $>34$ , between 20-50% of patients demonstrated no improvement in OKS following TKR.

**Figure 5-7** Histogram demonstrating the percentage of patients achieving improvement in OKS following TKR, according to their preoperative OKS. As preoperative OKS increases a greater proportion of patients achieve no postoperative improvement in their OKS.



Linear regression modelling confirmed the general trend between preoperative OKS and the postoperative change in OKS ( $R^2 = 0.177$ ,  $F(1, 50693) = 727.5$ ,  $p < 0.001$ ). The regression equation demonstrated that for each unit increase in preoperative OKS the postoperative change in OKS decreased by 0.55 ( $B = -0.55$ ; 95%CI (-0.56, -0.54);  $t(50693) = -100.46$ ,  $p < 0.001$ ). Similarly, a significant relationship was evident between preoperative OKS and postoperative OKS ( $R^2 = 0.189$ ,  $F(1, 50693) = 785.56$ ,  $p < 0.001$ ). For each unit increase in preoperative OKS the postoperative OKS increased by 0.45 ( $B = 0.45$ ; 95%CI (0.44, 0.46);  $t(50693) = 81.27$ ,  $p < 0.001$ ). Introduction of a quadratic term resulted in significant improvement in model fit confirming that the overall trend was not completely linear in nature ( $F$  change (50693) = 247.81,  $p < 0.001$ ).

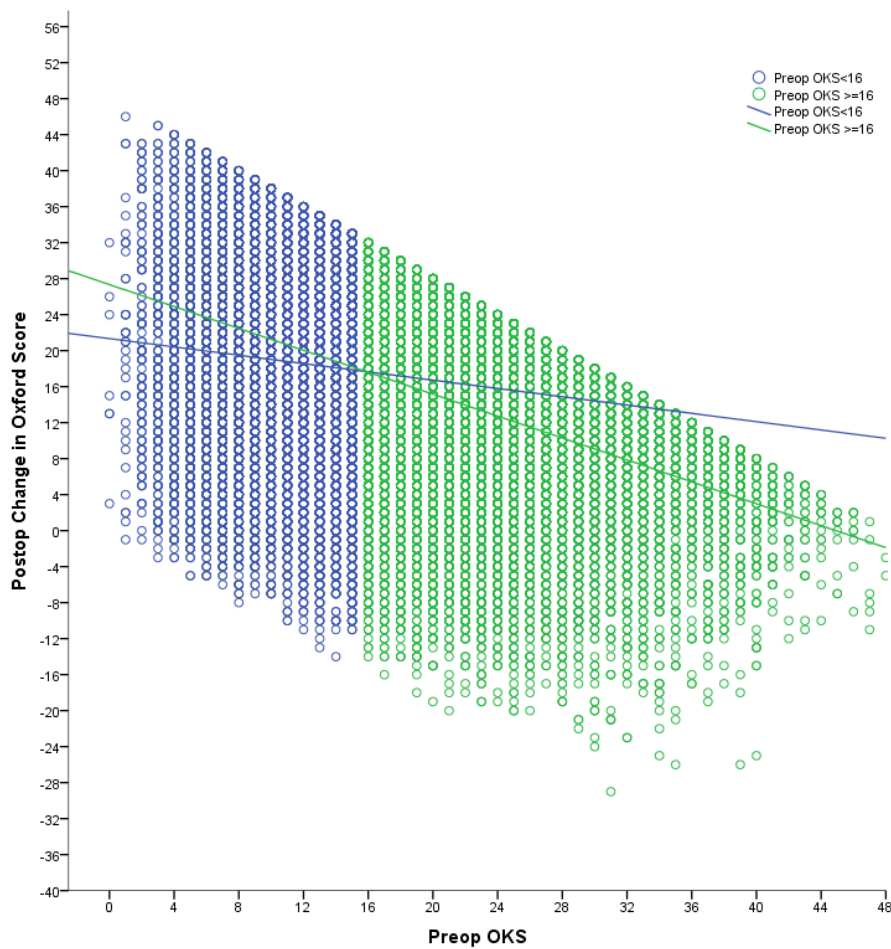
Further exploration of this deviation from linearity using a piecewise regression model demonstrated a significant change in relationship between postoperative change in OKS and the preoperative OKS at a break-point of  $OKS < 16$  versus  $OKS \geq 16$ . A significant linear trend remained evident for both preoperative OKS values  $< 16$  ( $B = -0.32$ ; 95% CI (-0.36, -0.28);  $t(50691) = -15.2$ ,  $p < 0.001$ ) and preoperative OKS values  $\geq 16$  ( $B = -0.65$ ; 95% CI (-0.67, -0.63);  $t(50691) = -70.9$ ,  $p < 0.001$ ). The difference in trend for preoperative OKS  $< 16$  and preoperative  $\geq 16$  was statistically significant ( $B = -0.33$ ; 95% CI (-0.38, -0.29);  $t(50691) = -14.7$ ,  $p < 0.001$ ), although there was no significant change in the intercept at this cut-point ( $B = -0.15$ ; 95% CI (-0.45, 0.15);  $t(50691) = -1.0$ ,  $p < 0.321$ ) (Figure 5-8).

### 5.3.2 Secondary postoperative outcomes

The overall relationship between the preoperative OKS and the secondary postoperative outcome measures is demonstrated in Figure 5-9 and Figure 5-10. Overall, 83.4% ( $n = 46,253$ ) patients reported that the TKR procedure was successful, and 88.3% ( $n = 55,463$ ) reported knee symptoms to be improved following TKR. Results of the logistic regression models examining the relationship between preoperative OKS and the secondary outcome measures are summarised in Table 5-3 and Table 5-4. Patients with a higher preoperative OKS were significantly more likely to report a successful outcome following TKR. Those with the highest preoperative OKS (44-48) were five times more likely to perceive the procedure to be successful ( $p < 0.028$ ). Furthermore, for those patients with a preoperative OKS value between 4-15, there was no significant difference in the likelihood of a successful outcome, compared to those with lowest preoperative OKS (0-3). Similarly, the likelihood of reporting an improvement in knee symptoms was

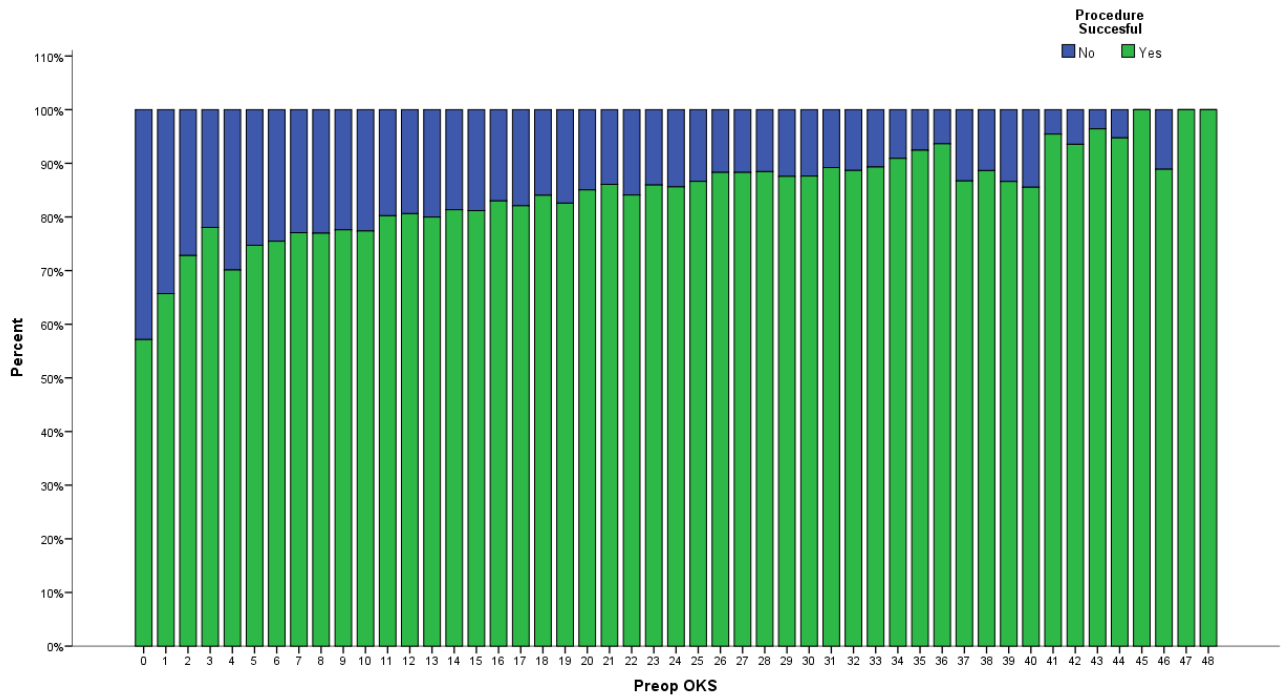
greater in those patients with a higher preoperative OKS, with the odds of improvement up to 3.5 times greater. This did not however reach statistical significance in the highest preoperative OKS group ( $p=0.337$ ).

**Figure 5-8** Scatter plot demonstrating the relationship between preoperative OKS and postoperative change in OKS. Linear regression trend lines demonstrate the difference in the relationship for those patients with preoperative OKS <16 (blue) compared to those with preoperative OKS  $\geq$ 16 (green).



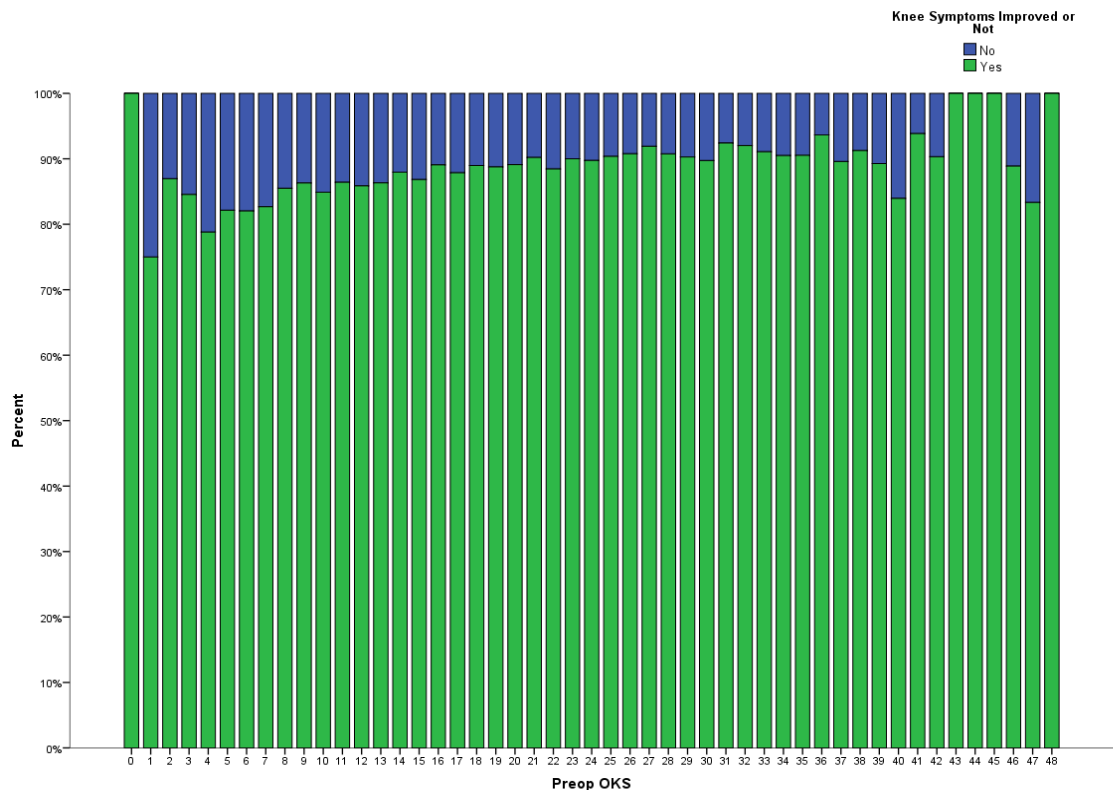
Preoperative symptom burden and TKR outcomes

**Figure 5-9** Histogram demonstrating percentage of patients reporting a successful outcome in relation to the preoperative OKS. As preoperative OKS increases a greater proportion of patients report a successful outcome.



Preoperative symptom burden and TKR outcomes

**Figure 5-10** Histogram demonstrating the percentage of patients reporting an improvement in knee symptoms following TKR according to the preoperative OKS. As preoperative OKS increases a greater proportion of patients report improvement in knee symptoms.



**Table 5-3 Summary of results from the logistic regression model examining relationship between preoperative OKS and the whether the procedure was deemed successful. Results are presented for both unadjusted and adjusted models presented using Odds Ratios (OR) of achieving a successful outcome.**

<b>Preoperative OKS</b>	<b>Number of cases</b>	<b>% successful</b>	<b>OR (unadjusted)</b>	<b>OR (adjusted*)</b>	<b>95% CI (adjusted*)</b>	<b>p-value</b>
<b>0-3</b>	375	75.2	Reference	Reference	-	-
<b>4-7</b>	2658	75.1	1.0	1.0	0.7-1.2	0.711
<b>8-11</b>	6117	77.8	1.2	1.1	0.8-1.4	0.561
<b>12-15</b>	8579	80.5	1.4	1.2	0.9-1.5	0.146
<b>16-19</b>	鏢	82.6	1.6	1.3	1.0-1.7	<b>0.022</b>
<b>20-23</b>	8892	84.8	1.8	1.5	1.2-1.9	<b>0.001</b>
<b>24-27</b>	6714	86.8	2.2	1.7	1.4-2.2	<b>&lt;0.001</b>
<b>28-31</b>	4065	87.9	2.4	1.9	1.4-2.4	<b>&lt;0.001</b>
<b>32-35</b>	2038	89.9	2.9	2.3	1.7-3.1	<b>&lt;0.001</b>
<b>36-39</b>	713	88.9	2.6	2.1	1.5-2.9	<b>&lt;0.001</b>
<b>40-43</b>	175	91.4	3.5	2.8	1.5-5.0	<b>0.001</b>
<b>44-48</b>	40	95.0	6.3	5.0	1.2-21.4	<b>0.028</b>

\*Model adjusted for age, gender, ethnicity, deprivation index, & comorbidities

**Table 5-4 Summary of results from the logistic regression model examining relationship between preoperative OKS and reported improvement in knee symptoms postoperatively. Results presented for both unadjusted and adjusted models presented using Odds Ratios(OR)**

Preoperative OKS	Number of cases	% improved	OR (unadjusted)	OR (adjusted*)	95% CI (adjusted*)	p-value
<b>0-3</b>	377	85.1	Reference	Reference	-	-
<b>4-7</b>	2662	81.7	1.1	1.0	0.9-1.3	0.469
<b>8-11</b>	6141	85.5	1.3	1.2	1.0-1.5	<b>0.027</b>
<b>12-15</b>	8622	86.5	1.5	1.3	1.1-1.7	<b>0.001</b>
<b>16-19</b>	9427	88.5	1.7	1.5	1.2-1.8	<b>&lt;0.001</b>
<b>20-23</b>	8920	89.0	1.9	1.7	1.3-1.9	<b>&lt;0.001</b>
<b>24-27</b>	6727	90.4	2.2	1.9	1.5-2.2	<b>&lt;0.001</b>
<b>28-31</b>	4069	90.4	2.2	1.3	1.4-2.2	<b>&lt;0.001</b>
<b>32-35</b>	2044	91.0	2.2	2.1	1.4-2.4	<b>&lt;0.001</b>
<b>36-39</b>	711	91.3	1.9	1.6	1.1-2.2	<b>0.010</b>
<b>40-43</b>	173	90.8	4.2	3.5	1.5-8.2	<b>0.004</b>
<b>44-48</b>	40	95.0	2.5	2.1	0.5-8.9	0.337

\*Model adjusted for age, gender, ethnicity, deprivation index, & comorbidities

## 5.4 Discussion

This study examining a cohort of 55,886 patients from the national arthroplasty PROMs database has revealed new detailed findings in relation the preoperative OKS and subsequent postoperative outcomes. The primary aim of this study was to examine the postoperative outcome amongst those patients with a high preoperative OKS, and arguably of limited potential to benefit from TKR. The general findings of this study support earlier studies, confirming that the greatest postoperative gain in PROMs is achieved amongst those patients

with the worst preoperative status, [120,121] whereas patients with better preoperative status generally achieved better postoperative outcomes [134,137]. However, it was also noted that many of those with a high preoperative OKS actually gained no improvement or even saw deterioration in OKS postoperatively. This would certainly seem to support arguments to limit TKR in those with a relatively high preoperative status in terms of pain and function, as measured by the OKS.

Interestingly though, this study also identified an unexpected observation in relation to the preoperative OKS which raises significant concerns about the apparent benefits achieved in those patients with the worst preoperative status. Previous studies have argued that those with the worst preoperative status have the greatest potential to benefit [86], and this seems to be supported by the greater gain in PROMs evident in these patients, both in this and earlier studies. Unlike previous studies though, this study has highlighted that this relationship is not entirely linear in nature. Lim et al investigated the relationship between pre and postoperative OKS in a small cohort of 45 patients, and although the data was extremely limited, greater variability in postoperative OKS was evident in those with the worst preoperative OKS [137]. Similarly, in this study, as preoperative OKS decreased it was observed that variability in postoperative OKS became greater, indicating a less predictable postoperative outcome. Furthermore, for patients with a preoperative  $OKS < 16$  there was a significant change in the linear trend, whereby for each unit change in preoperative OKS the postoperative *change in OKS* showed very little further improvement. Although it is accepted that those with worse preoperative status generally achieve lower postoperative total OKS, this is offset by the seemingly greater gains achieved as preoperative OKS continues to decrease. Judge et al found that for patients with lower preoperative OKS, satisfaction was associated with a greater postoperative change in OKS, but compared to patients with a higher preoperative

OKS they required a lower total postoperative OKS [138]. In the absence of a constant linear trend this may not necessarily hold true. In simple terms, for patients with preoperative OKS < 16, for each unit decrease in the preoperative OKS there appears to be very little further improvement in the postoperative change in OKS, whilst their final OKS outcome continues to deteriorate. This suggests that for most patients at the lower end of the preoperative OKS scale there is a limit to the amount of improvement achievable. As preoperative status deteriorates, these patients are therefore at a disadvantage, and are effectively locked into a poorer outcome, despite a seemingly large potential gain in OKS score.

Evaluation of the secondary outcomes revealed further evidence to support the primary findings, as well as cast further doubt on the suitability of the preoperative OKS as a threshold criterion for TKR. Despite demonstrating very minimal improvements in pre to postoperative OKS, those with the highest preoperative OKS were significantly more likely to report a successful outcome, and to achieve resolution of their knee symptoms compared to patients with worst preoperative status. This is an extremely important consideration in relation to arbitrary preoperative OKS based thresholds and proposals for rationalisation of TKR. Use of such thresholds would actually exclude those patients reporting the greatest self-perceived success, and highest rate of symptom resolution. More importantly these OKS thresholds would restrict surgical intervention to those who may be most likely to report poor outcome despite seemingly large postoperative improvements in pain and function. This is of even greater significance when we consider those CCGs proposing very low preoperative OKS thresholds. Scott et al in their study of 1217 patients found that ongoing pain after TKR was the greatest predictor of dissatisfaction [130]. The lower level of self-reported success

and symptom resolution in those with lowest preoperative OKS despite large OKS gains is likely to reflect the fact that their overall final OKS remains low.

Rather than supporting an upper OKS threshold for TKR, the findings of this study would seem to suggest that a lower threshold might be more appropriate, to identify and ensure patients are not allowed to deteriorate preoperatively to point where postoperative outcome are compromised. Despite an apparently milder symptom burden, and lesser perceived benefits from surgery this study would suggest that those with a high preoperative OKS still achieve good clinical outcomes. The findings also highlight an important element as far as outcome measures are concerned, in that appreciation of the metric being utilised is extremely important [66]. Even though assessment of pain and function using PROMs is important, and provides an objective insight into the patient's improvement, they may not necessarily reflect the patient's own perception of success or failure. Regardless of the reduction in pain or gain in function achieved, a dissatisfied patient remains a poor outcome.

The strengths of this study include the large population available for analysis, drawn from the national UK arthroplasty PROMs database, therefore providing a reflection of current UK practice. Although data was drawn from national UK databases, it is recognised that accuracy of coding within such datasets may not be perfect [140]. Linkage between the HES and PROMS datasets in the UK, has also been identified as a potential problem, with approximately 10% of cases lost through linkage ineligibility [141]. Accounting for the impact of additional factors, the true recruitment rate for TKR PROMs data in the UK is estimated to be as low as 72.4%, with social deprivation and non-whites less likely to be recruited. Although these concerns should be acknowledged, the data quality of national datasets are still considered sufficiently robust for application in research studies [140].

Despite the large size of the overall sample it should be noted that at the extreme ends of the preoperative OKS spectrum, the number of cases available for analysis remained low making statistical inferences slightly more difficult. This was further influenced by additional loss of cases when adjusted regression models were generated, due to missing sociodemographic data in some cases. However, the consideration of these additional covariates was felt to be important, and this did not significantly change the overall results of the study. In the absence of any other large scale PROMs databases, this currently provides the best available option for analysis of this specific group of patients, and has identified important issues in relation to the preoperative OKS at both upper and lower ends of the spectrum.

In summary, this study has shown that despite seemingly limited gain in the OKS following TKR, those patients proceeding to TKR in the presence of a high preoperative OKS are more likely to report a successful outcome and achieve resolution of symptoms. Additionally, patients with a lower preoperative OKS appear to report inferior overall outcome despite large gains in OKS, with a suggestion that no further benefit is achieved once preoperative  $OKS < 16$ . This would suggest that caution is needed before considering introduction of OKS based thresholds as a criterion for TKR, and further focus should be given to preventing excessive delay in surgery.

## **Chapter Six**

### **6 General Discussion & Conclusions**

#### **6.1 Summary of the problem**

Knee OA represents a global problem with significant impact both on individuals and society as a whole. The management of knee OA encompasses a wide spectrum of treatment options, with TKR representing the main modality for end-stage symptomatic OA [18,23]. As the demand for TKR increases, on a background of limited healthcare resources, so have the pressures on healthcare commissioners to ensure that TKR remains both cost-effective and clinically effective. As part of this drive, concerns have been raised regarding the variation in access to surgery, as well as the appropriateness of TKR selection process [26,115,116]. Furthermore, the important role of PROMs in defining outcome following TKR has increasingly been recognised in recent years. In the UK this has led to a formal PROMs arthroplasty outcome programme, of which the OKS is considered the PROM tool of choice for TKR [41]. The OKS is a well-researched and widely accepted PROM in the assessment of TKR outcomes [41,71,73,112]. It is generally considered to be a good disease-specific measure of pain and function in relation to knee pathology, and is increasingly being validated and utilised for additional roles outside of its original remit of knee arthroplasty [71,142]. In a bid to further rationalise healthcare expenditure, and reduce perceived variations in access to TKR, there have recently been proposals in the UK to implement threshold criterion based on the preoperative OKS, to define those patients considered the most appropriate candidates for surgical intervention [94,95,97,98,103].

Despite the existence of extensive research in relation to knee OA, TKR outcomes, as well as PROMs, there remains a paucity of data in the current literature regarding the preoperative symptom burden of patients proceeding to TKR, and more importantly limited information as to how this relates to the preoperative OKS. Within this thesis a number of topics have been explored, providing important and highly relevant insight into the preoperative symptom burden of the knee replacement population, and the role of the preoperative OKS in defining this burden. It has highlighted a number of wider ranging issues surrounding the role of preoperative OKS as a measure of symptom burden, and as a threshold criterion for TKR.

## **6.2 Summary of Main Findings**

In Chapter 2 it was observed that there is great variability in the preoperative OKS amongst patients proceeding to knee replacement (UKR and TKR) in the UK, indicating substantial variation in individual symptom burden amongst patients. A normal distribution profile was observed for the preoperative OKS in all three centres. Of particular interest was the observation that a small cohort of patients appeared to be proceeding to TKR in the presence of a high preoperative OKS suggesting that the overall pain and functional disability levels may be also less severe than generally predicted. The normal distribution profile and wide variability in preoperative OKS was observed within all three centres and amongst individual surgeons however, there was some suggestion of a possible difference between individual centres, seemingly supporting previous concerns raised about variations in TKR practice.

This apparent variation was further investigated in Chapter 3, where the UK national arthroplasty database was explored in greater detail. This more expansive study again confirmed the general variability in preoperative OKS amongst the TKR population previously observed in Chapter 2. Analysis of the data across different populations showed

that although some variation does indeed exist between individual centre and surgeons, these were not major contributors to the general variability observed in preoperative OKS, suggesting that the selection process may not be entirely to blame. Despite the commonly perceived perception that variation in TKR surgery may be influenced by poor patient selection criteria, these findings suggested that wide variability in preoperative symptom burden is common practice across the UK.

Even though this variability in preoperative OKS was found to be a common observation throughout all the UK centres, the presence of patients pursuing TKR with a seemingly minimal symptom burden remained a concern, forming the basis of the study detailed in Chapter 4. In this study it was seen that patients proceeding to TKR with a higher preoperative OKS appeared to have a lower level of pain and greater function, suggesting the OKS was indeed an appropriate measure of preoperative symptom burden. Despite this though, it was also noted that the vast majority of patient with a high preoperative OKS still had evidence to suggest that the individual symptom burden remained significant, and that they remained clinically indicated. This was further supported by the review of the consultation process which found no obvious differences to suggest that inappropriate selection criteria may have accounted for these patients proceeding to TKR.

Finally, having identified wide variation in preoperative status throughout the TKR population, together with the existence of patients undergoing TKR with a seemingly lower symptom burden, a further analysis was undertaken to gain an understanding as to the impact of this apparently lower preoperative symptom burden on subsequent outcomes. It was observed that those patients proceeding to TKR with the highest preoperative OKS were most likely to report alleviation of their symptoms, report a successful outcome, and attain the

highest postoperative OKS, despite minimal change in OKS as a result of surgery. Contrasting with this group, those with the lowest preoperative OKS achieved large change in OKS following surgery, but despite this apparent benefit were least likely to perceive the result to be successful, and were more likely to report persistent knee symptoms. Of greater interest, once preoperative OKS dropped below 16 points, there appeared to be very little additional benefit achieved by these patients, suggesting that there may be an argument to support surgical intervention prior to this point.

It is evident from the previous literature, and additional work presented in this thesis that the preoperative symptom burden is an extremely complex issue to evaluate and utilise in the TKR population. It has been suggested that the preoperative OKS may be a useful surrogate marker of this symptom burden, however its application in this role may not be as simple as it first appears. A number of important features have been identified in relation to the role of the preoperative OKS as an indicator of symptom burden in the TKR population, and it is important to further consider how these may impact clinical care. A greater understanding of the preoperative OKS and the symptom profile of TKR patients is important for healthcare commissioners, health professionals and patients alike, and of major clinical relevance for both current and future arthroplasty practice.

## **6.3 Clinical relevance**

### **6.3.1 Understanding the normal ‘footprint’**

Inappropriate variation in practice is a major concern amongst healthcare commissioners, and although there are a number of potential causes for variation, patient selection is considered a

major factor [57,115,116]. Currently, there is only limited data published regarding the preoperative symptom burden of TKR in general, and the presence of variations in individual practice [28,58,63,120,121]. Similarly, there is very little data readily available to individual surgeons or centres, regarding the profile of the preoperative OKS, that can be easily accessed and interpreted in order to gain a meaningful insight into their individual practice and how this compares to other surgeons, centres or even the national data [107,108,137]. The Atlas of Variation report, based on UK data, examined variation in the Eq5D, which is a generic PROM tool, and reviewed data across PCTs only [58]. As a generic PROM, this is not specific to the knee symptom burden, and the limited data presented provides very little information regarding the true variation present. Even the most recent full National PROMs report, provides only basic information regarding preoperative OKS, precluding any useful comparisons for individual surgeons [143]. In light of the concerns regarding variation in practice it is essential that the normal ‘footprint’ is recognised and appreciated as there appears to be some mismatch between perceived practice and reality.

Most surgeons may imagine that the majority of TKR patients in their practice proceed to surgery in the presence of quite severe pain and disability, and thus considered to be clinically justified. In a survey of 234 Canadian orthopaedic surgeons, examining their perceived patient profiles, 97% of respondents thought that patients had moderate or severe pain, and that 99% required assistance walking or could not manage stairs [29]. A similar survey in the US exploring preoperative symptom profiles indicative of need for TKR, found that most surgeons believed that daily severe pain was required, and that patient should be able to walk less than 3 blocks or unable to manage stairs [27]. This thesis has however highlighted that the true symptom burden prior to TKR may not be as severe as anticipated. An appreciation of the wide variability in preoperative OKS may allow individual surgeons

to better appreciate how their individual practice relates to the general population, providing a reference to support or refute concerns about inappropriate variation in patient selection criteria. An appreciation of individual profiles and recognition of the variability in the preoperative OKS, may also enable surgeons to more easily compare and interpret their postoperative outcome data by better understanding their specific patient population.

Basic information on the normal ‘footprint’ may also be of value within the primary care setting. It is essential that GPs appreciate how the preoperative OKS reflects the preoperative symptom burden, and to realise that extreme pain and disability is not necessarily the ‘norm’. It is already recognised that GPs exhibit higher subjective thresholds than surgeons, delaying referrals for TKR until they perceive the symptoms to be severe enough, and that many lack confidence in making decision to refer on to secondary care [26,56]. Introduction of formal referral and threshold could in theory therefore help standardise the process. Many CCGs incorporate the OKS only as a component of their threshold criteria, with additional subjective factors also included. However, a lack of confidence with subjective indicators amongst GPs may lead to over-reliance on the preoperative OKS as an indicator of TKR eligibility. There is a risk though that unless individuals fully appreciate how the preoperative OKS relates to normal arthroplasty practice that it further compounds delays in referral to secondary care.

As the potential applications of the OKS continue to expand [142], it becomes even more important that we fully understand how the population profile of this PROM tool may differ across different populations and disease processes. This is especially important where the OKS is utilised to monitor response to non-operative treatment of knee OA. Although, not necessarily utilised for any formal threshold criterion in this role, again unless the individuals

interpreting the OKS in this scenario appreciates the normal TKR profile, they cannot fully understand how these measurements in these alternative roles relate to the TKR population, which may be essential when making further decision about progression up the treatment ladder.

### **6.3.2 Guiding Treatment Decisions for Patients**

In addition to guiding GPs and surgeons, this information may act as a useful preoperative symptom map for patients alike. Shared decision-making has become a key concept within the NHS [144]. In turn patient decision aids have come to play an increasing role in the decision processes around management of knee OA, ensuring that patients are fully informed about their treatment options [145]. A common problem for patients considering TKR is that individual perception of symptom burden is very subjective, and there is very little information available to act as a comparative benchmark to aid their decision making process [50,146,147]. As a result, they may mistakenly consider their individual symptom burden to be less (or more) severe relative to the general TKR population, and in turn defer consultations and surgery [50,148]. Although there are numerous considerations when proceeding to TKR, this is not currently incorporated into TKR decision aids [149]. An ability to place an individual on the preoperative OKS scale may enable them to gain a better appreciation how their preoperative symptom burden compares with other individuals currently undergoing TKR.

Dissatisfaction following TKR remains a major concern, especially when many of these patients are perceived to have achieved a good clinical outcome [150]. It is commonly reported that those with the worst preoperative status have the greatest potential gains following TKR. It is important though to realise that improvement in status on one metric

may not necessarily reflect a good outcome from the patient's perspective. The ability to tailor the preoperative consultation according to the individual's preoperative OKS, and provide a more realistic indication of likely outcome with regards pain, function and overall success rate might be of greater value than current generic predictive models simply quoting a single cut-point associated with success. Scott et al found that those with the lowest preoperative OKS generally had highest expectations with regards pain relief and function [151]. However as identified in this thesis, those at the lowest end of the preoperative OKS scale are likely to still have considerable degree of pain and disability. This thesis highlighted that a good outcome was achievable, even amongst those patients with the worst preoperative status. However, amongst this group of patients there was greater variability in postoperative OKS, implying that outcome is less predictable. It is therefore important that these patients are specifically educated about the likely outcome, and to realise that these gains may not take them back to a 'normal' state. Understanding the preoperative OKS profile and its relationship with outcome may be extremely useful. The ability to identify these patients and ensure they are fully informed may negate the potential for unrealistic expectations about likely postoperative outcome. In turn this may help reduce some of the dissatisfaction that is often seen following TKR.

### **6.3.3 Preoperative Symptom Burden and TKR Thresholds**

A key prompt leading to the development of the work presented in this thesis were proposals by some CCGs to introduce OKS based thresholds to define eligibility for TKR [94-98]. The OKS is widely accepted and fully validated as a PROM tool to assess postoperative TKR outcomes, as well as increasingly playing a role in other clinical scenarios [41,71,142]. However, at present there are no validation studies to support the OKS as an indicator of surgical need and/or eligibility, and no evidence to guide commissioners as to the most

appropriate threshold cut-off. Current threshold proposals do not provide any rationalisation for the selected cut-points, and consequently there is wide variation with OKS threshold values ranging from as low as 20 all the way up to 34. Some outcome studies have previously attempted to partition the OKS into 4 groups to better define the relative success postoperatively; OKS>41=excellent, OKS 34-41= good, OKS 27-33=fair, OKS<27 = poor [137,152]. The ‘good’ and ‘fair’ groups seem to match some of the proposed threshold cut-points, although it is unclear whether they actually formed the basis for the proposed thresholds scores or not. Even these defined cut-points though remain entirely arbitrary in nature, and furthermore were originally described in relation to the OHS [153]. As was evident in Chapter 3 the impact of such arbitrary thresholds is potentially immense, especially if the preoperative OKS profile of the local population is not known. Between 4-37% of patients currently undergoing TKR would be excluded by current proposed OKS thresholds, creating huge potential inequities in access rather than reducing them.

Most current threshold or priority criteria generally assume that patients expressing the greatest level of pain and disability represent individuals with the greatest clinical need [84,154,155]. However, there are numerous published reports showing that disease severity does not necessarily correlate with an individual’s perceived need for surgery [45,50,92,156,157]. Hawker et al surveyed a group of patients with OA identified to have potential need for arthroplasty [156]. They utilised the WOMAC score as an indicator of disease severity, and found that it was not related to patients’ willingness to undergo surgical intervention. Amongst those with severe OA no more than 15% were ‘definitely willing’ to consider knee arthroplasty. Approximately 50% were ‘probably’ or ‘definitely’ unwilling to consider an arthroplasty procedure. McHugh et al examined patients referred to secondary care for consideration of TKR in the UK, and found that only 33% of those referred

subsequently proceeded to a TKR [158]. It was noted though that those proceeding to knee replacement had worse pain and function as measured by the OKS, pain scores and the WOMAC. Toye et al explored the potential need for TKR amongst patients deemed to have a lower than average disease burden. They identified that symptoms alone were not the only factors important in the decision process, and that ‘personal meanings’ were important to some sub-groups [147].

An additional potential problem surrounding the use of the OKS to define surgical eligibility is that it was developed at a time when TKR indications were considerably more restricted. As a result, only simple ADLs were considered important in defining the symptom burden. As the indications and success of knee arthroplasty have changed, so have the functional expectations. Patients are increasingly seeing TKR as a route to restore quality of life and leisure activities, rather than simply being an intervention to eliminate pain [159]. This has already been recognised as a potential weakness of the OKS as an outcome measure, with development of a modified high activity version to further assess the more active patients [160]. The OKS in its basic form may therefore fail to take into consideration the additional impact of knee OA on quality of life, occupations, and elements such as family and social life.

There is also the question of whether the preoperative symptom burden is in itself a valid indicator of priority or eligibility for TKR, which may not necessarily correspond with the individual’s perceived need. A comparison of physician derived priority scores versus the WOMAC amongst a cohort of TKR patients in New Zealand, demonstrated poor correlation between the two scores [92]. It was concluded that the priority score failed to accurately prioritise cases. However, an alternative conclusion not considered, might be that the

WOMAC did not accurately reflect the true surgical need. The OARSI-OMERACT group explored the correlation between subjective opinion of surgeons and a number of different PROM tools, measuring pain and function, in 1130 TKR patients [45]. Although they found that those recommended for TKR generally had higher pain and disability levels, this alone did not fully discriminate between those where TKR was or was not indicated. More importantly, in a separate sub-analysis it was found that they could not identify a specific cut-point for the PROMs analysed, that reflected eligibility for TKR [31]. The results presented in this thesis seem to suggest that preoperative pain and disability alone do not appear to be the only indications for TKR, evident by the wide variability in the preoperative OKS observed throughout the UK. It further supports these earlier studies, casting doubt on the ability of standard knee PROMs to accurately determine the relative need for TKR amongst patients, based on pain and disability alone. Over reliance on the preoperative OKS as a sole indicator of surgical need may exclude some patients where alternative indications are not fully encapsulated, whereas in other cases might encourage healthcare professionals to pursue the arthroplasty route, without full consideration of the individual's preferences.

One justification for implementation of PROM-based thresholds is that subjective criteria are open to interpretation, and in turn potentially allow for inappropriate surgery and potentially inferior outcomes. Riddle et al examined postoperative outcomes in 167 patients that had undergone TKR, and determined that 34.3% could be classed as inappropriate [62]. The inappropriate group exhibited better preoperative KOOS and WOMAC function scores, suggesting a lower symptom burden. Furthermore, they concluded that this group of patients experienced minimal benefit from surgery, and that postoperative improvements were delayed. Qunitana et al found in a cohort of 792 TKR patients, that 12.4% were deemed inappropriate and again demonstrated better preoperative WOMAC and SF-36 scores. These

patients did achieve an improvement in both WOMAC and SF-36 scores postoperatively, but these were significantly lower compared to the appropriate cases [61]. Of greater interest though was their observation in relation to patient perception of satisfaction and general improvement. Inappropriate TKR cases reported lower levels of satisfaction with the intervention, as well as a lower global satisfaction rating. These are strong arguments in support of explicit threshold criteria for TKR, and suggest that preoperative symptom burden as measured by PROM tools may be important. Based on these findings one might therefore suspect that the cohort of patients proceeding to TKR in the presence of a high OKS may also reflect inappropriate surgical intervention. Formal assessment of appropriateness could not be undertaken on the study populations described in this thesis. However, the results presented in Chapter 4 contrast with these earlier studies, suggesting that most cases undergoing TKR remained clinically justified despite a high OKS, and that all had followed an appropriate preoperative consultation process. In Chapter 5 it was noted that the largest gains in postoperative OKS was observed amongst those with the worst preoperative OKS. However, those patients with the highest preoperative OKS achieved the highest level of success and symptom resolution postoperatively based on the secondary anchoring questions, despite achieving very little change in postoperative OKS. This would suggest that the preoperative OKS does not adequately allow us to determine appropriateness, and is a further important consideration relating to its proposed role as a threshold tool.

Although many may see threshold scores primarily as a clinical decision aid, ongoing financial restrictions and the need to reduce health care expenditure and improve cost effectiveness across the health service forms another argument in their favour [85]. Health economic analyses seem to suggest though that those with a high OKS may still be appropriate for TKR casting further doubt on the role of preoperative OKS as a financial

threshold criterion. Dakin et al demonstrated in their analysis of 2,131 patients, based on conservative estimates, that TKR was still justified on economic grounds with a preoperative OKS up to 40, and when considering less conservative analyses, TKR remained cost effective for all except approximately 100 patients per year [101]. A simulated threshold of  $OKS < 27$  in their study sample would deny cost effective surgery to over 10,000 patients. Jenkins et al demonstrated that TKR was more cost-effective in those with the lowest preoperative OKS in terms of QALY gained [161]. However, TKR remained generally cost-effective, even when additional factors such as future revision costs associated with younger and more active patients were considered, and fell well below NICE cost effectiveness thresholds. They also identified that if OKS based thresholds were utilised as a financial criterion, there was a need for an 8-point difference between THR and TKR cut-points. Although this thesis is focussed on TKR thresholds only, all the current CCG threshold proposals include identical cut-off values for both the OKS and the OHS, and highlights yet another potential issue with introduction of arbitrary values in the absence of formal validation studies.

#### **6.3.4 The Impact of OKS Thresholds on Postoperative Outcomes**

A key consideration in measuring and utilising the preoperative symptom burden as an indicator of eligibility, is the subsequent impact it may have on the postoperative outcomes. The general relationship between pre and postoperative outcome scores have been well documented in the literature [121,162], but data relating specifically to the OKS is more limited [130,135,137,138]. Previous studies have focussed on the general trends, and have not specifically focussed on those patients with the greatest or lowest preoperative symptom burden. Although these studies highlight the importance of considering preoperative status, and confirm the general trend, they do not help the individual surgeon to determine the likelihood of success according to preoperative status. As has already been discussed, the

current data provides an indication of likely outcome based on the individual preoperative OKS, and thus is of direct relevance to surgeons and patients.

Some studies suggest those with worst preoperative OKS achieve the greatest gains from surgery and that a larger postoperative change in OKS is associated with postoperative satisfaction [130,138]. Care is needed in interpreting these findings particularly when only a single metric is utilised. The results presented in this thesis showed that despite generally attaining large changes in OKS following TKR, those with the lowest preoperative scores had a poorer self-perceived outcome. Conversely, those with a seemingly lower preoperative symptom burden, proceeding to surgery with a high OKS actually had the highest reported success and most likely to report symptom resolution. The key feature identified in this thesis, not previously reported, was the non-linear relationship between pre and postop OKS. Patients with a preoperative  $OKS < 16$  did not appear to achieve any further benefit as the OKS continued to decrease. Hamilton et al prospectively investigated the predictors of satisfaction following TKR in 4709 patients at a single centre [163]. Five key factors were identified, one of which included the preoperative OKS. Detailed breakdown was not provided, but a mean preoperative OKS of 16, which is similar to the cut point identified here, was found to be associated with the dissatisfied group. This is an extremely important finding, especially in light of proposed introduction of OKS based thresholds. Introduction of arbitrary OKS-based thresholds, would specifically exclude a large cohort of patients who are most likely to report a successful outcome, most likely to report resolution of knee symptoms, and achieve the highest postoperative OKS. Delaying referrals until a specific threshold is reached also risks patients subsequently overshooting, and ultimately undergo surgery with a much lower preoperative OKS. In turn patient could be destined to a poorer outcome, the complete opposite of what most healthcare professionals and commissioners strive for.

The indirect impact of OKS based thresholds on postoperative outcomes should also be considered. Introduction of priority scores and threshold criteria in other healthcare systems has resulted in mixed responses [91,92,164,165]. One particular phenomenon that has been observed secondary to priority scores is that of ‘gaming the system’, whereby patients and/or doctors artificially manipulate the threshold scores in order to bypass the restrictions on access to surgery [166,167]. Such behaviour further promotes inequities in access, as not all patients or surgeons may participate in this activity [91,164,165]. In the context of the preoperative OKS as a threshold criterion, a much greater concern needs consideration regarding this potential for ‘gaming’ the system.

The OKS is the main PROM tool currently utilised in the UK to monitor TKR outcomes and additionally provides valuable data for research purposes [143]. It is therefore critical that the measured outcomes are an accurate representation of the results of surgery. If the preoperative OKS is artificially manipulated in order to gain access to surgery, this will in turn directly impact on our interpretation of the subsequent postoperative outcomes, particularly when calculating the postoperative change in OKS, which may in turn be overestimated. It was observed in the previous chapter that patients with the highest preoperative OKS were most likely to report a successful outcome. If these patients had artificially manipulated their preoperative OKS to bypass proposed thresholds, we might incorrectly conclude that lower scores were associated with the best outcomes. The potential impact of such manipulation of the preoperative OKS is immense, compromising outcome interpretation as well as any research studies utilising this data, and might seriously invalidate the national arthroplasty database. Caution is therefore required before introducing OKS-based thresholds and full consideration needs to be given to both direct and indirect effects.

The scope of this thesis was not to formally evaluate the role of the preoperative OKS as a threshold criterion. However, a number of significant issues have been identified which cast serious doubt as to the validity of such an application for this PROM tool. The general findings presented would suggest that in the currently proposed format the preoperative OKS based thresholds appear to be inappropriate as a tool to define eligibility for TKR. If such a role is to be implemented, then formal validation studies are urgently required. There may however be a role as a priority indicator at the lower end of the preoperative OKS spectrum. Regardless of thresholds scores, patient decision aids, and various other measures to try and optimise the selection process for TKR, there will always be cases where surgery does not occur at the optimum time. Identification of these cases at risk of inferior outcome may be much more important than restricting access to those deemed to have a lesser symptom burden.

## **6.4 Future directions**

The work presented in this thesis has provided an interesting snapshot of current arthroplasty practice in the UK, and the role of the preoperative OKS as a marker of symptom burden. Although it has expanded on the current literature pertaining to this topic, several supplementary questions and issues have been identified that are worthy of further investigation.

Appropriateness of any surgical decision is a difficult concept to measure, and due to the limitations of retrospective analyses this could not be fully evaluated in the work presented here. Limited analyses suggested that in the main, cases did appear to still be clinically indicated despite a high preoperative OKS, but further prospective studies should be undertaken to formally evaluate this. It is essential to understand how clinical decision

making relates to the preoperative OKS, and to determine whether those with a high preoperative OKS truly reflect appropriate case selection.

In the current economic climate ongoing rationalisation of health services is inevitably here to stay, and likely to become a major factor in healthcare commissioning policies. Consequently, ongoing changes to practice including arthroplasty surgery, whether secondary to formal threshold criteria or not seems unavoidable. Further comparative review of more recent data from the national database, may yield further insights into the potential impact of stricter referral criteria. It would be interesting to formally evaluate whether those CCGs that have applied stricter referral criteria demonstrate any shift in the distribution profile of the preoperative OKS to reflect these changes. Having highlighted the potential impact of preoperative OKS on subsequent outcome it would also be important to specifically focus on the postoperative outcomes within these same CCGs to determine whether there has been any deterioration in postoperative outcomes as a result of stricter referral criteria. This would provide evidence to further support or refute some of the conclusion presented in this thesis

This work represents the first detailed analysis of the preoperative OKS profile in the TKR population, however the OKS is now increasingly being accepted as a PROM tool for additional roles outside of this original remit [74-77,142]. As the preoperative OKS continues to contribute to the monitoring of non-operative management of knee OA, it is important that we extend such analyses to this setting as well in order to fully understand how the OKS distribution profiles of patients at different stages of treatment, or indeed with alternative pathologies may differ, and the potential impact of this on future management options.

The results presented here suggest that the OKS alone does not appear to fully encapsulate the indications for pursuing TKR, particularly amongst those with the highest preoperative OKS. The importance of higher level function, particularly amongst younger patients, has already been recognised, leading to the development of the modified high activity OKS [160]. It may be useful to further explore how this modified OKS system compares with the standard OKS, and whether it provides better feedback in relation to the indications for surgery, and thus may be a more useful tool at the higher end of the preoperative OKS spectrum.

## **6.5 Final Conclusions**

A number of significant issues and findings have been discussed within this thesis in relation to the preoperative OKS and its role in determining the preoperative symptom burden prior to TKR. The normal distribution profile and general variability in preoperative OKS evident throughout the UK suggests that the contribution of pain and disability as an indication for TKR may be less important than has been previously considered, and highlights that a significant cohort of patients proceed to TKR in the presence of a seemingly minimal preoperative symptom burden. Despite a high preoperative OKS these patients still seem to be clinically indicated for TKR. Furthermore, these patients subsequently achieved the best overall postoperative outcomes, whereas those with the greatest preoperative symptom burden appear to be compromised postoperatively once preoperative OKS < 16.

Although the preoperative OKS may certainly have a useful and important role in defining some of the symptom burden prior to TKR, it would appear that it does not fully encapsulate the clinical indications for TKR. There is a risk that exclusion of patients with a lower symptom burden, using OKS-based thresholds, and limiting TKR to those with the worst

preoperative status, could actually lead to inferior outcomes in the longer term. There may be a role for utilising the preoperative OKS as a priority indicator, particularly to identify those at the lowest end of the spectrum, but its role as an indicator of preoperative symptom burden and potential eligibility for TKR should be used with extreme caution.

## Appendix

### A. Oxford Knee Score

#### PROBLEMS WITH YOUR KNEE

During the past 4 weeks..

✓ tick one box  
for every question

1	<p><i>During the past 4 weeks.....</i></p> <p>How would you describe the pain you <u>usually</u> have from your knee?</p> <p>None                  Very mild                  Mild                  Moderate                  Severe</p> <p><input type="checkbox"/>                  <input type="checkbox"/>                  <input type="checkbox"/>                  <input type="checkbox"/>                  <input type="checkbox"/></p>
2	<p><i>During the past 4 weeks.....</i></p> <p>Have you had any trouble with washing and drying yourself (all over) <u>because of your knee</u>?</p> <p>No trouble at all                  Very little trouble                  Moderate trouble                  Extreme difficulty                  Impossible to do</p> <p><input type="checkbox"/>                  <input type="checkbox"/>                  <input type="checkbox"/>                  <input type="checkbox"/>                  <input type="checkbox"/></p>
3	<p><i>During the past 4 weeks.....</i></p> <p>Have you had any trouble getting in and out of a car or using public transport <u>because of your knee</u>? (whichever you would tend to use)</p> <p>No trouble at all                  Very little trouble                  Moderate trouble                  Extreme difficulty                  Impossible to do</p> <p><input type="checkbox"/>                  <input type="checkbox"/>                  <input type="checkbox"/>                  <input type="checkbox"/>                  <input type="checkbox"/></p>
4	<p><i>During the past 4 weeks.....</i></p> <p>For how long have you been able to walk before <u>pain from your knee</u> becomes <b>severe</b>? (<i>with or without a stick</i>)</p> <p>No pain/ More than 30 minutes                  16 to 30 minutes                  5 to 15 minutes                  Around the house <u>only</u>                  Not at all - pain severe when walking</p> <p><input type="checkbox"/>                  <input type="checkbox"/>                  <input type="checkbox"/>                  <input type="checkbox"/>                  <input type="checkbox"/></p>
5	<p><i>During the past 4 weeks.....</i></p> <p>After a meal (sat at a table), how painful has it been for you to stand up from a chair <u>because of your knee</u>?</p> <p>Not at all painful                  Slightly painful                  Moderately painful                  Very painful                  Unbearable</p> <p><input type="checkbox"/>                  <input type="checkbox"/>                  <input type="checkbox"/>                  <input type="checkbox"/>                  <input type="checkbox"/></p>
6	<p><i>During the past 4 weeks.....</i></p> <p>Have you been limping when walking, <u>because of your knee</u>?</p> <p>Rarely/ never                  Sometimes, or just at first                  Often, not just at first                  Most of the time                  All of the time</p> <p><input type="checkbox"/>                  <input type="checkbox"/>                  <input type="checkbox"/>                  <input type="checkbox"/>                  <input type="checkbox"/></p>

**During the past 4 weeks...** ✓tick one box  
for every question

<b>7</b>	<p><i>During the past 4 weeks.....</i></p> <p><b>Could you kneel down and get up again afterwards?</b></p> <p>Yes, Easily <input type="checkbox"/>      With little difficulty <input type="checkbox"/>      With moderate difficulty <input type="checkbox"/>      With extreme difficulty <input type="checkbox"/>      No, Impossible <input type="checkbox"/></p>
<b>8</b>	<p><i>During the past 4 weeks.....</i></p> <p><b>Have you been troubled by <u>pain from your knee</u> in bed at night?</b></p> <p>No nights <input type="checkbox"/>      Only 1 or 2 nights <input type="checkbox"/>      Some nights <input type="checkbox"/>      Most nights <input type="checkbox"/>      Every night <input type="checkbox"/></p>
<b>9</b>	<p><i>During the past 4 weeks.....</i></p> <p><b>How much has <u>pain from your knee</u> interfered with your usual work (including housework)?</b></p> <p>Not at all <input type="checkbox"/>      A little bit <input type="checkbox"/>      Moderately <input type="checkbox"/>      Greatly <input type="checkbox"/>      Totally <input type="checkbox"/></p>
<b>10</b>	<p><i>During the past 4 weeks.....</i></p> <p><b>Have you felt that your knee might suddenly 'give way' or let you down?</b></p> <p>Rarely/never <input type="checkbox"/>      Sometimes, or just at first <input type="checkbox"/>      Often, not just at first <input type="checkbox"/>      Most of the time <input type="checkbox"/>      All of the time <input type="checkbox"/></p>
<b>11</b>	<p><i>During the past 4 weeks.....</i></p> <p><b>Could you do the household shopping <u>on your own</u>?</b></p> <p>Yes, Easily <input type="checkbox"/>      With little difficulty <input type="checkbox"/>      With moderate difficulty <input type="checkbox"/>      With extreme difficulty <input type="checkbox"/>      No, Impossible <input type="checkbox"/></p>
<b>12</b>	<p><i>During the past 4 weeks.....</i></p> <p><b>Could you walk down one flight of stairs?</b></p> <p>Yes, Easily <input type="checkbox"/>      With little difficulty <input type="checkbox"/>      With moderate difficulty <input type="checkbox"/>      With extreme difficulty <input type="checkbox"/>      No, Impossible <input type="checkbox"/></p>

## B. List of Abbreviations

ADL	Activities of Daily Living
CCG	Clinical Commissioning Group
CI	Confidence Intervals
CV	Coefficient of Variation
DALY	Disability Adjusted Life Years
GP	General Practitioner
HES	Hospital Episode Statistics
ICC	Intraclass Correlation Coefficient
IMD	Index of Multiple Deprivation
IQR	Interquartile Range
MRI	Magnetic Resonance Imaging
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NSAID	Non-steroidal Anti Inflammatory Drugs
OA	Osteoarthritis
OKS	Oxford Knee Score
OHS	Oxford Hip Score
OR	Odds Ratio
PCT	Primary Care Trust
PROM	Patient Reported Outcome Measure
QOL	Quality of Life
RLL	Restricted Log Likelihood
SD	Standard Deviation
THR	Total Hip Replacement
TKR	Total Knee Replacement
UK	United Kingdom
UKR	Unicompartmental Knee Replacement
US	United States
VPC	Variance Partition Coefficient
WHO	World Health Organization
WOMAC	Western Ontario and McMaster Universities Osteoarthritis Index
YLD	Years Lived with Disability

## Bibliography

1. **Arden N., Nevitt M. C.** Osteoarthritis: epidemiology. *Best Pract Res Clin Rheumatol.* 2006;20(1):3-25.
2. **Arthritis Research UK.** Osteoarthritis in General Practice 2013.
3. **Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al.** Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet.* 2012;380(9859):2163-96.
4. **Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al.** Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet.* 2012;380(9859):2197-223.
5. **Murray CJ, Richards MA, Newton JN, Fenton KA, Anderson HR, Atkinson C, et al.** UK health performance: findings of the Global Burden of Disease Study 2010. *Lancet.* 2013;381(9871):997-1020.
6. **Litwic A, Edwards MH, Dennison EM, Cooper C.** Epidemiology and burden of osteoarthritis. *Br Med Bull.* 2013;105:185-99.
7. **Cross M, Smith E, Hoy D, Nolte S, Ackerman I, Fransen M, et al.** The global burden of hip and knee osteoarthritis: estimates from the global burden of disease 2010 study. *Ann Rheum Dis.* 2014;73(7):1323-30.
8. **Wolf AD, Pfleger B.** Policy and Practice Burden of major musculoskeletal conditions. *B World Health Organ.* 2003;81(03).
9. **Sowers MR, Karvonen-Gutierrez CA.** The evolving role of obesity in knee osteoarthritis. *Curr Opin Rheumatol.* 2010;22(5):533-7.
10. **Losina E, Weinstein AM, Reichmann WM, Burbine SA, Solomon DH, Daigle ME, et al.** Lifetime risk and age at diagnosis of symptomatic knee osteoarthritis in the US. *Arthritis Care Res.* 2013;65(5):703-11.
11. **Murphy L, Schwartz TA, Helmick CG, Renner JB, Tudor G, Koch G, et al.** Lifetime risk of symptomatic knee osteoarthritis. *Arthritis Rheum.* 2008;59(9):1207-13.
12. **Ruiz Jr D, Koenig L, Dall TM, Gallo P, Narzikul A, Parvizi J, et al.** The direct and indirect costs to society of treatment for end-stage knee osteoarthritis. *J Bone Joint Surg Am.* 2013;95(16):1473-80.

13. **Chen A, Gupte C, Akhtar K, Smith P, Cobb J.** The Global Economic Cost of Osteoarthritis: How the UK Compares. *Arthritis*. 2012;2012:698709.
14. **Gupta S, Hawker GA, Laporte A, Croxford R, Coyte PC.** The economic burden of disabling hip and knee osteoarthritis (OA) from the perspective of individuals living with this condition. *Rheumatology*. 2005;44(12):1531-7.
15. **Leardini G, Salaffi F, Caporali R, Canesi B, Rovati L, Montanelli R, et al.** Direct and indirect costs of osteoarthritis of the knee. *Clin Exp Rheumatol*. 2004;22(6):699-706.
16. **Xie F, Thumboo J, Fong KY, Lo NN, Yeo SJ, Yang KY, et al.** A study on indirect and intangible costs for patients with knee osteoarthritis in Singapore. *Value Health*. 2008;11 Suppl 1:S84-90.
17. **Bennell KL, Hunter DJ, Hinman RS.** Management of osteoarthritis of the knee. *BMJ*. 2012;345:e4934.
18. **National Institute for Health and Care Excellence.** Osteoarthritis: care and management in adults. 2014. [www.nice.org.uk/guidance/cg177](http://www.nice.org.uk/guidance/cg177) (date last accessed March 2016).
19. **National Institute for Health and Care Excellence.** Arthroscopic knee washout, with or without debridement, for the treatment of osteoarthritis. 2007. [www.nice.org.uk/guidance/ipg230](http://www.nice.org.uk/guidance/ipg230) (date last accessed March 2016).
20. **Swedish Knee Arthroplasty Register.** Annual Report 2013. 2013. [www.shpr.se/en/publications/documentsreports.aspx](http://www.shpr.se/en/publications/documentsreports.aspx) (date last accessed March 2016).
21. **Brouwer RW, Huizinga MR, Duivenvoorden T, van Raaij TM, Verhagen AP, Bierma-Zeinstra SM, et al.** Osteotomy for treating knee osteoarthritis. *Cochrane Database Syst Rev*. 2014;12:CD004019.
22. **National Joint Registry for England & Wales.** 9th Annual Report. 2012.
23. **NIH Consensus Panel.** NIH Consensus Statement on total knee replacement December 8-10, 2003. *J Bone Joint Surg Am*. 2004;86-A(6):1328-35.
24. **Culliford DJ, Maskell J, Beard DJ, Murray DW, Price AJ, Arden NK.** Temporal trends in hip and knee replacement in the United Kingdom: 1991 to 2006. *J Bone Joint Surg Br*. 2010;92(1):130-5.
25. **Kurtz S, Ong K, Lau E, Mowat F, Halpern M.** Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. *J Bone Joint Surg Am*. 2007;89(4):780-5.

26. **Dieppe P, Basler HD, Chard J, Croft P, Dixon J, Hurley M, et al.** Knee replacement surgery for osteoarthritis: effectiveness, practice variations, indications and possible determinants of utilization. *Rheumatology* 1999;38(1):73-83.
27. **Mancuso C, Ranawat CS, Esdaile JM, Johanson N, Charlson ME.** Indications for total hip and total knee arthroplasties. Results of orthopaedic surveys. *J Arthroplasty*. 1996;11(1):34-46.
28. **Löfvendahl S, Bizjajeva S, Ranstam J, Lidgren L.** Indications for hip and knee replacement in Sweden. *J Eval Clin Pract* 2010;(1):1-10.
29. **Wright JG, Coyte P, Hawker G, Bombardier C, Cooke D, Heck D, et al.** Variation in orthopedic surgeons' perceptions of the indications for and outcomes of knee replacement. *Can Med Assoc J* 1995;152(5):687-97.
30. **Herickhoff PK, Callaghan JJ, Johnston R, Marsh JL, Clark CR, Noiseux N.** Primary hip and knee replacement: "are we all operating on the same patients, even at the same institution?". *Iowa Orthop J*. 2010;30:109-14.
31. **Gossec L, Paternotte S, Bingham CO, Clegg DO, Coste P, Conaghan PG, et al.** OARSI/OMERACT Initiative to Define States of Severity and Indication for Joint Replacement in Hip and Knee Osteoarthritis. An OMERACT 10 Special Interest Group. *J Rheumatol*. [10.3899/jrheum.110403]. 2011;38(8):1765-9.
32. **Dieppe P, Lim K, Lohmander S.** Who should have knee joint replacement surgery for osteoarthritis? *Int J Rheumat Dis*. 2011;14(2):175-80.
33. **Bijlsma JW, Berenbaum F, Lafeber FP.** Osteoarthritis: an update with relevance for clinical practice. *Lancet*. 2011;377(9783):2115-26.
34. **Ahlback S.** Osteoarthrosis of the knee. A radiographic investigation. *Acta Radiol Diagn*. 1968:Suppl 277:7-72.
35. **Kellgren JH, Lawrence JS.** Radiological Assessment of Osteo-Arthrosis. *Ann Rheum Dis*. 1957;16(4):494-502.
36. **Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, et al.** Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. *Arthritis Rheum*. 1986;29(8):1039-49.
37. **Zhang W, Doherty M, Peat G, Bierma-Zeinstra MA, Arden NK, Bresnihan B, et al.** EULAR evidence-based recommendations for the diagnosis of knee osteoarthritis. *Ann Rheum Dis*. 2010;69(3):483-9.

38. **Outerbridge RE.** The etiology of chondromalacia patellae. *J Bone Joint Surg Br.* 1961;43-B:752-7.
39. **Peterfy CG, Guermazi A, Zaim S, Tirman PF, Miaux Y, White D, et al.** Whole-Organ Magnetic Resonance Imaging Score (WORMS) of the knee in osteoarthritis. *Osteoarthr Cartilage.* 2004;12(3):177-90.
40. **Hunter DJ, Lo GH, Gale D, Grainger AJ, Guermazi A, Conaghan PG.** The reliability of a new scoring system for knee osteoarthritis MRI and the validity of bone marrow lesion assessment: BLOKS (Boston Leeds Osteoarthritis Knee Score). *Ann Rheum Dis.* 2008;67(2):206-11.
41. **Dawson J, Fitzpatrick R, Murray D, Carr A.** Questionnaire on the perceptions of patients about total knee replacement. *J Bone Joint Surg Br.* 1998;80(1):63-9.
42. **Collins NJ, Misra D, Felson DT, Crossley KM, Roos EM.** Measures of knee function: International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form, Knee Injury and Osteoarthritis Outcome Score (KOOS), Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form (KOOS-PS), Knee Outcome Survey Activities of Daily Living Scale (KOS-ADL), Lysholm Knee Scoring Scale, Oxford Knee Score (OKS), Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), Activity Rating Scale (ARS), and Tegner Activity Score (TAS). *Arthritis Care Res (Hoboken).* 2011;63 Suppl 11:S208-28.
43. **Peat G, Thomas E, Duncan R, Wood L, Hay E, Croft P.** Clinical classification criteria for knee osteoarthritis: performance in the general population and primary care. *Ann Rheum Dis.* 2006;65(10):1363-7.
44. **Schiphof D, de Klerk BM, Koes BW, Bierma-Zeinstra S.** Good reliability, questionable validity of 25 different classification criteria of knee osteoarthritis: a systematic appraisal. *J Clin Epidemiol.* 2008;61(12):1205-15.
45. **Gossec L, Paternotte S, Maillefert JF, Combescure C, Conaghan PG, Davis AM, et al.** The role of pain and functional impairment in the decision to recommend total joint replacement in hip and knee osteoarthritis: an international cross-sectional study of 1909 patients. Report of the OARSI-OMERACT Task Force on total joint replacement. *Osteoarthritis Cartilage.* 2011;19(2):147-54.
46. **Riddle DL, Jiranek WA, Hull JR.** Validity and reliability of radiographic knee osteoarthritis measures by arthroplasty surgeons. *Orthopedics.* 2013;36(1):e25-32.

47. **Bedson J, Croft PR.** The discordance between clinical and radiographic knee osteoarthritis: a systematic search and summary of the literature. *BMC Musculoskelet Disord.* 2008;9:116.
48. **Singh JA, Lewallen DG.** Time trends in the characteristics of patients undergoing primary total knee arthroplasty. *Arthritis Care Res.* 2014;66(6):897-906.
49. **Holland R.** Population needs assessment and knee replacement surgery. *Rheumatology.* 2003;42(4):503-6.
50. **Sansom A, Donovan J, Sanders C, Dieppe P, Horwood J, Learmonth I, et al.** Routes to total joint replacement surgery: patients' and clinicians' perceptions of need. *Arthritis Care Res.* 2010;62(9):1252-7.
51. **Peat G, McCarney R, Croft P.** Knee pain and osteoarthritis in older adults: a review of community burden and current use of primary health care. *Ann Rheum Dis.* 2001;60(2):91-7.
52. **Smith TO, Purdy R, Lister S, Salter C, Flectcroft R, Conaghan PG.** Attitudes of people with osteoarthritis towards their conservative management: a systematic review and meta-ethnography. *Rheumatol Int.* 2014;34(3):299-313.
53. **Frankel L, Sanmartin C, Conner-Spady B, Marshall DA, Freeman-Collins L, Wall A, et al.** Osteoarthritis patients' perceptions of "appropriateness" for total joint replacement surgery. *Osteoarthritis Cartilage.* 2012;20(9):967-73.
54. **Dieppe PA, Cushnaghan J, Shepstone L.** The Bristol 'OA500' study: progression of osteoarthritis (OA) over 3 years and the relationship between clinical and radiographic changes at the knee joint. *Osteoarthritis Cartilage.* 1997;5(2):87-97.
55. **Mandl LA.** Determining who should be referred for total hip and knee replacements. *Nat Rev Rheumatol.* 2013;9(6):351-7.
56. **Dieppe P, Dixon D, Horwood J, Pollard B, Johnston M.** MOBILE and the provision of total joint replacement. *J Health Serv Res Policy.* 2008;13 (Suppl 3):47-56.
57. **Judge A, Welton NJ, Sandhu J, Ben-Shlomo Y.** Equity in access to total joint replacement of the hip and knee in England: cross sectional study. *BMJ.* 2010;341:c4092.
58. **Department of Health.** The NHS Atlas of Variation in Healthcare: Reducing Unwarranted variation to increase value and improve quality. 2010. <http://www.rightcare.nhs.uk/index.php/atlas/atlas-of-variation-2010/> (date last accessed 5 June, 2015).

59. **Riddle DL, Jiranek WA, Hayes CW.** Use of a validated algorithm to judge the appropriateness of total knee arthroplasty in the United States: a multicenter longitudinal cohort study. *Arthritis Rheumatol.* 2014;66(8):2134-43.
60. **Quintana JM, Arostegui I, Escobar A, Azkarate J, Goenaga JI, Lafuente I.** Prevalence of knee and hip osteoarthritis and the appropriateness of joint replacement in an older population. *Arch Intern Med.* 2008;168(14):1576-84.
61. **Quintana JM, Escobar A, Arostegui I, Bilbao A, Azkarate J, Goenaga JI, et al.** Health-related quality of life and appropriateness of knee or hip joint replacement. *Arch Intern Med.* 2006;166(2):220-6.
62. **Riddle DL, Perera RA, Jiranek WA, Dumenci L.** Using surgical appropriateness criteria to examine outcomes of total knee arthroplasty in a United States sample. *Arthritis Care Res.* 2015;67(3):349-57.
63. **Ackerman IN, Dieppe PA, March LM, Roos EM, Nilsson AK, Brown GC, et al.** Variation in age and physical status prior to total knee and hip replacement surgery: a comparison of centers in Australia and Europe. *Arthritis Rheum.* 2009;61(2):166-73.
64. **Dieppe P, Judge A, Williams S, Ikwueke IF, Guenther K, Floeren M, et al.** Variations in the pre-operative status of patients coming to primary hip replacement for osteoarthritis in European orthopaedic centres. *BMC musculoskeletal disorders.* 2009;10:19-.
65. **Carr AJ, Morris RW, Murray DW, Pynsent PB.** Survival analysis in joint replacement surgery. *J Bone Joint Surg Br.* 1993;75:178-82.
66. **Price AJ, Longino D, Rees J, Rout R, Pandit H, Javaid K, et al.** Are pain and function better measures of outcome than revision rates after TKR in the younger patient? *Knee.* 2010;17(3):196-9.
67. **Devlin NJ, Appleby J.** Getting the most out of PROMs - Putting health outcomes at the heart of NHS decision-making. 2010. [www.kingsfund.org.uk/publications/getting-most-out-proms](http://www.kingsfund.org.uk/publications/getting-most-out-proms) (date last accessed March 2016).
68. **Xie F, Li SC, Lo NN, Yeo SJ, Yang KY, Yeo W, et al.** Cross-cultural adaptation and validation of Singapore English and Chinese Versions of the Oxford Knee Score (OKS) in knee osteoarthritis patients undergoing total knee replacement. *Osteoarthritis Cartilage.* 2007;15(9):1019-24.
69. **Naal FD, Impellizzeri FM, Sieverding M, Loibl M, von Knoch F, Mannion AF, et al.** The 12-item Oxford Knee Score: cross-cultural adaptation into German and assessment of its psychometric properties in patients with osteoarthritis of the knee. *Osteoarthritis Cartilage.* 2009;17(1):49-52.

70. **Dunbar MJ, Robertsson O, Ryd L, Lidgren L.** Translation and validation of the Oxford-12 item knee score for use in Sweden. *Acta Orthop Scand.* 2000;71(3):268-74.
71. **Murray DW, Fitzpatrick R, Rogers K, Pandit H, Beard DJ, Carr AJ, et al.** The use of the Oxford hip and knee scores. *J Bone Joint Surg Br.* 2007;89(8):1010-4.
72. **Harris K, Dawson J, Doll H, Field RE, Murray DW, Fitzpatrick R, et al.** Can pain and function be distinguished in the Oxford Knee Score in a meaningful way? An exploratory and confirmatory factor analysis. *Qual Life Res.* 2013;(1573-2649 (Electronic)).
73. **Beard DJ, Dawson J, Harris K, Doll H, Murray DW, Carr AJ, et al.** Minimal important change or difference for the Oxford hip and knee scores. ISAKOS Biennial Congress; Toronto2013.
74. **Choi WJ, Hwang SJ, Song JG, Leem JG, Kang YU, Park PH, et al.** Radiofrequency treatment relieves chronic knee osteoarthritis pain: a double-blind randomized controlled trial. *Pain.* 2011;152(3):481-7.
75. **Norrish AR, Jibri ZA, Hopgood P.** The LISS plate treatment of supracondylar fractures above a total knee replacement: a case-control study. *Acta Orthop Belg.* 2009;75(5):642-8.
76. **Haviv B, Bronak S, Thein R, Kidron A, Thein R.** Mid-term outcome of opening-wedge high tibial osteotomy for varus arthritic knees. *Orthopedics.* 2012;35(2):e192-6.
77. **Williamson L, Wyatt MR, Yein K, Melton JT.** Severe knee osteoarthritis: a randomized controlled trial of acupuncture, physiotherapy (supervised exercise) and standard management for patients awaiting knee replacement. *Rheumatology.* 2007;46(9):1445-9.
78. **Harris K, Dawson J, Jones LD, AJ Price, DJ Beard,** editors. Validation of the Oxford Knee Score for patients undergoing non-operative treatment for knee OA. ISAKOS Biennial Congress; 2013; Toronto.
79. **Xie F, Ye H, Zhang Y, Liu X, Lei T, Li SC.** Extension from inpatients to outpatients: validity and reliability of the Oxford Knee Score in measuring health outcomes in patients with knee osteoarthritis. *Int J Rheum Dis.* 2011;14(2):206-10.
80. **National Audit Office.** Delivering efficiency savings in the NHS. 2011.
81. **Briggs TWR.** Getting it right first time - improving the quality of orthopaedic care within the National Health Service in England. 2012. [http://www.timbriggs-gettingitrightfirsttime.com/downloads/BriggsReportA4\\_FIN.pdf](http://www.timbriggs-gettingitrightfirsttime.com/downloads/BriggsReportA4_FIN.pdf) (date last accessed June 2015).
82. **Klein R.** Rationing in the fiscal ice age. *Health Econ Policy Law.* 2010;5(4):389-96.

- 83. The Medical Technology Group.** Hip and Knee Replacement: Combating Patient Lotteries. 2012.
- 84. Arnett G, Hadorn DC.** Developing priority criteria for hip and knee replacement: results from the Western Canada Waiting List Project. *Can J Surg.* 2003;46(4):290-6.
- 85. Dew K, Cumming J, McLeod D, Morgan S, McKinlay E, Dowell A, et al.** Explicit rationing of elective services: implementing the New Zealand reforms. *Health policy* 2005;74(1):1-12.
- 86. Escobar A, Quintana JM, Gonzalez M, Bilbao A, Ibanez B.** Waiting list management: priority criteria or first-in first-out? A case for total joint replacement. *J Eval Clin Pract.* 2009;15(4):595-601.
- 87. Ham C.** Health care rationing. *BMJ.* 1995;310(6993):1483-4.
- 88. Escobar A, Quintana JM, Bilbao A, Ibañez B, Arenaza JC, Gutiérrez L, et al.** Development of explicit criteria for prioritization of hip and knee replacement. *J Eval Clin Pract* 2007;13(3):429-34.
- 89. De Coster C, McMillan S, Brant R, McGurran J, Noseworthy T.** The Western Canada Waiting List Project: development of a priority referral score for hip and knee arthroplasty. *J Eval Clin Pract.* 2007;13(2):192-6.
- 90. Hadorn DC, Holmes AC.** The New Zealand priority criteria project. Part 1: Overview. *BMJ* 1997;314(7074):131-4.
- 91. McLeod D, Morgan S, McKinlay E, Dew K, Cumming J, Dowell A, et al.** Use of, and attitudes to, clinical priority assessment criteria in elective surgery in New Zealand. *J Health Serv Res Po.* 2004;9(2):91-9.
- 92. Coleman B, McChesney S, Twaddle B.** Does the Priority Scoring System for Joint Replacement really identify those in most need? *J NZ Med Assoc* 2005;118(1215):1-6.
- 93. Royal College of Surgeons of England.** Is access to surgery a postcode lottery? 2014.  
[www.rcseng.ac.uk/news/docs/Is%20surgery%20to%20surgery%20a%20postcode%20lottery.pdf/view](http://www.rcseng.ac.uk/news/docs/Is%20surgery%20to%20surgery%20a%20postcode%20lottery.pdf/view) (date last accessed March 2016).
- 94. No Authors Listed.** NHS Manchester Topics for Inclusion or Change in the PCT Non-commissioned Procedures List. . 2010.  
[https://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CDwQFjAA&url=http%3A%2F%2Fwww.manchester.nhs.uk%2Fdocument\\_uploads%2F19%2520January%25202011%2FC2turnaroundAppendix%25202%2520TR.doc&ei=ssOQUeO7IMrB0QW6lYEQ&usg=AFQjCNG5Gonb3k9osSM-](https://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CDwQFjAA&url=http%3A%2F%2Fwww.manchester.nhs.uk%2Fdocument_uploads%2F19%2520January%25202011%2FC2turnaroundAppendix%25202%2520TR.doc&ei=ssOQUeO7IMrB0QW6lYEQ&usg=AFQjCNG5Gonb3k9osSM-)

[NqmnWqqQuNVWzQ&sig2=mkYkgamEZ8LF5fpm\\_tMsNA&bvm=bv.46340616,d.d2k](http://www.oxfordshireccg.nhs.uk/wp-content/uploads/2013/03/PS188-Knee-replacement-surgery.pdf)

(date last accessed 28 September 2013).

**95. NHS Oxfordshire.** INTERIM Treatment Threshold Statement: Knee Arthroplasty (Reference PS 188). 2010. <http://www.oxfordshireccg.nhs.uk/wp-content/uploads/2013/03/PS188-Knee-replacement-surgery.pdf>

(date last accessed 9 June 2015).

**96. City NHS Nottinghamshire County & NHS Nottingham.** Commissioning policy for procedures of limited clinical value. 2011.

<http://www.nottinghamnortheastccg.nhs.uk/wp-content/uploads/2014/04/10.-Policy-for-Procedures-of-Low-Clinical-Value-PLCV-Version-D-March-2011-NNE.pdf>

(date last accessed 10 June 2015).

**97. No Authors Listed.** 2014/15 South West London effective commissioning initiative. 2014.

[https://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CDEQFjAA&url=http%3A%2F%2Fwww.swlphnetwork.org.uk%2Fuploads%2Fdocuments%2Fswl-eci%2Feffective-commissioning-document-master-version-2010-11-final%5B1%5D.doc&ei=8cKQUd7FEpO10QXUIIC4BA&usg=AFQjCNEqnSbunJQCGNmM0tZWpTHvZ5C\\_FQ&sig2=6wpW4hqUJS35Z4qUkI5lsg&bvm=bv.46340616,d.d2k](https://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CDEQFjAA&url=http%3A%2F%2Fwww.swlphnetwork.org.uk%2Fuploads%2Fdocuments%2Fswl-eci%2Feffective-commissioning-document-master-version-2010-11-final%5B1%5D.doc&ei=8cKQUd7FEpO10QXUIIC4BA&usg=AFQjCNEqnSbunJQCGNmM0tZWpTHvZ5C_FQ&sig2=6wpW4hqUJS35Z4qUkI5lsg&bvm=bv.46340616,d.d2k)

(date last accessed 10 June 2015).

**98. Central and Eastern Cheshire PCT.** Cheshire & Merseyside prior approval scheme incorporating Procedures of Lower Clinical Priority (PLCP). 2011.

[https://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&cad=rja&ved=0CD8QFjAC&url=http%3A%2F%2Fwww.liverpoolccg.nhs.uk%2FLibrary%2FAbout\\_us%2FPublications%2FCheshire%2520and%2520Merseyside%2520Procedures%2520of%2520Low%2520Clinical%2520Priority%2520policy.doc&ei=Zs-QUZDfGMWo0AWw-IEI&usg=AFQjCNFlchOqfzR3uKFZNaXxXX9711HIUQ&sig2=AvGuNJW4ajH0E6fcI2yCMg&bvm=bv.46340616,d.d2k](https://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&cad=rja&ved=0CD8QFjAC&url=http%3A%2F%2Fwww.liverpoolccg.nhs.uk%2FLibrary%2FAbout_us%2FPublications%2FCheshire%2520and%2520Merseyside%2520Procedures%2520of%2520Low%2520Clinical%2520Priority%2520policy.doc&ei=Zs-QUZDfGMWo0AWw-IEI&usg=AFQjCNFlchOqfzR3uKFZNaXxXX9711HIUQ&sig2=AvGuNJW4ajH0E6fcI2yCMg&bvm=bv.46340616,d.d2k)

(date last accessed 9 June 2015).

**99. Mid Essex Clinical Commissioning Group.** Service Restriction Policy. 2014. [http://midessexccg.nhs.uk/your-nhs/medicines-management/clinical-priorities-policies/doc\\_download/592-service-restriction-policy-meccg78](http://midessexccg.nhs.uk/your-nhs/medicines-management/clinical-priorities-policies/doc_download/592-service-restriction-policy-meccg78)

(date last accessed 6 June 2015).

**100. North Staffordshire Clinical Commissioning Group.** Commissioning policy - excluded and restricted procedures. 2013.

<http://www.northstaffsccg.nhs.uk/download.cfm?doc=docm93jjm4n2767.pdf&ver=4814>

(date last

**101. Dakin H, Gray A, Fitzpatrick R, Maclennan G, Murray D.** Rationing of total knee replacement: a cost-effectiveness analysis on a large trial data set. *BMJ Open*. 2012;2(1):e000332.

**102. Hossain FS, Konan S, Patel S, Rodriguez-Merchan EC, Haddad FS.** The assessment of outcome after total knee arthroplasty: are we there yet? *Bone Joint J*. 2015;97-B(1):3-9.

**103. Acute Commissioning Team.** Commissioning policy for musculoskeletal surgical interventions. 2010. (June):

<https://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CDYQFjAA&url=http%3A%2F%2Fwww.worcestershire.nhs.uk%2FEasySiteWeb%2FGatewayLink.aspx%3FalId%3D3603&ei=CeWIUf-TIquT0AXe8YHQDw&usg=AFQjCNHHf17kuwmpxXA9eFpSLJaDlCitqg&sig2=OcUxkRM23j5ji4LzoNajIw&bvm=bv.47008514,d.d2k> (date last accessed 10 June 2015).

**104. Gloucestershire Clinical Commissioning Group.** Effective Clinical Commissioning Policies List (April 2013). 2013.

<https://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=11&ved=0CC4QFjAAOAO&url=http%3A%2F%2Fwww.gloucestershireccg.nhs.uk%2Fwp-content%2Fuploads%2F2012%2F03%2FEffective-Clinical-Commissioning-Policies-List-April-2013.doc&ei=pdeQUaftKeTK0AWrkYAw&usg=AFQjCNEY2AQfh5ihRIYsYwHPN419j1BoFA&sig2=Nz6NBkFwj0yY0bUhWxKymQ> (date last accessed 9 June 2015).

**105. Solihull CCG.** Commissioning policy - hip and knee replacement for patients with osteoarthritis. 2011. <https://solihullccg.nhs.uk/publications/our-policies-and-procedures-1/treatments-and-procedures-of-lower-priority-1/1165-hip-and-knee-replacement-for-patients-with-osteoarthritis/file> (date last accessed 6 June 2015).

**106. The Health and Social Care Information Centre.** Finalised Patient Reported Outcome Measures (PROMs) in England. April 2009 – April 2010: Pre- and postoperative data: Experimental statistics, 2010. 2010. <http://www.hscic.gov.uk/proms> (date last accessed 7 May 2015).

**107. Pynsent PB, Adams DJ, Disney SP.** The Oxford hip and knee outcome questionnaires for arthroplasty: outcomes and standards for surgical audit. *J Bone Joint Surg Br*. 2005;87-B(2):241-8.

- 108. Jenny JY, Diesinger Y.** The Oxford Knee Score: compared performance before and after knee replacement. *Orthop Traumatol Surg Res.* 2012;98(4):409-12.
- 109. The Health and Social Care Information Centre.** Provisional Monthly Patient Reported Outcome Measures (PROMs) in England , April to November 2009 (Pre-operative data only). 2009. <https://catalogue.ic.nhs.uk/publications/hospital/proms/prom-data-apr-09-nov-09/prom-data-arch-pre-op-0804-2010-rep1.pdf> (date last accessed 24 September 2013).
- 110. Field A.** Discovering Statistics using SPSS. 3rd ed: Sage; 2009.
- 111. Kelly KD, Voaklander DC, Johnston WC, Suarez-Almazor ME.** Equity in waiting times for major joint arthroplasty. *Can J Surg.* 2002;45(4):269-76.
- 112. Beard DJ, Harris K, Dawson J, Doll H, Murray DW, Carr AJ, et al.** Meaningful changes for the Oxford hip and knee scores after joint replacement surgery. *J Clin Epidemiol.* 2015;68(1):73-9.
- 113. Baker PN, Petheram T, Jameson SS, Avery PJ, Reed MR, Gregg PJ, et al.** Comparison of patient-reported outcome measures following total and unicondylar knee replacement. *J Bone Joint Surg Br.* 2012;94(7):919-27.
- 114. Yong PFK, Milner PC, Payne JN, Lewis PA, Jennison C.** Inequalities in access to knee joint replacements for people in need. *Ann Rheum Dis.* 2004;63:1483-9.
- 115. Appleby J, Raleigh V, Frosini F, Bevan G, Gao H, Lyscom T.** Variations in health care: The good, the bad and the inexplicable. 2011. <http://www.kingsfund.org.uk/sites/files/kf/Variations-in-health-care-good-bad-inexplicable-report-The-Kings-Fund-April-2011.pdf> (date last accessed 24 September 2013).
- 116. Department of Health.** The NHS Atlas of Variation in Healthcare: Reducing unwarranted variation to increase value and improve quality. . 2011. <http://www.rightcare.nhs.uk/index.php/nhs-atlas/atlas-downloads/> (date last accessed 6 June 2015).
- 117. Twisk JWR.** Applied multilevel analysis : a practical guide. Cambridge, UK ; New York: Cambridge University Press; 2006.
- 118. Leckie G.** Three-Level Multilevel Models - Concepts. LEMMA VLE Module 11. 2013. <http://www.bristol.ac.uk/cmm/learning/course.html> (date last accessed June 2013).
- 119. Gusi N, Olivares PR, Rajendram R.** The EQ-5D Health Related Quality of Life Questionnaire. Handbook of Disease Burdens and Quality of Life Measures: Springer, New York; 2010. p. 87-99.
- 120. Lingard EA, Katz JN, Wright EA, Sledge CB.** Predicting the outcome of total knee arthroplasty. *J Bone Joint Surg Am.* 2004;86-A(10):2179-86.

121. Fortin PR, Clarke AE, Joseph L, Liang MH, Tanzer M, Ferland D, et al. Outcomes of total hip and knee replacement: preoperative functional status predicts outcomes at six months after surgery. *Arthritis and Rheumatism*. 1999;42(8):1722-8.
122. Cobos R, Latorre A, Aizpuru F, Guenaga JI, Sarasqueta C, Escobar A, et al. Variability of indication criteria in knee and hip replacement: an observational study. *BMC Musculoskelet Disord*. 2010;11(1):249-.
123. Blackett J, Carslaw A, Lees D, Fargher S, Rao S, Pohl M. The impact of the 6-month waiting target for elective surgery: a patient record study. *N Z Med J*. 2014;127(1405):45-53.
124. Neogi T, Frey-Law L, Scholz J, Niu J, Arendt-Nielsen L, Woolf C, et al. Sensitivity and sensitisation in relation to pain severity in knee osteoarthritis: trait or state? *Ann Rheum Dis*. 2013.
125. King CD, Sibille KT, Goodin BR, Cruz-Almeida Y, Glover TL, Bartley E, et al. Experimental pain sensitivity differs as a function of clinical pain severity in symptomatic knee osteoarthritis. *Osteoarthritis Cartilage*. 2013;21(9):1243-52.
126. Steultjens MP, Dekker J, Bijlsma JW. Coping, pain, and disability in osteoarthritis: a longitudinal study. *J Rheumatol*. 2001;28(5):1068-72.
127. McHugh GA, Campbell M, Silman AJ, Kay PR, Luker KA. Patients waiting for a hip or knee joint replacement: is there any prioritization for surgery? *J Eval Clin Pract*. 2008;14(3):361-7.
128. Baker PN, van der Meulen JH, Lewsey J, Gregg PJ. The role of pain and function in determining patient satisfaction after total knee replacement. Data from the National Joint Registry for England and Wales. *J Bone Joint Surg Br*. 2007;89(7):893-900.
129. Robertsson O, Dunbar M, Pehrsson T, Knutson K, Lidgren L. Patient satisfaction after knee arthroplasty: a report on 27,372 knees operated on between 1981 and 1995 in Sweden. *Acta Orthop Scand*. 2000;71(3):262-7.
130. Scott CEH, Howie CR, MacDonald D, Biant LC. Predicting dissatisfaction following total knee replacement: a prospective study of 1217 patients. *J Bone Joint Surg Br*. 2010;92(9):1253-8.
131. Gandhi R, Davey JR, Mahomed NN. Predicting patient dissatisfaction following joint replacement surgery. *J Rheumatol*. 2008;35(12):2415-8.
132. Noble PC, Conditt MA, Cook KF, Mathis KB. The John Insall Award: Patient expectations affect satisfaction with total knee arthroplasty. *Clin Orthop Relat Res*. 2006;452(452):35-43.

- 133. Escobar A, Quintana JM, Bilbao A, Azkárate J, Güenaga JI, Arenaza JC, et al.** Effect of patient characteristics on reported outcomes after total knee replacement. *Rheumatology*. 2007;46(1):112-9.
- 134. Judge A, Arden NK, Cooper C, Javaid MK, Carr A, Field RE, et al.** Predictors of outcomes of total knee replacement surgery. *Rheumatology* 2012;51(10):1804-13.
- 135. Baker PN, Deehan DJ, Lees D, Jameson S, Avery PJ, Gregg PJ, et al.** The effect of surgical factors on early patient-reported outcome measures (PROMS) following total knee replacement. *J Bone Joint Surg Br*. 2012;94(8):1058-66.
- 136. Clement ND, Jenkins PJ, MacDonald D, Nie YX, Patton JT, Breusch SJ, et al.** Socioeconomic status affects the Oxford knee score and short-form 12 score following total knee replacement. *Bone Joint J*. 2013;95-B(1):52-8.
- 137. Lim JTK, Luscombe KL, Jones PW, White SH.** The effect of preoperative symptom severity on functional outcome of total knee replacement- patients with the lowest preoperative scores achieve the lowest marks. *Knee*. 2006;13(3):216-9.
- 138. Judge A, Arden NK, Kiran A, Price A, Javaid MK, Beard D, et al.** Interpretation of patient-reported outcomes for hip and knee replacement surgery: identification of thresholds associated with satisfaction with surgery. *J Bone Joint Surg Br*. 2012;94(3):412-8.
- 139. Hildon Z, Neuburger J, Allwood D, van der Meulen J, Black N.** Clinicians' and patients' views of metrics of change derived from patient reported outcome measures (PROMs) for comparing providers' performance of surgery. *BMC Health Serv Res*. 2012;12:171.
- 140. Burns EM, Rigby E, Mamidanna R, Bottle A, Aylin P, Ziprin P, et al.** Systematic review of discharge coding accuracy. *J Pub Health*. 2012;34(1):138-48.
- 141. Hutchings A, Neuburger J, van der Meulen J, Black N.** Estimating recruitment rates for routine use of patient reported outcome measures and the impact on provider comparisons. *BMC Health Serv Res*. 2014;14:66.
- 142. Harris KK, Dawson J, Jones LD, Beard DJ, Price AJ.** Extending the use of PROMs in the NHS--using the Oxford Knee Score in patients undergoing non-operative management for knee osteoarthritis: a validation study. *BMJ Open*. 2013;3(8):e003365.
- 143. HES Development (PROMs) Team.** Finalised patient reported outcome measures (PROMs) in England. April 2012 to March 2013. 2014. <http://www.hscic.gov.uk/proms> (date last

- 144. Department of Health.** Equity and Excellence: Liberating the NHS. 2010. [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/213823/dh\\_117794.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213823/dh_117794.pdf) (date last accessed 6 June 2015).
- 145. Jayadev C, Khan T, Coulter A, Beard DJ, Price AJ.** Patient decision aids in knee replacement surgery. *Knee*. 2012;19(6):746-50.
- 146. Hawker GA, Wright JG, Badley EM, Coyte PC.** Perceptions of, and willingness to consider, total joint arthroplasty in a population-based cohort of individuals with disabling hip and knee arthritis. *Arthritis Rheum*. 2004;51(4):635-41.
- 147. Toye FM, Barlow J, Wright C, Lamb SE.** Personal meanings in the construction of need for total knee replacement surgery. *Soc Sci Med*. 2006;63(1):43-53.
- 148. Hudak PL, Grassau P, Glazier RH, Hawker G, Kreder H, Coyte P, et al.** "Not everyone who needs one is going to get one": the influence of medical brokering on patient candidacy for total joint arthroplasty. *Med Decis Making*. 2008;28(5):773-80.
- 149. BMJ Group.** Shared Decision Making. 2015 [<http://sdm.rightcare.nhs.uk/pda/>] (date last accessed April 24, 2015);
- 150. Wylde V, Dieppe P, Hewlett S, Learmonth ID.** Total knee replacement: is it really an effective procedure for all? *Knee*. 2007;14(6):417-23.
- 151. Scott CE, Bugler KE, Clement ND, MacDonald D, Howie CR, Biant LC.** Patient expectations of arthroplasty of the hip and knee. *J Bone Joint Surg Br*. 2012;94(7):974-81.
- 152. Rothwell AG, Hooper GJ, Hobbs A, Frampto CM.** An analysis of the Oxford hip and knee scores and their relationship to early joint revision in the New Zealand Joint Registry. *J Bone Joint Surg Br*. 2010;92(3):413-8.
- 153. Kalairajah Y, Azurza K, Hulme C, Molloy S, Drabu KJ.** Health outcome measures in the evaluation of total hip arthroplasties--a comparison between the Harris hip score and the Oxford hip score. *J Arthroplasty*. 2005;20(8):1037-41.
- 154. Woolhead GM, Donovan JL, Chard JA, Dieppe PA.** Who should have priority for a knee joint replacement? *Rheumatology*. 2002;41(4):390-4.
- 155. Naylor CD, Williams JI.** Primary hip and knee replacement surgery: Ontario criteria for case selection and surgical priority. *Qual Saf Health Care*. 1996;5(1):20-30.
- 156. Hawker GA, Wright JG, Coyte PC, Williams JI, Glazier R, Wilkins A, et al.** Determining the need for hip and knee arthroplasty: The role of clinical severity and patients' preferences. *Med Care*. 2001;39(3):206-16.

- 157. Jones LD, Bottomley N, Harris K, Jackson W, Price AJ, Beard DJ.** The clinical symptom profile of early radiographic knee arthritis: a pain and function comparison with advanced disease. *Knee Surg Sports Traumatol Arthrosc.* 2014.
- 158. McHugh GA, Campbell M, Luker KA.** GP referral of patients with osteoarthritis for consideration of total joint replacement: a longitudinal study. *Br J Gen Pract.* 2011;61(589):e459-68.
- 159. Nilsson AK, Toksvig-Larsen S, Roos EM.** Knee arthroplasty: are patients' expectations fulfilled? A prospective study of pain and function in 102 patients with 5-year follow-up. *Acta orthopaedica.* 2009;80(1):55-61.
- 160. Dawson J, Beard DJ, McKibbin H, Harris K, Jenkinson C, Price AJ.** Development of a patient-reported outcome measure of activity and participation (the OKS-APQ) to supplement the Oxford knee score. *Bone Joint J.* 2014;96-B(3):332-8.
- 161. Jenkins PJ, Clement ND, Hamilton DF, Gaston P, Patton JT, Howie CR.** Predicting the cost-effectiveness of total hip and knee replacement: a health economic analysis. *Bone Joint J.* 2013;95-B(1):115-21.
- 162. Kim TK, Chang CB, Kang YG, Kim SJ, Seong SC.** Causes and predictors of patient's dissatisfaction after uncomplicated total knee arthroplasty. *J Arthroplasty.* 2009;24(2):263-71.
- 163. Hamilton DF, Lane JV, Gaston P, Patton JT, Macdonald D, Simpson AH, et al.** What determines patient satisfaction with surgery? A prospective cohort study of 4709 patients following total joint replacement. *BMJ Open.* 2013;3(4).
- 164. McLeod D, Morgan S, McKinlay E, Dew K, Cumming J, Dowell A, et al.** Clinicians' reported use of clinical priority assessment criteria and their attitudes to prioritization for elective surgery: a cross-sectional survey. *ANZ J Surg.* 2004;74(11):1003-9.
- 165. Dew K, Stubbe M, Macdonald L, Dowell A, Plumridge E.** The (non) use of prioritisation protocols by surgeons. *Sociol Health Ill.* 2010;32(4):545-62.
- 166. MacCormick AD, Tan CP, Parry BR.** Priority assessment of patients for elective general surgery: game on? *ANZ J Surg.* 2004;74(3):143-5.
- 167. Tudor R.** Points for pain : Waiting list priority scoring systems : May be the way forward , but we need to learn more about their effects. *BMJ.* 1999;318(7181):412-4.