EUGEN LUNGU

Identification of patients at risk of poor outcomes following hip or knee arthroplasty

DEPARTMENT OF BIOMEDICAL SCIENCES
FACULTY OF MEDICINE

Memoir presented to the Faculty of Medicine
in partial fulfillment of the requirements
for the degree
Masters of Sciences (M.Sc.)
Sciences Biomédicales, option Musculosquelettique
Université de Montréal

December 2015

© Eugen Lungu, 2015

Abstract

Total joint arthroplasties (TJA) are commonly performed procedures for patients afflicted with hip and knee osteoarthritis (OA), and although successful, these surgeries can yield suboptimal results in a non-negligible proportion of patients. In order to improve surgical outcomes, patients at risk of poor results could be targeted with focused interventions. However, the evidence regarding the ability to identify which patients are at risk of poor outcomes is scarce. The objectives of this memoir were 1) to systematically review the literature of preoperative determinants of medium-term patient-reported pain and physical function after total hip arthroplasty (THA) and total knee arthroplasty (TKA) and 2) to develop clinical prediction models allowing the individual identification of patients at risk of poor outcomes following THA and TKA.

Systematic literature searches targeting studies evaluating all studied determinants of pain and function following THA and TKA were performed in four important databases until April 2015 and October 2014 respectively. Moreover, retrospective data from 265 patients having undergone THA at the *Hôpital Maisonneuve-Rosemont* from 2004 to 2010 was used to develop a preliminary prediction algorithm (PA) to identify patients at risk of poor surgical results. Finally, prospective data from 141 patients recruited at their inclusion on a preoperative waitlist for TKA in three hospitals in Québec City, Canada and followed 6 months postoperatively was used to develop a clinical prediction rule (CPR) to identify patients at risk of poor outcomes

Twenty-two (22) studies evaluating determinants of poor pain and function after THA with moderate-to-excellent methodological quality found that preoperative levels of pain and function, higher body mass index (BMI), greater medical comorbidities, worse general health, lower education level, lower

OA radiographic severity and contralateral hip OA were consistently associated with poor THA outcomes. Thirty-four (34) studies evaluating determinants of poor pain and function after TKA with moderate-to-excellent methodological quality identified preoperative levels of pain and function, greater medical comorbidity, lower general health, greater levels of depression and/or anxiety, presence of back pain, greater pain catastrophizing and greater socioeconomic deprivation as consistently associated with worse outcomes.

A preliminary PA consisting of age, gender, BMI and three items of the preoperative Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) was able to identify patients at risk of suboptimal outcomes (worst quartile of the postoperative WOMAC score and perceiving their operated hip as artificial with minor or major limitations) on an average±standard deviation (SD) of 446±171 days after THA with a sensitivity of 75.0% (95% CI: 59.8 – 85.8), a specificity of 77.8% (95% CI: 71.9 – 82.7) and a positive likelihood ratio of 3.38 (98% CI: 2.49 – 4.57). A CPR consisting of five items of the preoperative WOMAC was able to predict the identity of patients awaiting TKA at the highest risk of poor outcomes (worst quintile of the postoperative WOMAC score) six months postoperatively with a sensitivity of 82.1 % (95% CI: 66.7 – 95.8), a specificity of 71.7% (95% CI: 62.8 – 79.8) and a positive likelihood ratio of 2.9 (95% CI: 1.8 – 4.7).

This memoir led to the identification of a list of determinants of pain and disability following TKA and THA with the highest level of evidence to date. Moreover, two clinical prediction models with good predictive capabilities were developed in order to allow the identification of patients at risk of poor outcomes following TKA and THA. These findings could help target the patients most likely to benefit

from interventions aimed at diminishing their risk profile and improving surgical outcomes of hip or knee arthroplasties. External validation of these rules is warranted before clinical implementation.

Keywords: Total joint arthroplasty, Systematic review, Determinants**,** Prediction**,** Postoperative pain**,** Postoperative function

Résumé

Les arthroplasties totales de la hanche (ATH) et du genou (ATG) sont souvent offertes aux patients atteints de dégénérescence articulaire sévère. Bien qu'efficace chez la majorité des patients, ces interventions mènent à des résultats sous-optimaux dans de nombreux cas. Il demeure difficile d'identifier les patients à risque de résultats sous-optimaux à l'heure actuelle. L'identification de ces patients avant la chirurgie pourrait permettre d'optimiser la gamme de soins et de services offerts et de possiblement améliorer les résultats de leur chirurgie. Ce mémoire a comme objectifs : 1) de réaliser une revue systématique des déterminants associés à la douleur et aux incapacités fonctionnelles rapportées par les patients à moyen-terme suivant ces deux types d'arthroplastie et 2) de développer des modèles de prédiction clinique permettant l'identification des patients à risque de mauvais résultats en terme de douleur et d'incapacités fonctionnelles suivant l'ATH et l'ATG.

Une revue systématique de la littérature identifiant les déterminants de la douleur et de la fonction suivant l'ATH et l'ATG a été réalisée dans quatre bases de données jusqu'en avril 2015 et octobre 2014, respectivement. Afin de développer un algorithme de prédiction pouvant identifier les patients à risque de résultats sous-optimaux, nous avons aussi utilisé des données rétrospectives provenant de 265 patients ayant subi une ATH à l'Hôpital Maisonneuve-Rosemont (HMR) de 2004 à 2010. Finalement, des données prospectives sur 141 patients recrutés au moment de leur inclusion sur une liste d'attente pour une ATG dans trois hôpitaux universitaires à Québec, Canada et suivis jusqu'à six mois après la chirurgie ont permis l'élaboration d'une règle de prédiction clinique permettant l'identification des patients à risque de mauvais résultats en terme de douleur et d'incapacités fonctionnelles.

Vingt-deux (22) études d'une qualité méthodologique moyenne à excellente ont été incluses dans la revue. Les principaux déterminants de douleur et d'incapacités fonctionnelles après l'ATH incluaient: le niveau préopératoire de douleur et de fonction, un indice de la masse corporelle plus élevé, des comorbidités médicales plus importantes, un état de santé générale diminué, une scolarité plus faible, une arthrose radiographique moins sévère et la présence d'arthrose à la hanche controlatérale. Trente-quatre (34) études évaluant les déterminants de douleur et d'incapacités fonctionnelles après l'ATG avec une qualité méthodologique moyenne à excellente ont été évaluées et les déterminants suivant ont été identifiés: le niveau préopératoire de douleur et de fonction, des comorbidités médicales plus importantes, un état de santé générale diminué, un plus grands niveau d'anxiété et/ou de symptômes dépressifs, la présence de douleur au dos, plus de pensées catastrophiques ou un faible niveau socioéconomique.

Pour la création d'une règle de prédiction clinique, un algorithme préliminaire composé de l'âge, du sexe, de l'indice de masse corporelle ainsi que de trois questions du WOMAC préopératoire a permis l'identification des patients à risque de résultats chirurgicaux sous-optimaux (pire quartile du WOMAC postopératoire et percevant leur hanche opérée comme artificielle avec des limitations fonctionnelles mineures ou majeures) à une durée moyenne ±écart type de 446±171 jours après une ATH avec une sensibilité de 75.0% (95% IC: 59.8 – 85.8), une spécificité de 77.8% (95% IC: 71.9 – 82.7) et un rapport de vraisemblance positif de 3.38 (98% IC: 2.49 – 4.57).

Une règle de prédiction clinique formée de cinq items du questionnaire WOMAC préopratoire a permis l'identification des patients en attente d'une ATG à risque de mauvais résultats (pire quintile du WOMAC postopératoire) six mois après l'ATG avec une sensibilité de 82.1 % (95% IC: 66.7 – 95.8),

une spécificité de 71.7% (95% IC: 62.8 – 79.8) et un rapport de vraisemblance positif de 2.9 (95% IC: 1.8 – 4.7).

Les résultats de ce mémoire ont permis d'identifier, à partir de la littérature, une liste de déterminants de douleur et d'incapacités fonctionnelles après l'ATH et l'ATG avec le plus haut niveau d'évidence à ce jour. De plus, deux modèles de prédiction avec de très bonnes capacités prédictives ont été développés afin d'identifier les patients à risque de mauvais résultats chirurgicaux après l'ATH et l'ATG. L'identification de ces patients avant la chirurgie pourrait permettre d'optimiser leur prise en charge et de possiblement améliorer les résultats de leur chirurgie.

Mots-clés: Arthroplastie de la hanche, Arthroplastie totale du genou, Revue systématique, Déterminant, Prédicteur, Douleur postopératoire, Fonction postopératoire

Foreword

This work encompasses a memoir comprising four articles published, accepted or in revision in peer-reviewed journals and is divided in seven chapters. Chapter 1 presents an introduction and a literature review of the main concepts underlying this work. Chapter 2 elaborates on the methodology that was employed. Four articles (Chapters 3, 4, 5 and 6) are subsequently presented. These articles are the result of the work under the supervision of my research director Dr. François Desmeules and my co-director Dr. Pascal-André Vendittoli. I am the main author of these articles. The article in Chapter 6 is based on previous work by Dr. François Desmeules; I participated in the interpretation of results and in the statistical analysis and led the manuscript writing. In the other three articles, I led the study design and coordination, collection of data, statistical analysis and interpretation of results as well as the editing of the manuscripts. Drs. Desmeules and Vendittoli, co-authors of all four articles, participated in the design and the realisation of the studies and supervised the drafting of the manuscripts. Sina Maftoon participated in the interpretation of results and the drafting of the manuscript in the chapter 3, and is one of the co-authors of the article. Chapter 7 encompasses the discussion as well as the limits, strengths, impacts and future avenues of the findings of this memoir.

I would like to thank everyone involved in the process of my research work. Dr. Desmeules provided inestimable guidance and support, and I am especially thankful for his ability to successfully reorient my work when solutions seemed limited. I would like to thank Dr. Vendittoli for his unequalled expertise and the unmatched advice regarding the practical application of my research. Many thanks to Sina Maftoon, whose help with the drafting of the manuscript in Chapter 3 greatly alleviated my workload during a particularly trying period of my studies.

I would like to express my gratitude to Odette Hinse for her assistance with the bibliographic search of the systematic reviews, to Janie Barry for her help with the grading of the methodological quality of the studies included in the systematic reviews and to Marie-Christine Stafford for her help with the statistical analysis. Special thanks to my parents and to Cristina for their inestimable support and incredible patience.

The *Fonds de Recherche en Santé - Québec* (FRSQ) and the *Faculté des études supérieures de l'Université de Montréal* have provided invaluable financial assistance during my studies, and I am greatly thankful for their support.

Table of contents

CHAPTER 5

List of tables

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

List of figures

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

List of appendices

List of abbreviations

ATG Arthroplastie totale du genou
ATH Arthroplastie totale de la hanche
AUC Area Under the ROC Curve
BKA bilateral knee arthroplasty
BMI body mass index
$CAD Canadian dollars
CART Classification and Regression Tree
CCI Charlson Comorbidity Index
CI confidence interval
CID clinically important difference
CINAHL Cumulative Index to Nursing and Allied Health Literature
CPR clinical prediction rule
CIRS Cumulative Illness Rating Scale
FRSQ Fonds de Recherche en Santé - Québec
HOOS Hip disability and Osteoarthritis Outcome Score
HHS Harris Hip Score
HMR Hôpital Maisonneuve-Rosemont
HRQoL health-related quality of life
HR hip resurfacing arthroplasty
LEFS Lower Extremity Functional Scale
MESH Medical Subject Heading
MCID minimal clinically important difference
MID minimally important difference
OA osteoarthritis
OHS Oxford Hip Score
OKS Oxford Knee Score
OMERACT-OARSI Outcome Measures in Rheumatology-Osteoarthritis Research Society International responder criteria
OR odds ratio
PA prediction algorithm
PASS patient acceptable symptom scale
PKA patellofemoral knee arthroplasty
PR Prediction rule
PROM patient-reported outcome measures
PSI Psychological Symptom Index
ROM range of motion
ROC receiver operating characteristics
STROBE Strengthening The Reporting of Observational Studies in Epidemiology
SF-12/36 Short Form 12/36 Questionnaire
SD standard deviation
THA total hip arthroplasty
TJA total joint arthroplasty
TKA total knee arthroplasty
UKA unicompartmental knee arthroplasty
WOMAC Western Ontario and McMaster Universities Osteoarthritis Index

CHAPTER 1

INTRODUCTION

1. Introduction

OA is a progressive condition affecting a more than 10% of the Canadian population aged 15 or older and has a predilection for the elderly population [1, 2]. Patients afflicted by hip or knee OA can experience significant pain and functional limitations [3-5]. Recent recommendations suggest both pharmacological and non-pharmacological approaches to the management of hip and knee OA [6-8]. The gold standard treatment for both hip and knee OA when all conservative methods have been exhausted is the replacement of the respective joints with a prosthesis (THA and TKA respectively) [8]. These interventions are generally associated with successful outcomes, and in most cases, lead to a dramatic amelioration of pain levels as well as an improvement in functional status [9-13]. This has steered a tremendous upsurge in the popularity of these interventions, as evidenced by an approximate six-fold increase in the demand for TKA in the following decades [14]. These statistics invariably suggest the economic burden of TJAs; effectively, just the in-hospital costs associated with these procedures are estimated at almost 1 billion Canadian dollars ($CAD) per year [15]. Clearly, strategies targeted at ameliorating the management of these patients are required in order to avert the rising demand and efficiently allocate resources to this clientele.

Although generally considered successful, TJAs can lead to poor pain-related and functional outcomes in 7-23% in those undergoing THA and in 10-30 % of those undergoing TKA [4, 12, 13, 16]. Moreover, 7 % of patients undergoing THA and up to 23% of those undergoing TKA are dissatisfied with their surgical outcomes [4, 17]. Identifying such patients could help clinicians and patients themselves in making the decision of going forward with such an intervention or it may lead to the implementation of medical and rehabilitation interventions to help these patients before and after the surgery [18]. For example, patients identified as potentially at risk of a poorer outcome before their

surgery could be enrolled in a prehabilitation program during pre-surgery wait or intensive rehabilitation could be planned postoperatively [19]. Moreover, targeting of modifiable risk factors in those at risk of poor outcomes could ultimately improve their surgical outcome. Eventually, in terms of health service organization, identification of patients risking poorer outcomes following TJA may allow stakeholders and clinicians to better plan healthcare resources required by patients who are most likely to benefit from these interventions, as well as diverting efforts from patients who are likely to fare well on their own [19, 20].

The identification of factors affecting the outcomes of TJAs and of patients at risk of poorer outcomes remains a challenge. The surgical outcomes are clearly complex and investigations of possible determinants have been primarily directed toward perioperative surgical complications and prosthetic-related factors [20]. Many personal, clinical, surgical or psychosocial factors have been associated with worse pain, function or poor satisfaction following TJA, however results have not been consistent across studies [20]. Several systematic reviews have attempted to summarize the determinants of pain and function following THA and TKA, however they included studies with low methodological quality and often only focused on one type of determinant [10, 18, 21, 22]. It thus remains a challenge to identify which TJA candidates will likely do well, or do poorly following surgery and may need targeted interventions [18, 20].

In light of the aforementioned shortcomings of TJA outcomes, identification of patients at risk of poor surgical results is paramount. Although prioritization tools have been developed to identify patients most in need of surgery [23], at the present time, no validated clinical prediction tools have been developed to identify patients at risk of suboptimal outcomes in terms of pain and disability. Such

tools could help manage more efficiently patients undergoing TJA and those waiting for their surgery. Development of new prediction tools are in line with the recommendations of the final report of the Canadian Federal Advisor on Wait Times, which suggested the implementation of new innovative models of care [24].

2. Literature review

2.1. Osteoarthritis

2.1.1. General principles and epidemiology

OA is a dynamic process characterized by a progressive loss of the articular cartilage in association with remodelling and sclerosis of the subchondral bone, resulting in a structural and functional deficiency of synovial joints [1, 25]. OA affects primarily the elderly, and as the population ages, the number of people expected to be afflicted with OA will increase, especially in the age group of 50 to 80 years old [26]. The prevalence of OA in Canada is anticipated to increase from 3 million (14%) in 2010 to 5.8 million (18%) in 2031 [1]. The total annual direct healthcare costs associated with OA in Canada are estimated to rise from 1.8 billion $CAD in 2010 to 8.1 billion $CAD in 2031 [1]. This rise can be attributed to an increase in the number of persons living with OA, to the increasing incidence of OA and greater longevity [1]. Other contributing factors are the increasing number of TJAs and greater use of services by patients treated surgically [1]. Disability attributed to OA in the population aged 15 and over is projected to increase from a prevalence of 2.3% (595,000) in 1991 to 3.3% (1.13 million) in 2031 [27].

2.1.2. Hip and knee OA

Hip and knee OA are important subgroups of articular degenerative diseases and represent the first cause of functional limitations when walking or in any other task involving the lower limb [28]. The natural history of knee and hip OA involve progressive degenerative changes of the articular surfaces, resulting in increased pain, loss of joint range of motion (ROM), and loss of function [29].

Recent guidelines suggest that an optimal management of hip and knee OA should rely on a combination of pharmacological and non-pharmacological approaches [6]. These include biomechanical interventions, administration of intra-articular corticosteroids, land-based and water-based exercise, self-management, education, strength and weight training and oral medication [6]. Whenever these interventions fail to alleviate joint OA-related symptomatology, it is almost universally recommended to proceed with a hip or knee replacement surgery, which are considered both effective and cost-effective [6].

2.2. Joint arthroplasties

2.2.1. Epidemiology

TJAs are commonly performed surgical interventions; for example, TKA ranks second among all orthopaedic procedures after fracture fixation [14]. In 2010-2011, it is estimated that more than 90,000 THAs and TKAs were performed in Canada [30]. There is an increasing trend in the number of TJAs performed in Canada; the rate at which these procedures were completed doubled from 1996 to 2006 [31]. It is estimated that these trends will continue to substantially evolve, with a 175% projected increase in the demand for THA and 673% for TKA [14]. The economic impact of these interventions is considerable. The total in-hospital costs of knee and hip replacements in Canada in 2011 was

estimated at 963 million $CAD [15]. Increasing trends are equally observed for hip and knee replacement-related acute care hospitalizations, a surrogate of the number of performed TJAs; 2014 data indicates a five-year increase of 16.5% for hip replacement and 19.2% for TKA [15]. A greater proportion of hip and knee replacements were performed on females based on data collected in 2014 (60.2% and 58.1% respectively) [15]. Age trends for males and females undergoing knee replacement were similar, the average values being 67.4 and 67.2 years respectively [15]. However, males undergoing hip replacement were on average younger than females (67.3 vs. 72.4 years old) [15]. The great majority of patients undergoing hip and knee arthroplasty were overweight or obese (74.4% and 86.7% respectively) [15]. The most important indication for primary hip arthroplasty is OA (76.5% in 2013), with acute hip fracture a distant second (13.7%) [15]. An even greater proportion of primary TKAs were conducted for OA (97.1%) [15].

2.2.2. Surgical techniques and types of prostheses

2.2.2.1. TKA and other types of knee arthroplasty

The TKA procedure typically involves the replacement of the articular surfaces of the femur, the tibia and the patella by implants made of metal alloy and polyethylene. The vast majority of implants today are fixed with the use of cement. Two types of implants and procedures are used for the stabilization of the prosthesis: the postero-stabilized implant with ablation of the posterior cruciate ligament or the non postero-stabilized type where the knee ligaments are still competent. A meta-analysis has not demonstrated any differences in terms of durability, either in the short or the long term, between the two types of implants [32]. There is no consistent evidence as to whether one type of implant, stabilization or fixation produces better short or long term outcomes.

Less invasive alternatives to TKA include unicompartmental knee arthroplasty (UKA) and bicompartmental knee arthroplasty (BKA). The most popular intervention, UKA, entails the replacement of the medial or the lateral articular surfaces of the tibiofemoral joint [33]. Traditionally, its indications are more strict when compared to TKA: a diagnosis of OA or osteonecrosis of the medial or the lateral compartment of the tibiofemoral joint, an age greater than 60, a low demand for activity, weight under 82 kg, minimal pain at rest suggesting a low possibility of an inflammatory component, a range of motion arc greater than 90° with less than 5° of flexion contracture and an angular deformity smaller than 15° that is passively corrected to neutral [34]. However, recently, indications have expanded in order to include younger and heavier patients [33]. When compared to TKA, UKAs tend to have higher revision rates, but possibly lead to a more normal feeling in the knee, smaller risk of infection and less extensive surgery [35]. Similar short-term and long-term clinical results are noted for UKA procedures when compared to TKAs [36, 37].

2.2.2.2 Hip arthroplasty

THA entails the replacement of the native femoral head and acetabulum with artificial materials. Four basic components are employed for THA: the acetabular component, the acetabular insert, the femoral head and the femoral component.

Different bearing or articulating surfaces have been developed. The most commonly used bearing surfaces in 2012 in Canada was metal-on-polyethylene (more than 80% of THAs) [38]. Because of the wear-related complications associated with the polyethylene components, alternative bearing couples have been developed. Ceramic components have the lowest wear rates among any bearing surface [38]. However, they are more brittle, and up to 17% of patients with ceramic-on-ceramic THAs

experience audible squeaking noise [38, 39]. Metal-on-metal bearings have low wear rates and are less brittle than ceramic components. They support the use of larger femoral head sizes, which are associated with improved stability and lower rates of dislocation. Nevertheless, metal-on-metal bearings raise a great concern due to local adverse reactions to metal debris, with host responses causing local inflammation and formation of pseudotumors, and occasionally, soft-tissue and bony destruction [40]. Although rates of complications and revisions differ between the types of bearings, recent meta-analyses reported no difference in terms of physical functioning and pain [40, 41].

Three different methods of fixation are employed by surgeons in order to secure the femoral stem and the acetabular components. Bone cement can be used to fix the implant to the patient's natural remaining bone stock. Cementless implants are initially secured with bone screws or are press-fit into position to allow secondary fixation with bone integration on the implant surface. A hybrid approach combines cemented and cementless implant components. Cemented implants are associated with higher rates of implant loosening, osteolysis and femoral stem fracture and are occasionally employed in older and lower-demand patients [38]. Cementless implants were the most frequently used fixation method in Canada in 2010 (84%), followed by hybrid (15%) and cemented (1%) implants. A recent meta-analysis showed unclear results regarding whether cemented or cementless fixation are related to better patient-reported clinical results [42].

Hip resurfacing arthroplasty (HR) is an alternative to THA with a femoral stem in patients who are younger, more active, with normal kidney function and appropriate proximal femoral bone morphology and quality [43]. It involves the resurfacing of the native femoral head with a metal cap and the implantation of an acetabular component with a polished inner cobalt-chrome metal surface

[38]. Theoretical advantages of HR include a greater potential of replicating physiological hip function, lower rates of dislocation, and improved levels of activity with minimal wear of the articular surfaces [40]. Multiple studies concluded that patient-reported outcomes in terms of pain and function are similar between THA with a femoral stem and HR [40].

2.3. Defining outcome after joint arthroplasty

Evaluating the surgical results after a procedure entails establishing a definition of what represents a good and a bad outcome. Measuring surgical results following hip and knee arthroplasty is clearly complex, as evidenced by the different outcomes used to evaluate its effectiveness. For the greatest part, assessment of joint arthroplasty outcomes has focused on surgical and technical aspects [20]. However, recently, a shift towards patient-reported outcome measures (PROMs) has been advocated, and this is evidenced in the latest literature [20, 44]. Disease-specific measures for domains such as pain, function and stiffness have received notable attention. Despite the fact that such measures do not necessarily provide a holistic view of the patients' health status on their own, they offer valuable information regarding aspects that are important from the patients' perspective [45]. There have been several approaches at quantifying the results of joint arthroplasty in terms of PROMs. One method is the concept of "change" in health status following the intervention when compared to the status preoperatively, which can be viewed as a measure of effectiveness of TJA [46]. Recently, attempts at identifying a value for the change in health status that can be considered appropriate have been undertaken, as evidenced by the development of tools such as the OMERACT-OARSI set of responder criteria and the minimal clinically important difference or improvement (MCID or MCII) (see Appendices A and B) [47-51]. These approaches entail the use of an anchor item evaluating the patients' own impression regarding the outcomes of the intervention, that is further employed to

identify a cut-off value for what represents an appropriate surgical results in terms of pain, stiffness and function (Appendices A and B). The other school of thought focuses on the concept of "patient state" following the intervention, regardless of the level of improvement through surgery [46]. This has led to the development of the concept of patient-acceptable symptom state (PASS) in order to quantify an appropriate level of symptoms following therapeutic interventions, including THA and TKA [48]. However, no consensus has been reached regarding which approach represents the most valid method of assessing the results of TJA. As can be seen Appendix B, the main difficulty resides in the different values obtained for each type of outcome assessment among different studies. This is probably due to different anchor items employed in each study, different methodologies of deriving these values, different follow-up time points as well as lack of validation of derived cut-off values in several studies.

2.4. Evolution of patients following a TJA

Both TKA and THA are generally considered effective and safe procedures and allow for adequate alleviation of pain and improvement of function as well as health-related quality of life (HRQoL) for the majority of patients undergoing the procedures [9, 11, 52-54]. These interventions are however associated with perioperative and immediate postoperative complications that can significantly jeopardize the well-being of the patients [52]. The most feared complication, infection of the joint, occurs in approximately 1% of TKA procedures and up to 2.5% of THA [39, 55, 56]. Up to 2% of patients can experience nerve injuries, most frequently peroneal nerve palsy for TKA and the peroneal division of the sciatic nerve in THAs [57, 58]. Despite adequate thromboprophylaxis, symptomatic thromboembolic events occur in 2-3% of patients, with up to 15% demonstrating evidence of deep vein thrombosis on venographic studies [59, 60]. Surgical revision of a failed implant occurs in only 2% of TKA cases at 5 years and 5% at 10 years for THA [61, 62].

A growing body of evidence indicates that a significant proportion of patients will experience poor outcomes; as many as 25% of THA patients and 30% of TKA patients will suffer from chronic pain and disability and up to 23% will experience dissatisfaction with the surgical outcomes [4, 12, 13, 16, 17]. The reasons behind such poor outcomes remain unclear, rendering the identification of factors affecting the outcomes of knee and hip arthroplasty and of patients at risk of poorer outcomes a significant challenge. The outcomes of joint arthroplasties are clearly complex and investigations of possible determinants have been primarily directed toward perioperative surgical complications and prosthetic-related factors [20]. Many demographic, socioeconomic, clinical or psychosocial factors have been associated with worse pain, function and HRQoL following TKA and THA, however results have not been consistent across studies and the precise impact or strength of the association between these factors and the outcomes remain elusive. It thus remains a challenge to identify which join arthroplasty candidates will likely do well, or do poorly following their surgery.

2.5. Determinants of TKA and THA outcomes in terms of pain and functional limitations

Several systematic reviews have attempted to summarize the determinants of pain and/or functional limitations after TKA and THA. Santaguida et al. (2008) concluded that older age and female gender are significant determinants of worse function following TKA [18]. Ethgen et al. (2004) suggested that female gender, presence of comorbidity and poor preoperative quality of life are all determinants of worse outcomes after surgery [10]. A systematic review by Vissers et al. (2012) identified pain catastrophizing and lower preoperative mental health as significant determinants of poor TKA outcomes [13]. The main limitations of the published reviews is their focus on solely one type of determinant, inclusion of studies with poor methodological quality, reporting of results based on conjoined TKA and THA cohorts and a large range of follow-up periods. Moreover, there has been an

increasing amount of novel evidence published recently on the topic, notably with the investigation of novel possible determinants, which can be generally classified as demographic, socioeconomic, psychosocial, clinical, surgical and healthcare-related. This motivates a comprehensive assessment of the literature of determinants of TJA in a structured manner.

2.6. Identification of patients at risk of poor outcomes following joint arthroplasty

2.6.1. General principles of prediction tools

By developing accurate and easy-to-use prediction tools, better case management of patients scheduled to undergo joint replacement may be achieved. A clinical prediction tool can be defined as "a tool that quantifies the individual contributions that various components of the history, physical examination, and basic laboratory results make toward the diagnosis, prognosis, or likely response to treatment in an individual patient" [63]. The prediction tools are intended to assist the clinician's diagnostic or prognostic assessment, and tend to be used when decision-making is difficult or when achieving cost-saving without compromising quality of patient care is possible [64].

Several phases are required in order to adequately implement a clinical prediction tool. Firstly, during the development phase, identification of predictors from an observational study is undertaken in order to choose a group with the best predictive capabilities [65]. Several statistical methods can be employed to develop a prediction tool, each presenting their own advantages and disadvantages. Scoring systems and predictive models derived from univariate or multivariate analyses allow the allocation of a "weight" to factors significantly associated with the desired outcome [65]. These models are easy to interpret, but are not fully reliable at eliminating the bias associated to the interaction between independent variables [65]. Nomograms are graphical calculating devices allowing

the graphical computation of complicated formulas [65]. Many nomograms are developed using regression analysis, and although their use may be more simplistic, they retain the disadvantages associated with their inherent methodology [65]. Artificial neural networks are computational models similar in structure and function to biological neural networks, where input flowing through the network generates an output, and subsequent inputted data sets lead to an adjustment of the relationship that predicts the outcome [65]. They can identify complex non-linear associations between variables, but can be difficult to use clinically due to their requirement of greater computational resources when compared to other types of prediction models [65]. Recursive partitioning analyses, of which the classification and regression tree (CART) approach is the most popular, is a non-parametric method allowing the development of a decision tree splitting the patient populations into progressively smaller samples based on risk factors [64]. The decision tree is generally easily understood by clinicians, and can be well translated into everyday practice. Moreover, this approach is more liberal regarding required assumptions as compared to more traditional methods such as logistic regression [66]. However, some authors suggest that the CART approach may be less accurate than other models, as the amount of data contained in the "leaves" may be insufficient to reliably predict the outcome [65]. Moreover, the developed trees suffer from instability, as small changes in the sample or different cutting points may significantly alter the entire structure [67].

The next phase of the establishment of a clinical decision tool is its validation in order to ensure its reliability. This can be achieved in different ways. Several statistical methods can be employed in order to determine whether the associations between the given predictors and outcomes are due purely to chance, i.e. establishing the internal validity of the prediction algorithm. The most straightforward method, the split-sample approach, involves developing the prediction rule in a random half of the

sample, termed the “training sample”, and testing its validity in the other half of the sample, namely the “validation sample”. An extension of this is the cross-validation approach, which entails removing a certain number of patients from the sample, generating the rule using the remainder of patients, testing its performance on the subsample that was initially removed from the sample, and repeating the procedure several times such as data on all subjects is used to develop the model [68]. The most efficient approach of internal validation is however bootstrapping, which involves drawing samples with replacement from the original data set that have the same sample size as the original sample set followed by testing of the predictive model in those samples [68]. Typically, 1,000 bootstrap resamples are completed. In the case of the CART approach, the bootstrap estimates of the predictive qualities along with their 95% confidence intervals are calculated, and subsequently compared to the respective asymptotic values. If the confidence intervals are similar, it is deemed that the model has a satisfactory internal validity.

Assessment of the external validity of the prediction rule can be undertaken by testing its performance in a different population from the one where it was initially developed. Several methods exist, each representing different levels of validity. The prediction model can be validated in a retrospective cohort, in a narrow prospective cohort, in a large prospective cohort including a broad spectrum of patients or in several smaller settings that differ from one another [64].

Following validation of the prediction model, the impact analysis phase entails measuring the usefulness of the rule in a clinical setting in terms of cost-benefit, patient satisfaction and time-resource allocation [65]. A proper implementation of the prediction rule assumes that its use changes physicians’ behaviours, and this leads to either improvement of the outcomes or optimization of costs

while maintaining the standard of care [64]. Only after all of these conditions are filled, a prediction model can be accepted and adopted in clinical practice.

2.6.2. Developed prediction tools in the musculoskeletal literature

To our knowledge, one model predicting the identity of patients at risk of poor outcomes after THA has been developed [69]. The model consisted of patient age, BMI and gender, and had a sensitivity of 87.5% (95% CI 52.9 – 97.8), a specificity of 72.4% (95% CI 54.3 – 85.3) and a positive likelihood ratio of 3.17 (95% CI 1.66 – 6.05). Unfortunately, the model was developed in a small sample of 37 patients, and surgical success was defined uniquely by functional outcomes (change between the preoperative and six-month Lower Extremity Functional Scale - LEFS- score). A priority-setting tool for TKA and THA has been developed and validated by the Western Canada Waiting List project partnership [23, 70]. This tool is able to quantify the level of urgency of the status of patients awaiting total knee or hip replacement; however, its predictive capabilities have not been studied. Other prediction tools intended at improving the management of patients with different musculoskeletal pathologies have been developed and validated. The Ottawa Knee and Ankle rules are well-known, validated clinical prediction tools that are extensively employed in the evaluation of the requirement of radiological assessment of suspected cases of knee and ankle fractures in the emergency department [71, 72]. The 5-item Cassandra Rule has been developed in order to identify patients with non-specific back pain who are most likely to develop or sustain long-term functional limitations [73].

3. Objectives

The objectives of this memoir were:

- To systematically review the literature of determinants of poor short- and medium-term surgical outcomes in terms of pain and functional limitations following THA and TKA in order to identify the determinants with the greatest level of evidence
- To develop preliminary clinical prediction tools allowing the preoperative identification of patients at risk of poor short- and medium-term outcomes following joint arthroplasties

CHAPTER 2

Methodology

1. Systematic reviews

This section describes the methodology of the systematic reviews of the literature of determinants of pain and function following THA and TKA (Chapters 3 and 5 respectively). Both reviews had similar methodologies, and are therefore presented conjointly.

1.1. Type of study

These studies entailed a systematic review of the literature of determinants of pain and functional outcomes following THA and TKA respectively.

1.2. Literature search and study identification

With the help of a qualified documentation technician, four databases (Medline, Pubmed, Embase and CINAHL) were searched from their respective inception dates using a combination of keywords and Medical Subject Headings (MESH) terms (see Appendices C and D). The search strategy was developed in order to identify all possible determinants of THA or TKA outcomes. Manual searches of previously published reviews as well as of the reference lists of representative articles were also conducted. Two reviewers (Eugen Lungu and Janie Bary) independently reviewed the titles, abstracts and full texts of articles in order to evaluate their eligibility. A third reviewer (François Desmeules) was available to remediate any differences, if required.

1.3. Selection criteria

The following selection criteria were applied in order to ensure homogeneity of the included studies as well as appropriate generalizability.

- Participants were patients undergoing primary unilateral THA or TKA with ≤ 10% of

the sample undergoing UKA (in the case of the review on determinants of TKA outcomes), bilateral or revision TJA.

- ≥90% of the study sample was diagnosed with hip or knee OA
- Results are presented for a follow-up between 6 weeks and 2 years
- The outcome measure was a disease-specific validated PROM assessing pain and/or function
- Identification of determinants was obtained using multivariate analysis
- Article was published in English or French due to the restrictions imposed by the languages mastered by the reviewers.

1.4. Data extraction

A standardized form was employed to extract data. Participants' characteristics (diagnosis, type of surgery, age and gender proportion), number of patients, follow-up period, outcome measure employed, statistical methods used and statistical adjustments, as well as significant and non-significant determinants reported by the study were recorded. Each article was extracted by one of the raters and verified by another in order to reduce the risk of extraction errors.

1.5. Appraisal of the methodological quality

Two trained reviewers independently performed the appraisal of the methodological quality of the included studies and results were discussed in order to reach consensus. In case of disagreement, a third reviewer was available for mediation of differences.

The methodological quality of the included studies was assessed by evaluating the risk of bias using a modified version of the *Methodology Checklist for Prognostic Studies* developed by Hayden et al. (2003) [74] (see Appendix E). Six items define this tool: "Study participation", "Study attrition", "Prognostic factor measurement", 'Outcome measurement", "Confounding measurement and account" and 'Analysis". Each item evaluates the risk of potential methodological bias: "yes" indicates a low risk of bias, "no" indicates a high risk of bias and "unclear" indicates an unclear or an unknown risk given the information available from the article. A score of 2 was attributed if a low risk of bias was present, a score of 1 if the risk was judged unclear and 0 if the risk was high. For the 'Study participation" item, a score of 1 was given if the study was retrospective in nature and that information regarding patients not included in the study because of incomplete data was lacking. For the "Study attrition" item, a score of 0 was given automatically if the follow-up proportion at the relevant time-point was inferior to 80%. A score of 0 was given for the "Confounding measurement and account" item if confounding factors such as age, gender and BMI were not accounted for in the multivariate analysis. A total score was calculated by summating the individual scores of items, and a final standardized score was provided.

1.6. Data synthesis

Determinants of THA and TKA outcomes were summarized based on whether pain and function were assessed as separate or combined constructs. Moreover, due to a greater number of included studies yielding more data compared to the THA review, determinants of TKA outcomes were also reported based on whether pain and function were measured as postoperative change (difference between postoperative and preoperative levels) or postoperative status. A quantitative analysis was envisaged in order to pool the magnitude of association of the significant determinants with the outcomes. Given the

nature of the study designs and the heterogeneity of the included studies regarding variables' constructs and definitions, only a qualitative synthesis of results was eventually performed.

2. Prediction models

This section describes the methodology employed for the development of the prediction models presented in the chapters 4 and 6.

2.1. Preliminary prediction algorithm – hip arthroplasty

2.1.1. Type of study

This study has entailed a retrospective review of prospectively collected repeated-measures data of patients undergoing hip arthroplasty, with the follow-up ranging from just before the intervention until 12 to 24 months following surgery. This type of study allowed the assessment of the patient evolution from the preoperative to the postoperative period.

2.1.2. Recruitment location

The prospectively collected arthroplasty database of the HMR was consulted in order to select patients eligible for inclusion in the study. The database contains extensive preoperative and follow-up information on patients undergoing numerous orthopaedic hip interventions, including THA and HR. Independent assessors not involved in the medical care of the patients collect the prospective data. All patients provide informed consent prior to their inclusion in the database.

2.1.3. Patient selection

Patients were eligible for inclusion in the study if they satisfied the following criteria: 1) patients undergoing primary unilateral THA or HR, 2) diagnosis of primary hip OA, 3) complete preoperative and one to two-year postoperative self-reported outcome questionnaire responses.

Patients were excluded based on the following exclusion criteria: 1) THA or HR of the contralateral hip before the relevant follow-up evaluation, 2) revision of the implant before the 12 to 24 month follow-up, 3) diagnosis of inflammatory hip arthritis, paediatric hip disease, post-traumatic hip or any hip disease other than primary OA.

2.1.4. Variables and measures

The arthroplasty database of the HMR contains an extensive set of variables collected both preoperatively and postoperatively. However, a non-negligible proportion of data is missing. Therefore, in order to ensure the development of a prediction model with proper accuracy, we aimed to maximize the number of patients with complete preoperative and postoperative data for the greatest number of variables. Because we intended to build a model with an utmost clinical applicability, the limiting factors for the inclusion of patients were the presence of complete data for each item of the preoperative WOMAC, a number of individual answers of the postoperative WOMAC that would allow imputation of the total score as well as presence of an answer to the joint perception question. In the case of availability of complete data at 12 and at 24 months postoperatively, data at 12 months was included. Potential resulting selection bias associated with present design is addressed in the discussion sections of Chapter 4.

Dependent variable

Pain and functional status were assessed preoperatively and 12 to 24 months postoperatively with the WOMAC, a disease-specific patient-reported outcome measure, specifically designed to assess pain, stiffness and function in patients with hip or knee OA [75] (See Appendix F). It consists of the following domains: pain (5 items), stiffness (2 items) and functional limitation (17 items). Items are scored on a 5-point Likert scale representing different degrees of intensity (none, mild, moderate, severe or extreme). The scores of each domain as well as the total score were standardized on a 0 to 100 scale, with a greater score indicating more pain, stiffness or functional limitation. The psychometric qualities of the WOMAC, including its responsiveness, convergent construct validity and reliability have been found excellent for evaluating patients with hip and knee OA undergoing TJA [75, 76]. Moreover, a French-Canadian version of the WOMAC has been successfully validated [77]. Missing postoperative values were imputated as described in the WOMAC User Manual [78].

Patients' perception of their operated joint was measured at the 12 to 24 month follow-up with a multiple-choice question: "How do you perceive your operated hip?". The possible responses were "Like a native or natural joint", "Like an artificial joint with no restriction", "Like an artificial joint with minimal restriction", "Like an artificial joint with major restriction" and "Like a non-functional joint" (See Appendix G). Evaluation of joint perception has been strongly correlated with validated clinical scores of patient-reported outcome measures in patients undergoing TJA and can be employed as a measure of patient satisfaction [79]. Moreover, joint perception was found to be similar among patients undergoing THA and HR [79].

No consensus exists regarding what constitutes a poor outcome following hip arthroplasty. We therefore attempted to classify patients as having a poor outcome based on different definitions, and build prediction models accordingly, with the intent to ultimately choose the best algorithm (see Statistical Analysis). Definitions included a combination of patients being in the worst teritle or quartile of the total postoperative WOMAC score and different levels of joint perception. Ultimately, because this categorization yielded the best model in terms of predictive capabilities, clinical applicability and ease-of-use, patients were classified as having a poor outcome if they were in the worst quartile of the postoperative WOMAC score and perceived their joint "Like an artificial joint with minimal restriction", "Like an artificial joint with major restriction" or "Like a non-functional joint".

Independent variables

Demographic variables

Among the available demographic variables, age at the time of surgery, gender and employment status (employed, household, retired or other) were collected from the arthroplasty database. Complete data was only available for age and gender, which were included in the final analyses.

Clinical variables

BMI and presence of back pain were available from the arthroplasty database. Comorbidities that were obtainable included diabetes, gastrointestinal disease, immunosuppression, cardiac disease, obesity, osteoporosis, pulmonary disease, neurological disease, urological disease and other conditions. Pain localization was defined as presence of pain at any of the following locations: buttocks, greater

trochanters, groin, thigh, knee, calf, radicular or elsewhere. Presence of hip pain at rest, after first few steps, after a long walk and during sexual intercourse was also surveyed.

Several other variables that had a missing proportion of up to 15% were not included in the analysis but were collected in order to better characterize the study population. Contralateral hip status was defined as unaffected by OA or affected by OA but not operated (THA, HR nor other intervention). Knee status was defined as affected or unaffected by OA. Level of walking performance was determined according to whether the patient was unable to walk with aid, used crutches, two canes, one cane on a permanent basis with instability, one cane exclusively for outdoor activities or one cane for long distance walking. The level of activity in the three months before surgery was quantified as performing heavy work or sport, moderate work, mild work/sport, sedentary or immobile. Duration of walk before eliciting pain was determined as follows: walking unaffected, 31-60 minutes, 11-30 minutes, 2-10 minutes, less than 2 minutes or walking impossible. Patients' Charnley class was recorded and defined as follows: patients assigned to the class A had a single joint arthroplasty and no significant medical comorbidity; patients in class B had one other joint in need of arthroplasty or an unsuccessful or failing arthroplasty in another joint; class C patients have multiple joints in need of arthroplasty, multiple failing arthroplasties or significant medical or psychosocial impairment [80]. The Charnley classification is a commonly used scale in the orthopaedic literature that allows the stratification of patients according to their variability in outcome, which is quantified as the walking ability [80]. It has been shown that the patient's Charnley class also influences their outcome as measured by the WOMAC scale [81].

2.1.5. Statistical analysis

Descriptive statistics (mean, standard deviation and frequencies) were employed accordingly to describe the included patients' characteristics. Baseline and follow-up mean individual subscale and total WOMAC scores along with their standard deviations were calculated. Graphical representations of the scores were used in order to observe their distribution and assess their normality. Differences between time points in relation to total WOMAC scores and the respective domains was assessed using paired samples Student-t tests, with a significance level set at 0.05. Analyses were carried out with the IBM SPSS Statistics Version 20.0 (SPSS Inc., Chicago).

One of the most effective algorithms of recursive partitioning is the CART approach, and it was therefore employed to develop the preliminary PA [82]. It is based on maximizing the within-node homogeneity and minimizing the within-node error by evaluating all combinations of potential predictors. The Gini impurity measure was used as a splitting criterion for the development of the decision trees [83]. Because we aimed to develop a PA with the highest possible validity and due to the relatively small sample size, data for all the patients in the training set was used to develop the models. Firstly, all the variables were included in the analysis and models were developed using an automated approach. Secondly, several potential predictor variables were manually chosen based on several criteria, including statistical, clinical and ease-of-use qualities, in order to develop further models. For example, information on variables such as age and gender would be more readily attainable clinically compared to the number of comorbidities. Sensitivity, specificity, positive and negative predictive values as well as positive and negative likelihood ratios were calculated for all the developed models along with their 95% confidence intervals [84]. The model that showed the highest level of sensitivity and an appropriate level of specificity as well as fitting the ease-of-use criterion was selected among

all of the proposed models. This method was employed seeing that we aimed to develop a screening tool, i.e. a model able to identify patients at risk of poor surgical outcomes with the smallest proportion of false negatives. Internal validity of the model was then appraised using 1,000 bootstrap resamples [85].

2.1.6. Ethics and confidentiality

This study was part of a larger research protocol that was reapproved annually by the *Comité d'éthique de la recherche* of the HMR entitled *"Évaluation de vos capacities fonctionnelles et de votre qualité de vie en pré chirurgie, péri chirurgie et post chirurgie"* (see Appendix H). The ethics committee approved annually the information and consent form, as well as the database politics (see Appendix I).

2.2. Clinical prediction rule – total knee arthroplasty

This study was based on a prospectively collected cohort that recruited patients undergoing TKA from February 2006 to July 2007. As mentioned in the Foreword section, I have not participated in the study design, patient recruitment or data collection, but the methodology will be summarized briefly in the following sections.

2.2.1. Type of study

A prospective longitudinal design with repeated-measures was employed. This study was part of a larger study evaluating the effect of waiting time on the preoperative state and the postoperative outcome of patients awaiting TKA.

2.2.2. Settings and patient recruitment

Patients who were newly enrolled on the waiting lists for TKA in three teaching hospitals in Québec City, Canada (CHUL, HSFA and HDQ) were recruited weekly between February 2006 and July 2007. All seven orthopaedists practicing TKA in these hospitals participated in the study. Patients were followed until September 2010.

2.2.3. Participants

A research nurse contacted patients who were newly enrolled on the surgical wait lists of the three hospitals weekly via telephone. The following inclusion criteria were employed: (1) age $\geq$ 40 years old; (2) scheduled for primary unilateral TKA; 3) understands, reads and speaks French. Patients were excluded from the study if they were suffering from a severe cardiac condition, a severe degenerative disease (other than OA) such as Parkinson's disease, Alzheimer's disease, any type of dystrophies or other type of medical condition potentially interfering with recovery after TKA or any severe mental disorder (severe depression, bipolar disorder, schizophrenia or dementia) that could impede the ability to answer the required questionnaires. Patients with a previous hip or knee arthroplasty were also excluded. Subjects who suffered major knee trauma in the previous year or who underwent urgent surgery within 30 days of inclusion on the waiting list were further excluded.

2.2.4. Data collection and variables

Structured 45 minute telephone interviews conducted by three trained interviewers and review of the subjects' medical files were employed in order to collect data. The interviews took place several days after enrolment on the wait lists (mean ± SD: 12.6 ± 4.7 days) and six months following the intervention (mean ± SD: 188.7 ± 5.4 days).

Dependent variables

Pain, stiffness and function at enrolment and six months after surgery were measured with the WOMAC. The WOMAC score was transformed in order to obtain a score that varied from 0 to 100, 0 indicating no pain, no functional limitations nor knee joint stiffness. In the light of a lack of consensus regarding the definition of a poor outcome following TKA, it was defined as the poorest quintile of the six-month postoperative WOMAC score (i.e. WOMAC score >40.4); an optimal outcome was defined as a WOMAC score in the best four quintiles of the distribution (i.e. score ≤ 40.4).

Independent variables

In order to develop a model with the highest possible validity, independent variables that were collected for the purpose of being considered as potential predictors included known important determinants of TKA outcomes reported in the literature (see Chapter 1). Variables were collected at the time of the patients' inclusion on the wait list and 6 months following TKA.

Potential predictors at enrolment on surgical wait list

Initial diagnosis, anthropometric data and comorbidities were extracted from the patients' medical files. The CIRS was employed in order to determine the burden of comorbidities [86]. During the initial interview, questions drawn from the questionnaire of the 1998 Quebec Health Survey were used to determine the level of education, the employment status, the household income and the level of social support [87]. The level of social support was also measured with questions from the Quebec Health Survey [87]. Marital status, household living status, and clinical variables such as duration of disease symptoms were also recorded at the time of the first interview. Psychological distress was documented by using a modified version of the Psychological Symptom Index (PSI). This version

includes 13 questions assessing depression and anxiety during the past week (range: 0-42) [88]. In an effort to minimize the number of questions to include in the final tool, individual items from the validated questionnaires (social support tool, PSI and WOMAC) were also employed as independent variables in order to develop the CPR.

Other variables

Several surgical variables such as type of implant, bearing type, implant fixation, patella resurfacing and the number and type of in-hospital complications (wound infection, dislocation, knee ankylosis and manipulation, cardiovascular/ pulmonary/ circulatory complications, peripheral/ central nervous system involvement, urinary infection, acute confusion, tendon and ligament rupture and blood transfusion) following TKA were documented by reviewing the patients' medical files. Hospital length of stay and discharge to a rehabilitation or recovery facility were also available through a review of the medical files. The pre-surgery wait times were calculated using the information in the wait list database of each hospital. Patients reported use of walking aid and the number of community physiotherapy treatment hours received since discharge from the hospital six months following the intervention.

2.2.5. Statistical analysis

The statistical methodology employed by this study was identical to the one described in section 2.1.5 in this chapter.

2.2.6. Ethics and confidentiality

Participants signed an informed consent form. The research protocol was approved annually by the Research ethics board of all three participating hospitals (CHUL, HSFA, HDQ).

CHAPTER 3

A systematic review of preoperative determinants of patient-reported pain and physical function up to two years following primary total hip arthroplasty

E. Lungu[1,2§], S. Maftoon [2], P-A. Vendittoli [2,3], F. Desmeules[2,4]

[1]Department of Biomedical Sciences, Faculty of Medicine, University of Montreal, Montreal, QC, Canada

[2]Centre de recherche de l'Hôpital Maisonneuve-Rosemont, 5415 Boul. L'Assomption, Montreal, Québec, Canada, H1T 2M4.

[3]Surgery Department, Maisonneuve-Rosemont Hospital, University of Montreal affiliated Research Center, 5415 Boul. L'Assomption, Montréal, Québec, Canada, H1T 2M4

[4]School of Rehabilitation, Faculty of Medicine, University of Montréal, Montreal, QC, Canada

Manuscript accepted for publication by *Orthopaedics & Traumatology: Surgery and Research* (see Appendix J)

1. Abstract

Background

A thorough understanding of the determinants of patient-reported pain and function following THA can help plan interventions directed at improving surgical results. Therefore, the objective was to summarize the preoperative determinants of pain and disability up to two years following THA in studies with appropriate methodological quality.

Methods

Four databases (Medline, Pubmed, Embase and CINAHL) were screened from their respective inception dates until April 2015 using a combination of keywords and MESH terms. Criteria for inclusion were 1- participants with primary unilateral THA followed up to 2 years, 2- validated disease-specific patient-reported outcome measures assessing pain and/or disability, 3-identification of determinants obtained via multivariate analyses. Methodological quality was assessed using a modified version of the *Methodology Checklist for Prognostic Studies*.

Results

Twenty-two manuscripts were included. Mean score of the methodological quality was 81.0±10.3 %. Among socioeconomic determinants, a lower educational level was significantly related to worse pain and function (three out of three studies evaluating the relationship). Clinical determinants of poor outcomes included preoperative levels of pain and physical function (nine out of 12 studies), higher BMI (six out of ten studies), presence/greater level of comorbidities (five out of seven), worse general health (four out of four studies) and lower radiographic OA severity (three out of four studies). Study

heterogeneity limited the pooled assessment of the strength of association between the preoperative variables and THA outcomes.

Discussion

Moderate-to-high quality of evidence allowed to elaborate a list of determinants of THA pain and function in the medium term with the greatest amount of evidence. This knowledge may assist the management of patients at risk of suboptimal results. Further research is required to clarify the force of association between determinants and THA outcomes.

Keywords: Total hip arthroplasty, Hip osteoarthritis, Determinants, Pain, Disability

2. Résumé

Introduction

Une bonne compréhension des déterminants de la douleur et des incapacités fonctionnelles suivant une arthroplastie totale de la hanche (ATH) peut aider à la prise en charge de ces patients et améliorer ultimement le succès de la chirurgie. L'objectif de cette étude était de synthétiser les déterminants de la douleur et des incapacités fonctionnelles jusqu'à deux ans suivant une ATH.

Méthodes

Une recherche bibliographiques dans quatre banques de données (Medline, Pubmed, Embase et CINAHL) a été réalisée jusqu'en avril 2015 en utilisant une combinaison de mots-clés et de termes MESH. Les critères d'inclusion étaient 1 – patients subissant une ATH primaire unilatérale avec d'au plus deux ans de suivi, 2 – emploi de questionnaires validés auto-rapportés évaluant la douleur et la fonction et 3 – identification dans ces études, des déterminants potentiels à l'aide d'analyses multivariées. La qualité méthodologique des différentes études a été évaluée à l'aide d'une version modifiée de *Methodology Checklist for Prognostic Studies*.

Résultats

Vingt-deux études ont été incluses. Le score méthodologique total moyen des études était de 81.0±10.3%. Parmi les facteurs psychosociaux identifiés dans les études incluses, un niveau de scolarité plus bas a été significativement associé avec des douleurs et des incapacités fonctionnelles plus sévères post-chirurgie (trois des trois études investiguant cette association). Les déterminants cliniques significativement associés avec des niveaux des douleurs et des incapacités post-arthroplastie incluaient les niveaux préopératoires de douleur et de fonction (neuf études sur 12), un indice de masse

corporelle plus élevé (six études sur dix), la totalité et/ou la sévérité des comorbidités (cinq études sur sept) et une arthrose radiographique plus sévère (trois études sur trois). Les différentes méthodes statistiques utilisées afin d'identifier les déterminants limitaient l'évaluation de la force des associations entre les différents déterminants et les douleurs et la fonction post-ATH.

Discussion

Des données probantes de qualité moyenne à haute suggèrent la présence de plusieurs déterminants associées à la douleur et à la fonction suivant l'ATH. L'identification de ces différents déterminant pourrait améliorer la prise en charge des patients à risque de résultats chirurgicaux sous-optimaux. Davantage de recherches sont nécessaires afin d'élucider formellement les déterminants significativement associés ainsi que la force de ces différentes associations avec la douleur et les incapacités suivant l'ATH.

Mots-clés : Arthroplastie totale de la hanche, Arthrose de la hanche, Déterminant, Douleur, Incapacité fonctionnelle

3. Introduction

Hip OA is a chronic debilitating condition, limiting the affected individuals in terms of functioning and causing important levels of physical pain [3, 89, 90]. Numerous pharmacological and non-pharmacological approaches aimed at relieving the ailments accompanying hip OA have been advocated. THA is currently the mainstay treatment in candidates experiencing important levels of pain and physical limitations who are unresponsive to other treatments [8]. Although generally effective in reducing pain and disability, THA may yield suboptimal results in up to 25% of patients [3]. A sound knowledge of determinants of THA results can ultimately provide an estimate of the likelihood of surgical success. Moreover, it can assist in the creation of an efficacious plan in order to improve outcomes.

Hence, there is a clear necessity of comprehensively summarizing the determinants of pain and function levels after THA with the greatest amount of high-quality evidence. This is more so motivated by the increasing amount of evidence published in the recent years on the topic, including the identification of novel determinants. An appropriate knowledge of determinants of THA outcomes can be achieved by reviewing studies employing validated and disease-specific PROMs of pain and function evaluating the independent effect of THA determinants by using multivariate analysis [91]. Identification of determinants in a timeframe up to two years following THA is particularly relevant, as patients are thoroughly monitored by their surgeons, and any outcome considered unsatisfactory can potentially be addressed promptly.

Consequently, the purpose of this systematic review was to identify the preoperative determinants of patient-reported pain and disability up to two years following primary unilateral THA for hip OA in studies with appropriate methodological quality.

4. Methods

4.1. Literature search and study identification

Four databases (Medline, Pubmed, Embase and CINAHL) were reviewed from their respective inception dates until April 2015 using a combination of keywords and MESH terms (see Appendix C for detailed search strategy). References of previously published reviews and relevant articles were scanned manually. In order to evaluate the eligibility, two authors independently reviewed the titles, abstracts and full texts of the articles.

4.2. Study selection

The eligibility of the studies was evaluated by considering the following criteria:

1. Participants underwent primary unilateral THA for hip OA
2. Results are presented for a follow-up of up to two years
3. The outcome measure was a disease-specific validated PROM assessing pain and/or function
4. Identification of determinants was obtained using multivariate analyses
5. Full-text article was published in English or French

4.3. Data extraction

The following characteristics were recorded using a standardized form: participants' age and gender proportion of the sample, number of patients, follow-up period, outcome measure, statistical methods

used and adjustments, as well as statistically significant and non-significant determinants reported by the study. Data from each article was recorded by one of the raters and verified by another.

4.4. Methodological quality appraisal

Two trained evaluators independently appraised the methodological quality of the studies. Subsequently, results were discussed to facilitate consensus. A third evaluator intervened in case of differences. The appraisal of the risk of bias was undertaken with a modified version of the *Methodology Checklist for Prognostic Studies* [74]. This tool evaluates the following characteristics: "Study participation", 'Study attrition", "Prognostic factor measurement", 'Outcome measurement", "Confounding measurement and account" and "Analysis". Each item assesses the risk of potential bias: low (0), unclear or unknown given the information available in the article (1) and high (2), with a maximal total score of 12, a higher score indicating a better methodological quality. If the included study was retrospective and no information regarding patients excluded from the study was provided, a score of 1 was automatically attributed to the "Study participation" item. A follow-up proportion inferior to 80% prompted the attribution of a score of 0 to the "Study attrition" item. Studies not accounting for either age, gender or BMI in their multivariate analysis received a score of 0 for the "Confounding measurement and account" item. Total methodological score was subsequently standardized.

4.5. Data synthesis

In their multivariate analyses, studies employed two approaches when defining the dependent variables: either pain and function were evaluated separately (such as the pain and the function subscale of the WOMAC) or as part of a combined construct (total WOMAC score). Therefore, data

on determinants was summarized according to both approaches. Solely a qualitative analysis was performed due the heterogeneity of included studies in terms of study designs, variables' constructs and definitions as well as statistical analyses.

5. Results

5.1. Description of the included studies

After exclusion of titles and abstracts, 129 full-text articles were further evaluated. One hundred and seven full-text articles were subsequently excluded, leaving 22 manuscripts published from 1997 to 2015 for inclusion (Figure 1). Details of study characteristics can be found in Table 1. Results from two studies are shown conjointly because of results based on the same cohort [52, 53]. The WOMAC was the most frequently employed validated tool (14 studies), followed by the HHS, the OHS and the LEFS (two studies each) and the HOOS used in one study.

5.2. Methodological quality

Details of the methodological quality of the included studies can be found in Table 2. Mean total score was 81.0% (SD: 10.3%), representing moderate-to-high methodological quality. No study received a score lower than 66.7% and four studies were graded higher than 90% [11, 89, 92, 93]. The "Study attrition" domain received the lowest mean score of 35.7% (SD 42.3%), with 11 studies having a follow-up proportion under 80%.

5.3. Preoperative determinants of pain and function levels following THA

5.3.1. Demographic determinants

Among the demographic variables that were investigated, older age was found to be associated with poor pain and functional outcomes in four studies [89, 93-95]. Kessler et al. showed that being older was associated with a lower odds of achieving a higher total WOMAC score three months postoperatively [93]. Quintana et al. found that in their cohort of 590 patients, an age of over 70 was associated with a smaller change in the WOMAC pain subscale score six months after the surgery, but did not find similar results for the two-year change in pain and in function [94]. Dowsey et al. showed that higher age was associated with worse function at one and two years postoperatively as measured with the Harris Hip Function score [89]. In addition, Stevens et al. showed that being older than 70 years old which was associated worse function at one year as per the function subscale of the WOMAC score [95]. Nevertheless, seven studies did not identify a significant association between age and THA outcomes [11, 45, 52, 53, 96-99].

Only two studies concluded to a significant relationship between gender and postoperative pain and function, and their findings are contradictory. Being female was associated with a higher change in the WOMAC pain score six months following THA in the first study [11], but to a lower total WOMAC score one year postoperatively in the other [95]. Seven studies were unable to establish such a significant association [11, 52, 53, 93, 94, 98-100].

The living arrangements was the sole other demographic determinant investigated in one study. Jones et al. showed that living alone was significantly associated with a lower change in the function subscale of the WOMAC score six months postoperatively [11].

5.3.2. Socioeconomic determinants

Among socioeconomic determinants of THA outcomes, a higher educational level was shown to be associated to lower levels of postoperative pain and to better functional status. Fortin et al. (1999 & 2002) report that a greater number of years of education was associated to lower levels of pain and better function six and 12 months after THA [52, 53]. Judge et al. (2010 & 2011) showed that patients with more education had a higher chance of being responders according to the OMERACT-OARSI criteria (Table 1), in addition to the greater likelihood of returning to a normal state (decrease in the total WOMAC score by 2 standard deviations compared to baseline) one year after THA.

Regarding the socioeconomic status, in their cohort of 1744 subjects followed for 18 months, Jenkins et al. found that a greater level of social deprivation predicted a poorer HHS 18 months after THA [101].

5.3.3. Psychosocial determinants

Only one study investigated psychosocial determinants of THA outcomes. Judge et al. (2011) found that having a greater number of expectations regarding functional, activity and pain levels following surgery were associated with higher odds of achieving the OMERACT-OARSI responder criteria and the MCID on the one-year function domain of the WOMAC score (Table 1), but not associated to the pain domain [98].

5.3.4. Clinical determinants

The associations between clinical variables and THA outcomes were the most studied relationships, with the greatest amount of evidence present for the preoperative levels of hip-related pain and

function. In the case where outcomes were measured as a function of postoperative state, worse levels of preoperative pain and function were associated with worse levels in the respective domains [11, 94, 102]. In contrast, studies that employed a change in status as a dependent variable showed that better preoperative levels of pain and function were associated with smaller changes [52, 53, 99, 103-105]. Only two studies showed no significant associations between preoperative and postoperative pain and function [45, 96].

A greater BMI at the time of surgery was associated with worse THA results in terms of pain and function. In a cohort of 707 THA patients, a BMI superior to 30 was associated with lower total WOMAC scores at one year [92]. Similar results were found for subjects having a BMI greater than 25 in another study [95] and comparable results have been observed by four other studies [11, 89, 96, 97], while two studies report a non-significant association between BMI and THA outcomes [45, 93].

Regardless of the method of measurement of comorbidity, i.e. either the presence of a specific medical condition or the number of concomitant disorders, comorbidities were significantly associated with worse pain and function following THA in a consistent manner. Peter et al. showed that a greater number of preoperative comorbidities were associated with worse HOOS pain and physical functioning scores up to 22 months after THA [106]. Moreover, the same study showed that, arteriosclerosis, cardiac disorders, dizziness in combination with episodes of falling, asthma, chronic obstructive pulmonary disease and cancer were all associated with worse pain and functioning. Four other studies concluded to similar findings regarding the association between greater comorbidity and poor THA pain and functional outcomes [11, 92, 95, 99], while two others found no association between comorbidity and THA results [94, 103].

Back pain and contralateral hip OA could be considered special instances of comorbidity and were evaluated in several studies. Presence of preoperative back pain was associated with a smaller change on the pain subscale of the two-year WOMAC score, and to a smaller change in function at six months and two years postoperatively [94]. Similarly, severe back pain was associated to poor HOOS pain and functional scores seven to 22 months postoperatively [106]. Regarding contralateral hip involvement, two studies found it to be associated to poor pain and function outcomes after THA [11, 94].

A worse general health level, as measured by the SF-36 and the SF-12 questionnaires, was associated with poor outcomes after THA in all the four studies that evaluated such an association [11, 89, 94, 100]. According to Dowsey et al, better physical and mental health are both independent determinants of hip pain and function, and were associated with better Harris Hip pain and function Scores at one and two year follow-ups [89].

A lower radiographic OA severity was associated to poor outcomes in three studies [89, 98, 100]. Judge et al. (2011) found that a Kellgren-Lawrence grade of 1, 2 or 3 was associated with a lower odds of being a responder according to the OMERACT-OARSI responder criteria (Table 1) 1 year postoperatively when compared to a grade of 4 [100].

Table 3 presents other clinical variables that were found significantly associated with pain and function after THA, including greater widespread pain sensitivity [107], radiographical medial concentric disease [89] and lower knee extensor muscles strength [90].

5.3.5. Surgical determinants

Surgical determinants of THA outcomes received modest attention. Braeken et al. investigated the type of implant fixation and found that having a cemented prosthesis was associated with a worse six to 12 month level of pain as measured by the WOMAC scale [96]. Jones et al. found that a cementless prosthesis was related to a lower change in WOMAC pain scores six months postoperatively [11]. Dowsey et al. showed that a greater femoral head size was significantly associated with worse functional outcomes, but not with pain levels [89].

5.3.6. Healthcare-related determinants

One study investigated the relationship between waiting time and THA outcomes. Vergara et al. showed that a waiting time longer than six months for THA was associated to lower gains in function one year after the surgery when compared to a waiting time shorter than three months [102].

6. Discussion

6.1. Highlight points

A proper understanding of the determinants of THA outcomes could lead to improved results in terms of pain and functional status outcomes in the significant proportion of patients experiencing suboptimal results following this intervention. The objective of our study was to systematically review the literature of all the studied determinants of patient-reported, disease-specific pain and functional limitation following primary unilateral THA in patients with hip OA. Twenty-two prognostic studies identified demographic, psychosocial, clinical, surgical and healthcare-related determinants with a moderate-to-high methodological quality and allowed the elaboration of a list of determinants with the highest amount of available evidence (Table 4).

6.2. Strengths and limitations of the review

The strictness of the eligibility criteria allowed for the inclusion of studies with the highest methodological quality from four important databases. Focus on all variables investigated for a possible association with THA outcomes permitted the elaboration of a comprehensive list of determinants with the highest level of evidence to date.

In terms of limitations, study heterogeneity reduced the ability to pool results in order to evaluate the strength of association between significant determinants and THA outcomes. Study findings do not apply to patients undergoing bilateral and revision THA, nor can be extended to determinants of long-term outcomes. Moreover, regardless of the sound statistical methods employed to identify determinants of THA results in the included studies, the extent of the clinical and practical significance of the determinants remains elusive.

6.3. Main findings

In contrast with previously published systematic reviews on the same subject [10, 18], the findings of the current study do not suggest a significant association between demographic variables such as age and gender with THA outcomes. Although moderate level evidence indicates that older age may be associated with worse pain and function, a non-negligible number of included studies point to no significant association. Regarding gender, the two studies that found a significant association with THA outcomes have contradictory results for males and females.

In terms of psychosocial variables, a lower educational level is associated with poor outcomes following THA in all the studies (n=4) where it was investigated. Although the reasons for such an

association are not directly discussed by the included studies, it is likely that the level of education is related to the patients' socioeconomic status, which has been consistently associated with outcomes of other musculoskeletal conditions [108, 109].

Clinical factors received the greatest amount of attention. The preoperative levels of pain and function are the determinants with the highest amount of evidence in the current review. The direction of the association depends on the approach employed for the evaluation of surgical results. Indeed, a lower preoperative status of pain and function is associated with a lower postoperative status, but with a higher change in these domains. Because the amount of pain and the extent of disability are often indications of proceeding with THA, this finding highlights the paradigm where patients with worse preoperative status have larger gains, but generally do not achieve the same levels as their counterparts who underwent THA earlier in the disease process. Unfortunately, there does not seem to be a clinical consensus regarding the necessity and timing for performing the surgery in patients with worse or better preoperative state [110].

BMI was significantly associated with worse THA outcomes in six out of ten studies in our review. Some of these studies emphasize that the clinical significance of such a relationship may however be limited, as the individual weight of other variables such as preoperative levels of pain, function or comorbidities is more substantial [95, 97]. For example, Stevens et al. report that a when compared to a BMI smaller than 25, a BMI greater than 25 was significantly associated with a poorer one-year total WOMAC score with a multivariate regression coefficient of -0.63, $p<0.001$, whereas having more than two comorbidities had a coefficient of -14.5, $p<0.001$.

A greater level of comorbidity and worse general health are somewhat related clinical factors that were found to have sizeable evidence. Traditionally, the level of comorbidity is evaluated preoperatively using the American Society for Anaesthesiology Scale in order to assess the risk of complications associated with the surgery [106]. The link between other medical conditions as well as general health with THA outcomes is increasingly recognized, and some authors suggest that addressing these before undergoing the surgery may be indicated [95].

A lower radiographic OA severity has been associated with worse changes in pain and function in three studies. Although the relationship between radiographic severity with preoperative hip pain and function is inconsistent, this finding may parallel the association between higher preoperative levels of pain and disability with lower changes in status after the surgery [111].

Some studies focused on evaluating determinants of THA for either pain or function independently, whereas others assessed the associations between preoperative variables and pain and function as part of a combined measure, such as in the case of the total WOMAC score. We attempted to compare whether the two approaches yielded different results in terms of the identity of the determinants. On occasion, individual studies report different determinants according to the method of outcome assessment; however, the results are ultimately similar when viewing the overall picture (Table 3).

7. Conclusion

Studies with a moderate-to-high methodological quality indicate that a lower educational status, worse or better preoperative levels of pain and function, greater BMI, more comorbidity and worse general health as well as a lower radiographical OA severity are significant determinants of pain and function

up to two years following primary unilateral THA with the highest amount of evidence. Knowledge of these determinants could aid the clinician and the patients in assessing the risks and benefits associated with the procedure. Moreover, interventions targeted at diminishing the risk profile of patients undergoing THA in order to potentially ameliorate their outcomes could be developed based on these determinants. More standardized approaches of future studies evaluating determinants of pain and function following THA could diminish the heterogeneity associated with the results, and improve the likelihood of establishing the strengths of the association between variables. Evaluating the extent of the clinical applicability of the relationship between determinants and THA outcomes should also be targeted in future studies.

8. Acknowledgements

EL is supported by a training award for M.D.-M.Sc. students from the FRQS.

9. Conflict of interest: none

Figure 1 Flowchart of the literature search

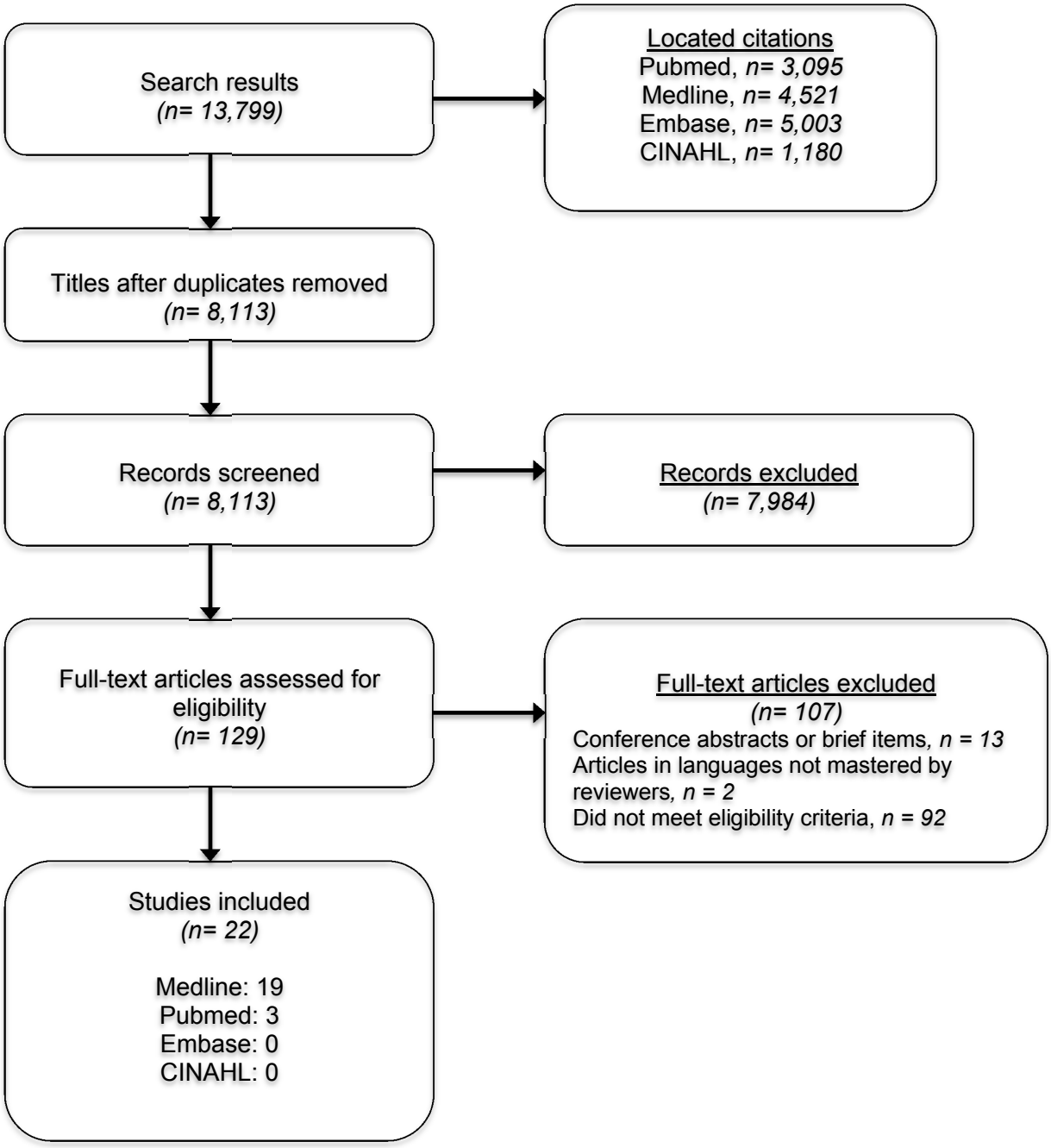


Table 1 Description of the included studies

Study	Participants				Number of patients	Follow-up period	Outcome Measure	Statistical method	Statistical adjustment	Results	
	Diagnosis	Type of Surgery	Mean Age(SD)	Gender (%female)						Significant Preoperative Determinants	Non-significant Preoperative Determinants
Braeken et al. (1997)	OA	Primary THA	63.5 (12.8)	61	193	6 months to 1 year	WOMAC pain and function scores	Stepwise multiple regression	None	**WOMAC pain** **Greater BMI (-)**: β = 0.077, p = 0.0001 **Greater preoperative pain relief (+)**: β = - 0.126, p = 0.0002 **Cemented prosthesis (-)**: β = 0.521, p = 0.0071 **WOMAC function** **Greater preoperative pain relief (+)**: β = - 0.221, p = 0.0001 **Greater BMI (-)**: β = 0.096, p = 0.0002 **Greater physical therapy intensity (+)**: β = -1.558, p = 0.0345	**WOMAC pain** **Preoperative pain** **Hybrid prosthesis** **WOMAC function** **Age**
Caracciolo et al. (2005)	OA	Primary THA	69.1 (8.1)	72	39	6 months	WOMAC function score	Logistic regression	None	**Higher preoperative function (+)**, OR = 1.44, 95% CI 1.09 to 1.92 compared to lower preoperative function (worst quartile of WOMAC function score)	**Preoperative osteoarthritis morbidity**
Dowsey et al. (2014)	Primary OA	Primary THA	68.9 (9.3)	60.7	382	1 and 2 years	Harris hip pain and function scores at 1 year and 2 years Minimum difference (MID) in pain and function scores at 1 or 2 years: 0.5*standard deviation of the mean change between preoperative and postoperative scores	Multivariate linear regression Multivariate logistic regression	None	**Harris Hip Pain score at 1 year** **Better physical health (+)**: β = 0.36, 95% CI 0.16 to 0.55, p < 0.001 **Better mental health (+)**: β = 0.16, 95% CI 0.07 to 0.26, p = 0.001 **Worse modified K-L grade (+)**: β = 1.12, 95% CI 0.13 to 2.12, p = 0.028 **Medial concentric (vs. supero-lateral) (+)**: β = 2.58, 95% CI 0.70 to 4.46, p = 0.07 **Harris Hip Pain score at 2 years** **Better physical health (+)**: β = 0.36, 95% CI 0.16 to 0.55, p < 0.001 **Better mental health (+)**: β = 0.16, 95% CI 0.07 to 0.26, p = 0.001 **Medial concentric (vs. supero-lateral) (+)**: β = 2.07, 95% CI 0.17 to 3.98, p = 0.07 **Harris Hip Function score at 1 year** **Higher age (-)**: β = -0.25, 95% CI -0.39 to -0.10, p = 0.001 **Higher BMI (-)**: β = -0.32, 95% CI -0.49 to -0.14, p = 0.001 **Higher ASA score (-)**: β = -1.99, 95% CI -3.74 to -0.24, p = 0.026 **Higher preoperative function (+)**: β = 0.35, 95% CI 0.23 to 0.48, p < 0.001 **Better physical health (+)**: β = 0.27, 95% CI 0.07 to 0.47, p = 0.01 **Better mental health (+)**: β = 0.20, 95% CI 0.10 to 0.30, p < 0.001 **Surgical approach (unclear)**: β = 3.71, 95% CI 1.67 to 5.75, p < 0.001	**Harris Hip Pain score at 1 year** **Age** **Gender** **BMI** **ASA score** **Preoperative hip pain** **Preoperative hip function** **Surgical approach** **Cementation** **Femoral head size** **Harris Hip Pain score at 2 years** **Age** **Gender** **BMI** **ASA score** **Preoperative hip pain** **Preoperative hip function** **Surgical approach** **Cementation** **Femoral head size** **Modified K-L grade** **Harris Hip Function score at 1 year** **Gender** **Preoperative hip pain** **Cementation** **Hypertrophic/atrophic radiographic disease pattern** **Chondrocalcinosis**

Greater femoral head size (-): β = -2.03, 95% CI -4.02 to -0.03, p = 0.047
Worse modified K-L grade (+): β = 1.43, 95% CI 0.34 to 2.52, p = 0.01

Harris Hip Function score at 2 years
Higher age (-): β = -0.33, 95% CI -0.49 to -0.17, p < 0.001
Higher BMI (-): β = -0.23, 95% CI -0.43 to -0.04, p = 0.021
Higher ASA score (-): β = -2.04, 95% CI -4.00 to -0.09, p = 0.04
Higher preoperative function (+): β = 0.34, 95% CI 0.20 to 0.49, p < 0.001
Better physical health (+): β = 0.31, 95% CI 0.09 to 0.53, p = 0.007
Better mental health (+): β = 0.23, 95% CI 0.11 to 0.34, p < 0.001
Surgical approach (unclear): β = 4.83, 95% CI 2.54 to 7.12, p < 0.001
Worse modified K-L grade (+): β = 1.64, 95% CI 0.44 to 2.85, p = 0.008

Harris Hip Function score at 2 years
Gender
Preoperative hip pain
Cementation
Femoral head size
Hypertrophic/atrophic radiographic disease pattern
Chondrocalcinosis

MID in Harris Hip Pain score at 1 year
Modified K-L ≤ 3a (vs. Modified K-L 4b) **(-)**: OR = 0.03, 95% CI 0.00 to 0.35

MID in Harris Hip Pain score at 1 year
Age
Gender
BMI
ASA score
Preoperative physical health
Preoperative mental health
Posterior vs. Hardinge approach
Cementation vs. uncemented
Femoral head > 28mm
Modiefied K-L 3b
Modified K-L 4a
Medial-concentric OA vs. supero-lateral

MID in Harris Hip Pain score at 2 years
Modified K-L ≤ 3a (vs. Modified K-L 4b) **(-)**: OR = 0.04, 95% CI 0.00 to 0.49

MID in Harris Hip Pain score at 2 years
Age
Gender
BMI
ASA score
Preoperative physical health
Preoperative mental health
Posterior vs. Hardinge approach
Cementation vs. uncemented
Femoral head > 28mm
Modified K-L 3b
Modified K-L 4a
Medial-concentric OA vs. supero-lateral

MID in Harris Hip Function score at 1 year
Modified K-L ≤ 3a (vs. Modified K-L 4b) **(-)**: OR = 0.22, 95% CI 0.08 to 0.64
Modified K-L 3b (vs. Modified K-L 4b) **(-)**: OR = 0.36, 95% CI 0.17 to 0.75
Modified K-L 4a (vs. Modified K-L 4b) **(-)**: OR = 0.30, 95% CI 0.14 to 0.64

MID in Harris Hip Function score at 1 year
Age
Gender
BMI
ASA score
Preoperative physical health
Preoperative mental health
Posterior vs. Hardinge approach
Cementation vs. uncemented
Femoral head > 28mm

MID in Harris Hip Function score at 2 years
Advanced age (-): OR = 0.96, 95% CI 0.92 to 1.00
Modified K-L ≤ 3a (vs. Modified K-L 4b) **(-)**: OR = 0.13, 95% CI 0.05 to 0.35
Modified K-L 3b (vs. Modified K-L 4b) **(-)**: OR =

MID in Harris Hip Function score at 2 years
Gender
BMI
ASA score
Preoperative physical health

										0.26, 95% CI 0.12 to 0.54 **Modified K-L 4a** (vs. Modified K-L 4b) **(-)**: OR = 0.37, 95% CI 0.17 to 0.80	**Preoperative mental health** **Posterior vs. Hardinge approach** **Cementation vs. uncemented** **Femoral head > 28mm**
Fortin et al. (1999) and Fortin et al. (2002)	OA	Primary THA	68 (9.1)	56	116 and 84	6 months and 2 years	WOMAC pain and function scores	Multiple linear regression	None	**<u>WOMAC pain at 6 months</u>** **Higher preoperative pain (-):** β = 0.27 ± 0.08 **Higher education (+)**: β = -0.18 ± 0.09 r^2 = 0.20 **<u>WOMAC function at 6 months</u>** **Higher preoperative function (+):** β = 0.36 ± 0.08 **Higher education (+)**: β = -0.41 ± 0.28 r^2 = 0.24 Similar results for (2002); data not shown	**Age** **Gender** **Center**
Gandhi et al. (2010)	Primary and secondary OA	Primary unilateral THA	66.5	77.4 %	707	1 year	WOMAC total score at 1 year	Linear regression modelling	Age Gender Baseline total WOMAC scores Comorbidity	**2 metabolic syndrome risk factors (-)**: β = 3.1, 95% CI 0.3 to 5.1 **4 metabolic syndrome risk factors (-)**: β = 15.0, 95% CI 1.4 to 28. **Hypertension (-):** β = 7.3, 95% CI 2.4 to 13.2 **Obesity (BMI > 30) (-)**: β = 2.4, 95% CI 1.4 to 4.2	**3 metabolic syndrome risk factors**: β = 0.4, 95% CI -6.1 to 5.3 r^2 for the number of metabolic syndrome risk factors model= 0.32 **Hypercholesterolemia**: β = 4.5, 95% CI -0.2 to 4.9 **Diabetes**: β = -2.3, 95% CI -8.3 to 3.7 r^2 for the individual metabolic syndrome factors model= 0.41
Garbuz et al. (2006)	OA	Primary THA	65	56.4	147	1 year	Better than expected WOMAC pain and function outcomes WOMAC pain and function scores	Logistic regression modelling	Age Gender Comorbidity	**<u>Better than expected WOMAC function</u>** **Each month of pre-surgical waiting time (-)**: OR = 0.92, p = 0.05 **WOMAC pain** ‘‘[…] the postoperative WOMAC pain score will increase by 30% when the preoperative WOMAC pain score increases by three points (pain score range, 0–20; p = 0.0036).‘‘ **WOMAC function** ‘‘The postoperative WOMAC functional score will increase by 35% when the preoperative WOMAC functional score increases by 10 points (functional score range, 0–68; p = 0.0005).‘‘	**<u>Better than expected WOMAC pain</u>** **Pre-surgical waiting time**
Holstege et al. (2011)	OA	Primary unilateral THA	72.1 (6.8)	74.5	37	12 weeks	WOMAC function score	Multiple linear regression	Age Sex Baseline WOMAC function score BMI	**Higher knee extensors strength (+)**: β = -0.105, p = 0.004	**Hip flexors strength** **Hip extensors strength** **Hip abductors strength** **Hip adductors strength** **Knee flexors strength**

Jones et al. (2001)	OA (93%)	Primary unilateral THA	71.4	61.2	197	6 months	Change in WOMAC pain and function scores at 6 months	Multiple linear regression models	None	**<u>Change in WOMAC pain:</u>** **Female gender (+)**: β = 6.49, 95% CI 1.10 to 11.91 **Higher preoperative bodily pain (SF-36) (-)**: β = -0.59, 95% CI -0.7 to -0.41 **Higher number of comorbid conditions (-)**: β = -1.67, 95% CI -2.83 to -0.50 **Cementless prosthesis (-):** β = -7.08, 95% CI -14.15 to -0.01 **<u>Change in WOMAC function:</u>** **Higher preoperative joint pain (WOMAC) (-)**: β = -0.59, 95% CI -0.76 to -0.41 **Higher BMI (-)**: β = -0.68, 95% CI -1.20 to -0.17 **Contralateral hip involvement (-)**: β = -6.83, 95% CI -11.00 to -2.65 **Living alone (-)**: β = -10.01, 95% CI -15.92 to -4.10 **Higher number of comorbid conditions (-)**: β = -2.06, 95% CI -3.28 to -0.84	**<u>Change in WOMAC pain:</u>** **Age** **Waiting time** **Length of stay** **<u>Change in WOMAC function:</u>** **Length of stay** **Age** **Gender** **Waiting time** **Preoperative bodily pain (SF-36)**
Jenkins et al. (2009)	OA	Unilateral THA	68.2 (10.2)	62.8	1744	18 months	Harris Hip Score at 18 monhts	Multivariate linear regression	Age Sex	**More socioeconomic deprivation (-)**: B = -0.63, p = 0.004	**None**
Judge et al. (2010)	OA	Primary THA	N/A	56.2	908	1 year	**Return to normal**: 2 SD below mean baseline WOMAC total scores **OMERACT-OARSI response criteria** Based on WOMAC total score **Minimally Important Difference**: 0.5*mean difference between 1 year WOMAC total score and baseline WOMAC total score	Logitstic regression analysis	None	**<u>Return to normal:</u>** **University degree** (vs. no education) **(+)**: OR = 2.9, 95% CI 1.4 to 5.9 **College diploma or equivalent** (vs. no education) **(+)**: OR = 2.1, 95% CI 1.3 to 3.4 **ASA status 1** (vs. ASA status 2) **(+)**: OR = 2.5, 95% CI 1.2 to 5.3 **<u>OMERACT-OARSI responder criteria:</u>** **University degree** (vs. No education) **(+)**: OR = 14.4, 95% CI 1.8 to 117.8 **<u>Minimally Important Difference</u>**: **K/L grade 3** (vs. K/L grade 4) **(-)**:OR = 0.5, 95% CI 0.4 to 0.8	**<u>Return to normal:</u>** **Age** **Sex** **Obesity** **Employment status** **Post-graduate degree** **ASA status 3** **ASA status 4** **K/L grade** **<u>OMERACT-OARSI responder criteria:</u>** **Age** **Sex** **Obesity** **Employment status** **College diploma or equivalent** **Post-graduate degree** **ASA status** **K/L grade** **<u>Minimally Important Difference</u>**: **Age** **Sex** **Obesity** **Employment status** **Education level** **ASA status** **K/L grade 1** **K/L grade 2**

Judge et al. (2011)	OA	Primary THA	65.9 (10.9)	56.2	1025 (complete data for 908 patients, 117 patients, data were imputated using multiple imputation)	1 year	**OMERACT-OARSI response criteria** Based on total WOMAC total score **Minimally Important Difference**: 0.5*mean difference between 1 year WOMAC pain score, function score and total score and baseline WOMAC total score	Logitstic regression analysis	**Minimally Important Difference**: Age Sex School education ASA grade K&L grade Obesity Medication use Pre-op WOMAC score Pre-op EQ5D score	**OMERACT-OARSI response criteria** **Higher number of expectations (+):** OR = 1.36, 95% CI 1.07 to 1.73 **University degree** (vs. No education) **(+)**: OR = 6.24, 95% CI 2.76 to 14.13 **College diploma or equivalent** (vs. no education) **(+)**: OR = 1.80, 95% CI 1.08 to 2.99 **K/L grade 0, 2, 3** (vs. K/L grade 4) **(-)**: OR = 0.61, 95% CI 0.38 to 0.95 **Minimally Important Difference** **WOMAC function** **Higher number of expectations (+):** OR = 1.20, 95% CI 1.09 to 1.32	**OMERACT-OARSI response criteria** **Age** **Sex** **ASA status** **BMI** **Number of medications** **Pre-op WOMAC total score** **Pre-op EQ5D score** **Minimally Important Difference** **WOMAC pain** **Number of expectations**
Judge et al. (2014)	OA	Primary THA	68.9	60.8	4413	1 year	Oxford Hip Score at 12 months	Multiple linear regression	Age Sex	**Higher BMI (-)**: 5-unit increase in BMI, the 12-month OHS decreases by 0.84 95%CI (0.59 to 1.08) points	**None**
Kennedy et al. (2011)	OA	Primary THA	61	42.6	75	1 year	LEFS at 1 year	Multiple linear regression	None	**None**	**Age** **BMI** **Preoperative LEFS**
Kessler et al. (2007)	OA	Primary THA	63.6	44.8	67	3 months	WOMAC total score	Multiple regression analysis	None	**Higher age (-)**: OR = 0.43, 95% CI 0.04 to 0.82	**Sex** **Affected side** **BMI** **Anchorage** **Duration of surgery** **Hospital length of stay** **WOMAC preoperatively**
Peter et al. (2014)	OA	Primary THA	69.8 (9.5)	64.0	281	7-22 months	HOOS pain and physical functioning scores	Multivariate regression analysis	Age Sex BMI	**HOOS pain score** **Presence of 3 and 4 comorbidities (-)**: β = -12.8, 95% CI -20.9 to -4.7 **Presence of 5 or more comorbidities (-)**: β = -15.9, 95% CI -25.3 to -6.4 **HOOS physical functioning score** **Presence of 1 and 2 comorbidities (-)**: β = -8.6, 95% CI -16.4 to -0.9 **Presence of 3 and 4 comorbidities (-)**: β = -14.9, 95% CI -23.8 to -6.1 **Presence of 5 or more comorbidities (-)**: β = -22.1, 95% CI -32.4 0to -11.7 **HOOS pain score** **Arteriosclerosis in abdomen or legs (-)**: β = -12.3, 95% CI -21.7 to -2.9 **Dizziness in combination with falling (-)**: β = -12.1, 95% CI -23.5 to -0.7 **Cardiac disorders (-)**: β = -10.6, 95% CI -17.6 to -3.6 **Severe back pain (-)**: β = -8.0, 95% CI -14.0 to -2.0 **Cancer (-)**: β = -6.7, 95% CI -12.6 to -0.7 R^2 = 0.147 **HOOS physical functioning score** **Arteriosclerosis in abdomen or legs (-)**: β = -14.9, 95% CI -25.3 to -4.5 **Dizziness in combination with falling (-)**: β = -	**HOOS pain score** **Presence of 1 and 2 comorbidities** **HOOS physical functioning score** None **HOOS pain score** None **HOOS physical functioning score** None

										22.0, 95% CI -34.2 to -9.9 **Asthma or COPD (-)**: β = -9.0, 95% CI -16.8 to -1.3 **Severe back pain (-)**: β = -12.9, 95% CI -19.3 to -6.4 **Cancer (-)**: β = -8.0, 95% CI -14.3 to -1.7 R^2 = 0.2015	
Quintana et al. (2009)	OA	THA	69.3 (8.5)	48.6	590	6 months and 2 years	WOMAC pain and function change scores	Multiple regression analysis	None	6 months **<u>WOMAC pain change score</u>** **Higher pre-intervention pain (+)**: β = -8.85, p < 0.001 **Better mental health (+)**: β = - 0.32, p < 0.001 **Age > 70 years (-)**: β = - 2.93, p = 0.04 **Contralateral hip OA (-)**: β = 3.95, p = 0.006 **<u>WOMAC function change score</u>** **Higher pre-intervention functional limitation (+)**: β = - 8.23, p < 0.001 **Better mental health (+)**: β = - 0.44, p < 0.002 **Contralateral hip OA (-)**: β = 6.39, p < 0.001 2 years **<u>WOMAC pain change score</u>** **Higher pre-intervention pain (-)**: β = 8.68, p < 0.001 **Better mental health (-)**: β = 0.71, p < 0.001 **Back pain (-)**: β = -5.32, p = 0.004 **<u>WOMAC function change score</u>** **Higher pre-intervention functional limitation (-)**: β = 8.37, p < 0.001 **Better mental health (-)**: β = 0.75, p < 0.001 **Contralateral hip OA (-)**: β = -5.62, p = 0.01 **Back pain (-)**: β = -7.23, p = 0.004	6 months **<u>WOMAC pain change score</u>** **Gender** **Burden of comorbidities (Charlson Comorbidity Index)** **Back pain** **<u>WOMAC function change score</u>** **Age** **Burden of comorbidities (Charlson Comorbidity Index)** 2 years **<u>WOMAC pain change score</u>** **Age** **Gender** **Contralateral hip OA** **Burden of comorbidities (Charlson Comorbidity Index)** **<u>WOMAC function change score</u>** **Age** **Gender** **Burden of comorbidities (Charlson Comorbidity Index)**
Ramaesh et al. (2013)	Arthrosis	THA	67.1	55	184	1 year	12-month Oxford Hip Score	Multiple regression analysis	None	**More comorbidity (-)**: B = -1.77, 95% CI -2.35 to -1.19 **Higher preoperative pain/function (+)**: B = 0.26, 95% CI 0.10 to 0.43	**Age** **Gender** **Personality type**
Slaven et al. (2012)	OA	THA	68.2 (8.2)	55	37	6 months	Increase in LEFS score by ≥9 points or ability to walk without and assistive device	Logistic regression analysis	None	**None**	**Walking speed** **Timed-Up and Go test score** **Functional Reach test score**
Stevens et al. (2012)	Primary OA	Primary THA	70.3 (8.2)	74.2	653	1 year	WOMAC total score	Linear regression analysis	Postoperative complications	**BMI >25 (-)**: Estimate = -0.63, p < 0.001 **Age > 70 years (-)**: Estimate = -0.18, p = 0.04 **Female (-)**: Estimate = -4.41, p = 0.01 **1 or 2 comorbidities (-)**: Estimate = -7.25, p = 0.01 **> 2 comorbidities (-)**: Estimate = -14.50	**None**

Wylde et al. (2015)	OA	Primary unilateral THR	66.5 (10)	58.7	254	12 months	WOMAC pain score at 12 months Change in WOMAC pain score at 12 months	Linear regression Linear mixed model	WOMAC pain score at 12 months: **Model 1**: gender, randomization (intraoperative local anesthetic wound infiltration and spinal anesthetic vs. spinal anesthetic alone) **Model 2**: Model 1 + age, height, weight, education, cohabitation, employment **Model 3**: Model 2 + preoperative pain Change in WOMAC pain score at 12 months: gender, age, height, weight, randomization, cohabitation, employment, education	**WOMAC pain score at 12 months** **Model 1:** **Greater widespread pain sensitivity (-)**: β = -0.110, 95% CI -0.193 to -0.027 **Model 2:** **Greater widespread pain sensitivity (-)**: β = -0.104, 95% CI -0.187 to -0.020 **Model 3:** **Greater widespread pain sensitivity (-)**: β = -0.091, 95% CI -0.176 to -0.006 **Change in WOMAC pain score at 12 months** None	**WOMAC pain score at 12 months** **Model 1:** None **Model 2:** None **Model 3:** None **Change in WOMAC pain score at 12 months** **Widespread pain sensitivity**
Vergara et al. (2011)	OA	Primary THA	69.0 (8.9)	49.2	527	1 year	Change in WOMAC function score Change in WOMAC function score ≥30/100	Multivariate linear regression Logistic regression	None	**Multivariate linear regression** **Waiting time > 6 months vs. <3 months (-)**: β = -4.79, p = 0.025 **Higher pre-intervention function (+)**: β = 0.67, p < 0.001 **Logistic regression** **Waiting time > 6 months vs. <3 months (-)**: OR = 0.47, 95% CI 0.27 to 0.81 **Higher pre-intervention function (+)**: OR = 1.06, 95% CI 1.04 to 1.07	**Multivariate linear regression** **Waiting time 3-6 months vs.<3 months** **Logistic regression** **Waiting time 3-6 months vs.<3 months**

* OMERACT-OARSI responder criteria - improvement in WOMAC pain or in function of 50% and an absolute change of 20/100, or, if the patient does not fulfill them, improvement in 2 of the 3 following: pain of 20% and an absolute change of 10/100, function of 20% and an absolute change of 10/100, and patient's global assessment (total WOMAC score) of 20% and an absolute change of 10/100.

(+): determinant of successful outcome; (-): determinant of poor outcome

ASA –American Society of Anaesthesiologists; β – regression coefficient; K-L – Kellgren-Lawrence; OR – Odds Ratio; r^2 – coefficient of determination of the linear regression model; SD – standard deviation

Table 2 Methodological appraisal of the included studies

Included studies (n= 22)	Study participation	Study attrition	Prognostic factor measurement	Outcome measurement	Confounding measurement and account	Analysis	Total score /12
Braeken et al. [96]	2	2	2	2	0	2	10
Caracciolo et al. [103]	2	1	2	2	0	2	9
Dowsey et al. [89]	2	2	2	2	2	2	12
Fortin et al. (1999 & 2002) [52, 53]	2	0	2	2	0	2	8
Gandhi et al. [92]	2	1	2	2	2	2	11
Garbuz et al. [105]	2	0	2	2	0	2	8
Holstege et al. [90]	2	0	2	2	2	2	10
Jenkins et al. [101]	1	1	2	2	0	2	8
Jones et al. [11]	2	1	2	2	2	2	11
Judge et al. (2010) [100]	2	0	2	2	2	2	10
Judge et al. (2011) [98]	2	0	2	2	2	2	10
Judge et al. (2014) [97]	2	0	2	2	2	2	10
Kennedy et al. [45]	2	1	2	2	0	2	9
Kessler et al. [93]	2	2	2	2	2	2	12
Peter et al. [106]	2	0	2	2	2	2	10
Quintana et al. [94]	2	0	2	2	0	2	8
Ramaesh et al. [99]	2	2	2	2	0	2	10
Slaven et al. [69]	2	2	2	2	0	2	10
Stevens et al. [95]	2	0	2	2	2	2	10
Vergara et al. [102]	2	0	2	2	0	2	8
Wylde et al. [107]	2	0	2	2	2	2	10

TOTAL (mean ± SD) /12	1.95 (0.22)	0.71 (0.85)	2.00 (0.0)	2.00 (0.0)	1.05 (1.02)	2.00 (0.0)	9.71 (1.23)
TOTAL (mean ± SD) /100	**97.6 (10.9)**	**35.7 (42.3)**	**100 (0.0)**	**100 ± 0.0**	**52.4 (51.2)**	**100 (0.0)**	**81.0 (10.3)**

Study participation: The study sample represents the population of interest with regard to key characteristics, sufficient to limit potential bias to the results
Study attrition: Loss to follow-up is unrelated to key characteristics (that is, the study data adequately represent the sample), sufficient to limit potential bias
Prognostic factor measurement: The prognostic factor of interest is adequately measured in study participants, sufficient to limit potential bias
Outcome measurement: The outcome of interest is adequately measured in study participants, sufficient to limit bias
Confounding measurement and account: Important potential confounders are appropriately accounted for, limiting potential bias with respect to the prognostic factor of interest
Analysis: The statistical analysis is appropriate for the design of the study, limiting potential for the presentation of invalid results
SD: standard deviation

Table 3 Significant determinants of poor outcome 12 weeks to 2 years following THA

Determinant type	**PAIN**	*Studies*	**FUNCTION**	*Studies*	**PAIN & FUNCTION combined**	*Studies*
Demographic	**Older age** **Male gender**	[94] [11]	**Older age** **Living alone**	[89] [11]	**Older age** **Female gender**	[93, 95] [95]
Socioeconomic	**Lower educational level**	[52, 53]	**Lower educational level**	[52, 53]	**Lower educational level** **More socioeconomic deprivation**	[98, 100] [101]
Psychosocial	NONE	NONE	**Fewer surgery expectations**	[98]	**Fewer surgery expectations**	[98]
Clinical	**Worse preoperative pain** **Better preoperative pain** **Worse bodily pain** **Higher BMI** **More comorbidity** **Worse mental health** **Worse physical health** **Presence of back pain** **Smaller preoperative pain relief** **Contralateral hip OA** **Greater widespread pain sensitivity** **Lower radiographic OA severity** **Medial concentric disease**	[52, 53, 94, 105] [94] [11] [96] [11, 106] [20, 94] [89] [89, 94] [96] [94] [107] [89] [89]	**Worse preoperative pain** **Worse preoperative function** **Better preoperative function** **Worse mental health** **Better mental health** **Smaller preoperative pain relief** **Higher BMI** **More comorbidity** **Worse physical health** **Worse general physical status** **Presence of back pain** **Contralateral hip involvement/OA** **Lower knee extensors strength** **Lower radiographic OA severity**	[11] [52, 53, 89, 94, 99, 102, 103, 105] [94] [89, 94] [94] [96] [11, 89, 96] [11, 106] [89] [89] [94] [11, 94] [90] [89]	**Worse preoperative pain/function** **Higher BMI** **More comorbidity** **Lower radiographic OA severity** **Worse general physical status**	[99] [92, 95, 97] [92, 95, 99] [98, 100] [100]
Surgical	**Cemented prosthesis** **Cementless prosthesis**	[96] [11]	**Greater femoral head size**	[89]	NONE	NONE
Healthcare-related	NONE		**Longer waiting time**	[102]	NONE	NONE

BMI – body mass index; OA– osteoarthritis

Table 4 Determinants of worse pain and function up to 2 years follow THA with the highest level of evidence

Determinant type	**PAIN AND/OR FUNCTION**	*Number of studies with significant results*	*Number of studies investigating the determinant*
Socioeconomic	**Lower educational status**	3	3
Clinical	**Better/worse preoperative pain/function***	9	12
	Higher body mass index	6	10
	Presence or greater level of comorbidity	5	7
	Lower general health	4	4
	Lower radiographic OA severity	3	4

* Better preoperative status (pain/function) associated with worse change in status; worse preoperative status associated with worse postoperative status

CHAPTER 4

Identification of patients with suboptimal results after hip arthroplasty: development of a preliminary prediction algorithm

E. Lungu[1,2], P-A. Vendittoli[2,3], F. Desmeules[2,4]

Published in BMC Musculoskeletal Disorders. 2015 October 5; 16(1):279.

[1]Department of Biomedical Sciences, Faculty of Medicine, University of Montreal, Montreal, QC, Canada
[2]Centre de recherche de l'Hôpital Maisonneuve-Rosemont, 5415 Boul. L'Assomption, Montreal, Québec, Canada, H1T 2M4.
[3]Surgery Department, Maisonneuve-Rosemont Hospital, University of Montreal affiliated Research Center, 5415 Boul. L'Assomption, Montréal, Québec, Canada, H1T 2M4
[4]School of Rehabilitation, Faculty of Medicine, University of Montréal, Montreal, QC, Canada

1. Abstract

Background

The ability to predict preoperatively the identity of patients undergoing hip arthroplasty who are at risk of suboptimal outcomes could help implement interventions targeted at improving surgical results. The objective was to develop a preliminary PA allowing the identification of patients at risk of unsatisfactory outcomes one to two years following hip arthroplasty.

Methods

Retrospective data on a cohort of 265 patients having undergone primary unilateral hip replacement (188 total arthroplasties and 77 resurfacing arthroplasties) from 2004 to 2010 were collected from our arthroplasty database. Hip pain and function, as measured by the WOMAC, were collected as well as self-reported hip joint perception after surgery. Demographic and clinical variables recorded at the time of the surgery were considered as potential predictors. Patients were considered as having a suboptimal surgical outcome if they were in the worst quartile of the postoperative total WOMAC score and perceived their operated hip as artificial with minimal or major limitations. The PA was developed using recursive partitioning.

Results

Mean postoperative surgical follow-up was 446 ± 171 days. Forty patients (15.1%) had a postoperative total WOMAC score in the worst quartile (≥ 11.5/100) and perceived their joint as artificial with minimal or major restrictions. A PA consisting of the following variables achieved the most acceptable level of prediction: gender, age at the time of surgery, BMI, and three items of the preoperative WOMAC (degree of pain with walking on a flat surface and during the night as well as degree of difficulty with putting

socks or stockings). The rule had a sensitivity of 75.0% (95% CI: 59.8-85.8), a specificity of 77.8% (95% CI: 71.9-82.7), a positive predictive value of 37.5% (95% CI: 27.7-48.5), a negative predictive value of 94.6% (95% CI: 90.3-97.0) and positive and negative likelihood ratios of 3.38 (95% CI: 2.49-4.57) and 0.34 (95% CI: 0.19-0.55) respectively.

Conclusions

The preliminary PA shows promising results at identifying patients at risk of significant functional limitations, increased pain and inadequate joint perception after hip arthroplasty. Clinical use should not be implemented before additional validation and refining.

Keywords: Hip arthroplasty**,** Osteoarthritis**,** Prediction**,** Surgical outcomes

2. Résumé

Introduction

La capacité d'identifier les patients à risque de mauvais résultats suivant une arthroplastie de la hanche pourrait cibler des interventions afin d'améliorer leurs résultats. L'objectif était de développer un algorithme de prédiction clinique (APC) afin d'identifier les patients à risque de résultats inadéquats un à deux ans suivant une arthroplastie de la hanche.

Méthodes

Des données rétrospectives sur une cohorte de 265 patients ayant subi une arthroplastie de la hanche unilatérale (188 arthroplasties totales et 77 arthroplasties de resurfaçage de la hanche) de 2004 à 2010 ont été colligées à partir de la banque de données de l'Hôpital Maisonneuve-Rosemont. La douleur, la raideur et la fonction de la hanche, telles que mesurées à l'aide du Western Ontario and McMaster Osteoarthritis Index (WOMAC) ont été recueillies, ainsi que la satisfaction des patients par rapport à leur hanche opérée. Des données démographiques et cliniques ont été considérées comme des prédicteurs potentiels. Les patients étaient considérés à risque s'ils avaient un score WOMAC total dans le quartile le plus faible de la cohorte et s'ils percevaient leur prothèse comme une articulation artificielle avec des limitations fonctionnelles minimales ou majeures. L'algorithme de prédiction a été bâti à l'aide de la partition récursive.

Résultats

Le suivi chirurgical moyen (± écart-type) a été de 446 ± 171 jours. Quarante patients (15.1%) avaient un score WOMAC postopératoire dans le pire quartile (≥ 11.5/100) et percevaient leur hanche comme une articulation artificielle avec des limitations fonctionnelles minimales ou majeures. Un APC utilisant les

variables suivantes a fourni le meilleur niveau de prédiction : le genre, l'âge au moment de la chirurgie, l'indice de la masse corporelle pré-chirurgie et trois items du WOMAC préopératoire (niveau de douleur à la marche sur une surface plane et durant la nuit ainsi que le niveau de difficulté à mettre des bas). L'APC présente une sensibilité de 75.0% (95% IC 59.8-85.8), une spécificité de 77.8% (95% IC 71.9-82.7), une valeur prédictive positive de 37.5% (95% IC: 27.7-48.5), une valeur prédictive négative de 94.6% (95% IC : 90.3-97.0) et des rapports de vraisemblance positif et négatif de 3.38 (95% IC 2.49-4.57) et 0.34 (95% IC: 0.19-0.55) respectivement.

Conclusions

Cet algorithme préliminaire démontre des capacités prédictives prometteuses pour identifier les patients à risque de douleurs et incapacités fonctionnelles sévères post-chirurgie. Une validation externe formelle de cet algorithme est nécessaire avant de recommander son utilisation clinique.

Mots-clés : Arthroplastie de la hanche, Arthrose, Prédiction, Douleur, Fonction

3. Background

Recent recommendations suggest that THA is indicated when the patients' functional limitations and pain levels due to hip OA are refractory to pharmacological and non-pharmacological treatments [7, 8]. HR is an alternative to THA in patients who are younger, more active, with normal kidney function and appropriate proximal femoral bone morphology and quality [43]. Both THA and HR are considered efficacious for the great majority of patients undergoing these procedures [10, 13, 112, 113]. Although generally successful at alleviating coxarthrosis-related ailments, hip arthroplasty can yield subpar results in terms of pain and functional outcomes as well as degree of satisfaction in a non-negligible proportion of patients. For example, a recent systematic review reports that 7 to 23% of the patients undergoing THA experience unfavourable pain outcomes three months to five years after the procedure [3]. Moreover, up to 15% of the patients report dissatisfaction with surgery [4, 17]. To our knowledge, no formal data on proportions of patients with poor pain, functional and satisfaction levels after HR exists. However, it can be posited that these proportions are similar to the ones observed among patients undergoing THA, as studies indicate that these outcomes are similar between the two procedures [114, 115].

In light of these observations, careful case management must be implemented in order to minimize unsuccessful outcomes. Potential interventions directed at improving surgical outcomes include patient education and intensive rehabilitation. However, identification of patients at risk of severe pain and functional limitations after THA or HR is difficult. A multitude of factors related to poor functional and pain outcomes following hip arthroplasty have been identified. These include worse preoperative levels of pain and function, lower educational level, comorbidities, presence of back pain or higher BMI, among others [11, 52, 53, 92, 94, 95, 97, 99, 103, 105]. Nevertheless, regardless of the quantity of the evidence of potential risk factors, no definitive consensus has been reached concerning their identity and the

magnitude of their association with postoperative pain, functioning and satisfaction. In light with these observations, an algorithm aimed at identifying with sufficient accuracy which patients present the greatest risk of unsuccessful outcomes may assist in the care process. Therefore, the objective of this study was to develop a preliminary PA used to identify patients at risk of unfavourable functional status, pain and joint perception one to two years following THA or HR.

4. Methods

4.1. Study design

This study entailed a retrospective analysis of longitudinal, prospectively collected data. The methodology adheres to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines for observational cohort studies (Appendix K).

4.2. Data collection

Our prospective arthroplasty database was consulted in order to identify patients eligible for inclusion in the study. The database contains extensive baseline and follow-up data on patients undergoing hip procedures, including THA and HR. All patients provide informed consent to participate. Independent assessors who are not involved in the medical care of the patients collect the prospective data.

Inclusion criteria were 1) patients undergoing primary unilateral THA or HR, 2) diagnosis of primary hip OA, 3) complete preoperative and one to two-year postoperative self-reported outcome questionnaire responses. The main exclusion criteria were 1) THA or HR of the contralateral hip before the relevant follow-up evaluation, 2) revision of the implant before the one to two-year follow-up, 3) diagnosis of inflammatory hip arthritis, pediatric hip disease, post-traumatic hip or any hip disease other than primary

OA. Data on all patients having undergone hip interventions were assessed for inclusion. All patients were initially interviewed just before their intervention. Postoperative outcomes were collected 12 to 24 months after the surgery.

4.3. Dependent variables

Functional status and pain levels were assessed preoperatively and at follow-up with the WOMAC [116]. The WOMAC consists of the following domains: pain (5 items), stiffness (2 items) and functional limitation (17 items). Items are scored on a 5-point Likert scale representing different degrees of intensity (none, mild, moderate, severe or extreme). The scores of each domain as well as the total score were standardized on a 0 to 100 scale, with a greater score indicating more pain, stiffness or functional limitation. The psychometric qualities of the WOMAC, including its responsiveness, convergent construct validity and reliability have been found excellent for evaluating patients with hip OA undergoing hip arthroplasty [75, 76].

At follow-up, self-perceived joint perception was measured by asking the patient a multiple-choice question: “How do you perceive your operated hip?” with the possible responses being “Like a native or natural joint”, “Like an artificial joint with no restriction”, “Like an artificial joint with minimal restriction”, “Like an artificial joint with major restriction” and “Like a non-functional joint” [79]. Evaluation of joint perception has been strongly associated with validated clinical scores of patient-reported outcome measures and can be employed as a measure of patient satisfaction [79].

No consensus exists regarding what represents poor outcome following hip arthroplasty. Hence, patients of risk of suboptimal outcomes were defined as the ones in the worst quartile of the total WOMAC score

at follow-up (i.e. WOMAC score >11.5) and perceiving their hip "like an artificial joint with minimal restriction", "like an artificial joint with major restriction" or "like a non-functional joint".

4.4. Independent variables

Several of the variables that were collected preoperatively and available in the database were considered as potential predictors of suboptimal hip arthroplasty outcome. Demographic variables included age and gender. Clinical variables included BMI, previous hip interventions and medical comorbidities (diabetes, gastrointestinal disease, immunosuppression secondary to corticosteroid use or other causes, cardiac disease, obesity, pulmonary disease, neurologic disease, urologic disease, and other comorbidities). Pain localization (back, radicular, buttocks, trochanter, groin, thigh, knee and/or calf) as well as whether hip pain was present at rest, after the first few steps, after a long walk and during sexual relations were also considered. Answers to the 24 individual items of the preoperative WOMAC questionnaire were additionally included in the analysis as potential predictors.

4.5. Statistical analysis

Baseline and follow-up mean WOMAC scores along with their standard deviations were calculated. Differences between time points in relation to total WOMAC scores and the respective domains were assessed using paired samples Student-t tests, with a significance level set at 0.05.

The CART approach was used to build the PA as it is one of the most effective algorithms of recursive partitioning [82]. It is based on maximizing the within-node homogeneity by evaluating all combinations of potential predictors, thus minimizing the within-node error. The Gini impurity measure was used as a splitting criterion to develop the decision trees [83]. Data for all the patients in the training set was used to

develop the PA. Firstly, all the potential predictor variables were employed to develop models using an automated approach. Secondly, a manual approach entailed the development of additional models by inputting independent variables that were judged to be more readily available and easier to employ in a clinical setting. For example, age and gender were favoured over the number of comorbidities and previous hip interventions because the latter two could be affected by a recall bias or would require extensive medical file review. The predictive values of every model were calculated along with their 95% confidence intervals, namely sensitivity, specificity, positive and negative predictive values as well as positive and negative likelihood ratios [84]. Among all the proposed models, the one that showed the highest level of sensitivity and an acceptable level of specificity and that fit the ease-of-use criterion was selected in order to develop the screening tool. Internal validity of the model was then evaluated by the use of 1,000 bootstrap resamples [85]. All analyses were performed using IBM SPSS Statistics Version 20.0 (SPSS Inc., Chicago).

4.6. Ethics

The research ethics committee of our centre approved the study annually.

5. Results

5.1. Participants

Our database yielded 2963 entries with at least some preoperative data on hip arthroplasty procedures performed from October 2004 to February 2014. Out of these, 1207 procedures (40.7%) fit the inclusion criteria. Incomplete preoperative and/or postoperative data required for the purposes of the current study obliged the exclusion of a further 942 entries. Thus, a total of 265 primary hip arthroplasty interventions (60 classical THAs, 128 large-femoral head diameter THAs, and 77 HRs) with complete preoperative and

postoperative data were included in the study (follow-up mean ± SD: 446.3 ± 171.1 days), representing a participation proportion of 22.0%.

Table 1 shows selected characteristics of the 265 patients included in the study. The mean age of the participants was 52.0 (SD 9.0) and 67.4 % were male. The mean BMI was 28.2 (SD 5.1) and each patient had on average 0.79 comorbidities (SD 0.96). Mean follow-up was 446 SD: ± 171.1 days and ranged from 253 to 1638 days. Postoperatively, the patients had significantly improved on pain (-44.9, SD: ± 22.6, 95% CI -42.1 to -47.6), stiffness (-44.6, SD: ±25.1, 95% CI -41.6 to -47.7), function (-43.6, SD: ±21.9, 95% CI -40.9 to -46.2) as well as total WOMAC score (-43.9, SD: ±21.1, 95% CI -41.4 to -46.5) (Table 2). Seventy-six patients (29%) reported that they perceived their prosthetic joint as artificial with minimal or major restrictions (Table 3).

Out of the 265 patients eligible for inclusion in the study, 40 (15.1%) had a total WOMAC score > 11.5 and perceived their joint as artificial with minimal or major restrictions. Hence, these patients were considered as having suboptimal surgical outcomes.

5.2. Final prediction algorithm

After developing several prediction rules, the algorithm with the highest level of sensitivity and an appropriate level of specificity was chosen. It consists of patient gender, age at the time of surgery, BMI and 3 items of the preoperative WOMAC, namely degree of pain with walking on a flat surface and during night and degree of difficulty with putting socks or stockings (Figure 1). Patients respond sequentially to the questions and their risk status is determined according to the classification algorithm (Figure 2).

The final PA correctly identified 30 out of the 40 patients considered at risk of suboptimal outcome based on their surgical outcomes and 175 patients out 225 were identified as not at risk of suboptimal outcome (Table 3). Therefore, the PA had a sensitivity of 75.0% (95% CI: 59.8-85.8), a specificity of 77.8% (95% CI: 71.9-82.7) and a positive likelihood ratio of 3.38 (95% CI: 2.49-4.57) (Table 4). The other prediction models that were also considered are presented in Appendix L.

5.3. Internal validation

Validation of the rule was established using 1,000 bootstrap re-samples. Table 4 shows the estimated bootstrap values of the predictive measures being close to the original ones, thus suggesting an appropriate accuracy of the proposed model.

6. Discussion

Since THA and HR can bring significant improvement in patients suffering from hip OA, careful management of subjects at risk of having unsuccessful outcomes is indicated. We aimed to develop a prediction tool in order to facilitate the preoperative identification of these patients, which could possibly ameliorate their surgical outcomes. With a cohort of 265 patients undergoing primary hip arthroplasty for OA, we were able to create a PA predicting the identity of patients that are at the highest risk of unsuccessful outcomes. Albeit preliminary in nature and requiring further development and validation, our PA has excellent predictive capacities, with a sensitivity of 75.0% (95% CI 59.8.4-85.8), a specificity of 77.8% (95% CI 71.9-82.7) and a positive likelihood ratio of 3.38 (95% CI 2.49-4.57).

To our knowledge, one model predicting the identity of patients at risk of poor outcomes after THA has been developed [69]. Consisting of patient age, BMI and gender, the model was able to correctly predict

patients' outcomes with a sensitivity of 87.5% (95% CI 52.9 – 97.8), a specificity of 72.4% (95% CI 54.3 – 85.3) and a positive likelihood ratio of 3.17 (95% CI 1.66 – 6.05). However, surgical success was determined solely based on functional outcomes (change between the preoperative and six-month LEFS score), and the results are based on a cohort of 37 patients.

The selection of patients for inclusion in the study was based on the availability of complete data for important determinants of hip arthroplasty outcomes as reported in the literature. This allowed the development of a prediction rule that is consistent with the clinical reality. Because there is no unequivocal definition of what represents suboptimal outcome following hip replacement, several criteria of classifying patients who are at risk have been considered and different prediction models were built accordingly. The choice of the final model was based on the principle of selecting a screening tool minimizing the number of false negatives that is easily employable in a clinical setting. Accordingly, an algorithm with a sensitivity of 75.0% and a specificity of 77.8% was deemed suitable. Although the positive likelihood ratio of 3.38 of the PA can be considered subpar when compared to accepted diagnostic standards, the PA performs similarly to other validated prediction models in the epidemiological literature. For example, the positive likelihood ratios of the Ottawa Knee and Ankle rules assessing the necessity of a roentgenographic evaluation in cases of acute knee and ankle injuries respectively are both inferior to the one reported by our PA [71, 72].

The algorithm with the most appropriate predictive capabilities contains two demographic variables (gender and age), one clinical (BMI) and 3 items of the preoperative WOMAC questionnaire (two pain-related and one function scale). All of these variables have been consistently related to hip arthroplasty outcomes [11, 93-97]. Moreover, the PA comprises all the predictors reported by Slaven et al. (2012) in

their model, namely age, gender and BMI, thus pointing towards the importance of these factors for prediction of hip arthroplasty results. It is noteworthy to mention that the prediction of surgical outcomes in women is achieved by age and BMI, with the body mass being the only modifiable risk factor. In the case of men, potential modifiable risk factors include BMI, degree of hip pain walking on a flat surface and during the night as well as degree of difficulty putting on socks or stockings. However, caution should be used, as recursive partitioning does not imply a causative relationship between variables [117]. Indeed, interventions targeted at ameliorating either of the items of the PA, such as weight loss in the case of high BMI, will not necessarily improve the outcome of the surgery; it will merely imply that the patient will be classified as not at risk of suboptimal outcomes by the PA. Further research in terms of appropriate interventions to improve surgical outcomes should be undertaken.

When developing the PA, we intended for it to be a clinically pertinent tool. The decision to include patients with different types of hip arthroplasties was taken in order to generate a PA that has the ability to perform successfully in a heterogeneous population. Moreover, we included subjects with complete information one to two years following the procedure, as patients are followed closely by their surgeons during this period, and the rehabilitation process can easily be altered if the progression is judged suboptimal.

In one instance, the interpretation process may yield a counterintuitive situation. For example, it is possible, in an extreme scenario, for a 49 year-old male patient with a BMI of 22 kg/m^2 and with no pain when walking on a flat surface as well as with no difficulty with putting on socks or stockings to be classified as at risk of suboptimal surgical outcomes. This pattern of answers was however shown to have

the best predictive capabilities when developing the algorithm with recursive partitioning. This situation underlines the concept that a predictor is not necessarily a determinant.

Strengths of the study

The developed PA is, to our knowledge, the first one of its kind to discriminate THA or HR results based on more than one parameter, namely patients' functional, stiffness and pain levels as well as their perception of the replaced hip joint. In the context of a lack of an accepted standard of surgical failure, this approach increases the likelihood of the patients thusly classified to truly present subpar outcomes. Moreover, this classification identified 15% of the patients as having unsuccessful outcomes, well in line with the published proportions of what can be considered a suboptimal outcome [3]. Finally, the rigorous statistical analysis employed in the development of the PA underlines the stringency of our approach.

Limitations of the study

Due to its retrospective design, the study has a certain risk of selection bias. Compared to the subjects excluded due to missing data, the included participants were, on average, younger, had a greater number of comorbidities and a greater proportion were male ($p < 0.05$, data not shown). Nevertheless, the preoperative baseline status as measured by the WOMAC domains as well as the total WOMAC score was not statistically significantly different between the included and the non-included subjects ($p > 0.05$, data not shown). Additionally, only 265 out of the 1207 procedures (22.0%) that were performed during the study time period met the inclusion criteria, therefore potentially limiting the generalizability of the results. Moreover, the population under study was patients undergoing primary unilateral hip replacement procedures, which precludes the utilization of this tool for patients undergoing revision or bilateral interventions. Recent evidence identifies other variables potentially associated with hip arthroplasty outcomes that were not included in our study, thus potentially limiting the pool of candidate predictor variables. Although there is no consensus regarding the optimal sample size for developing models

employing recursive partitioning, the progressively smaller number of cases in the leafs as the tree was built may limit the reliability of the findings, prompting further development in a subsequent study. Before employing it in a clinical setting, the decision rule has to be validated in a different sample of patients. Moreover, the performance of the PA has to be compared to clinical judgement alone and its financial impacts require evaluation.

7. Conclusions

The developed PA may discriminate with excellent capabilities the patients undergoing hip arthroplasty that are at the highest risk of suboptimal pain, functional limitations and joint perception outcomes on an average of 15 months following the intervention. Its implementation has the potential of targeting susceptible individuals such as to modify their risk profile, and eventually, improve surgical results.

8. Competing interests - The authors declare that they have no competing interests.

9. Authors' contributions

EL led the design, the coordination, the collection of data, the statistical analysis, the interpretation of results and the writing of the manuscript. PAV participated in the design, the coordination, the collection of data, the interpretation of results and the writing of the manuscript. FD participated in the design, the coordination, the statistical analysis, the interpretation of results and the writing of the manuscript. All authors read and approved the final manuscript.

10. Acknowledgements

EL is supported by a training award from the FRQS.

Table 1 Selected characteristics of the participants who underwent hip arthroplasty (n = 265)1

Variables considered for PA development	n (%)	Mean (SD)	Other collected variables	n (%)
Demographics			**Contralateral hip status**	
Age (years)		52 (9.0)	Unaffected	134 (50.4)
Female	89 (33.6)		Affected, not operated	105 (39.7)
Clinical characteristics			Unavailable	26 (9.9)
BMI ¯ (kg/m^2)		28.2 (5.1)	**Charnley class**	
Medical comorbidities			Charnley A	124 (46.8)
Diabetes	19 (7.2)		Charnley B	87 (32.8)
Gastrointestinal disease	16 (6.0)		Charnley C	13 (4.9)
Immunosuppression	3 (3.0)		Unavailable	41 (15.5)
Cardiac disease	21 (7.9)		**Employment status**	
Obesity	41 (15.5)		Employed	163 (61.5)
Osteoporosis	2 (0.8)		Household	44 (16.6)
Pulmonary disease	15 (5.7)		Retired	9 (3.4)
Neurological disease	1 (0.4)		Other	11 (4.2)
Urological disease	1 (0.4)		Unavailable	38 (14.3)
Other	91 (34.3)		**Walking aid**	
None	124 (46.8)		Incapable with aid	5 (1.9)
Presence of back pain	40 (15.1)		Crutches	1 (0.4)
Pain localization			Two canes	31 (11.7)
Buttocks	128 (48.3)		Cane on a permanent basis, instability	118 (44.5)
Trochanter	164 (61.9)		Cane for outdoor activities	42 (15.8)
Groin	177 (66.8)		Cane for long distance walking	43 (16.2)
Thigh	124 (46.8)		Unavailable	25 (9.5)
Knee	111 (41.9)		**Knee(s) status**	
Calf	36 (13.6)		Affected	28 (10.5)
Radicular	6 (2.2)		Unaffected	204 (77.0)
Elsewhere	3 (1.1)		Unavailable	33 (12.5)
Presence of hip pain			**Level of activity in the 3 months before surgery**	
At rest	148 (55.8)		Heavy work/sport	26 (9.8)
After first few steps	182 (68.8)		Moderate work	53 (20.0)
After a long walk	224 (84.5)		Mild work/walking	112 (42.3)
During sexual intercourse	156 (58.9)		Sedentary	34 (12.8)
			Immobile	6 (2.3)
			Unavailable	34 (12.8)
			Duration of walking before eliciting pain	
			Walking unaffected	37 (14.0)
			31-60 minutes	51 (19.2)
			11-30 minutes	82 (30.9)
			2-10 minutes	53 (20.0)
			< 2 minutes	16 (6.0)
			Walking impossible	1 (0.4)
			Unavailable	25 (9.5)

Table 2 Changes in WOMAC scores of the participants between preoperative measurement and following hip arthroplasty (n = 265)

	Mean score preoperatively† (SD)	Mean score postoperatively† (SD)	Change in score ‡ (SD)	95% CI	Comparison between time points (*p* value)
WOMAC					
Pain	55.4 (19.2)	10.5 (16.7)	- 44.9 (22.6)	- 42.1 to – 47.6	<0.001*
Stiffness	57.1 (19.4)	12.5 (18.1)	- 44.6 (25.1)	- 41.6 to – 47.7	<0.001*
Function	53.2 (20.0)	9.6 (15.3)	- 43.6 (21.9)	- 40.9 to – 46.2	<0.001*
Total score	54.0 (18.7)	10.1 (15.1)	- 43.9 (21.1)	- 41.4 to – 46.5	<0.001*

SD: standard deviation
CI: confidence interval
† Scores presented as standardised scores. Lower scores sign a better condition. Scores were measured on the day of the surgery.
‡ Negative changes in score indicate an improvement of the condition. Scores were measured on 446.3 ± 171.1 days following the intervention.
* $p < 0.05$

Table 3. Postoperative joint perception of the patients who underwent hip arthroplasty according to the distribution of their postoperative total WOMAC scores (n = 265)

	Joint perception					
WOMAC quartile	Native/Natural	Artificial with no restrictions	Artificial with minimal restrictions	Artificial with major restrictions	Non-functional	**TOTAL**
First	56	21	4	0	0	81
Second	32	9	6	0	0	47
Third	28	14	25	1	0	68
Fourth	17	12	**35***	**5***	0	69
TOTAL	133	56	70	6	0	**265**

* patients considered at risk of suboptimal outcome (n = 40); a higher quartile indicates a worse total WOMAC score at follow-up

Table 4 Two by Two table of predicted versus actual outcomes of the prediction algorithm

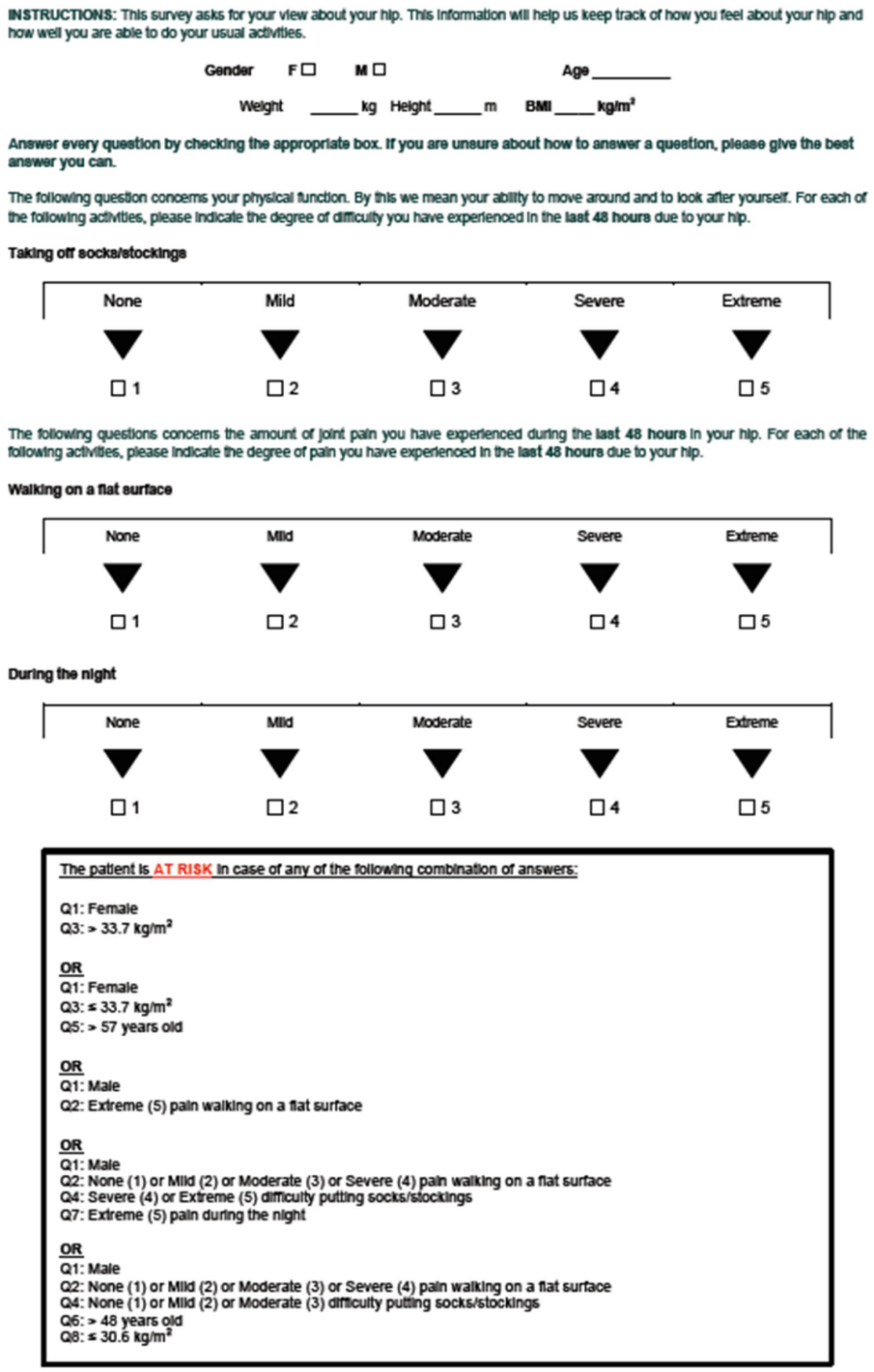
	Actual Outcome	
	AT RISK	**NOT AT RISK**
Predicted outcome	Worst postoperative WOMAC quartile (>11.5/100) & "Artificial with minimal or major limitations" joint perception	Postoperative WOMAC ≤ 11.5 &''Artificial with no limitations'' or ''Natural joint'' joint perception
AT RISK	30	50
NOT AT RISK	10	175
TOTAL	*40*	*225*

Table 5 Validity measures of the prediction algorithm

Measure	Estimates in training sample	Estimates with 1,000 bootstrap resamples
Sensitivity % (95% CI)	75.0 (59.8.4-85.8)	75.0 (60.0-88.0*)
Specificity % (95% CI)	77.8 (71.9-82.7)	77.8 (72.2-82.9*)
Positive predictive value % (95% CI)	37.5 (27.7-48.5)	37.2 (27.2-47.2*)
Negative predictive value % (95% CI)	94.6 (90.3-97.0)	94.7 (91.2 to 97.8*)
Positive likelihood ratio (95% CI)	3.38 (2.49-4.57)	3.38 (2.50 to 4.63*)
Negative likelihood ratio (95% CI)	0.32 (0.19-0.55)	0.32 (0.15 to 0.52*)

- * 95% asymptotic confidence intervals
- Sensitivity: number of participants classified at risk both by the PA and the postoperative WOMAC score and joint perception divided by all participants classified at risk by the postoperative WOMAC score and the joint perception (actual outcome).
- Specificity: number of participants classified not at risk by the PA and the postoperative WOMAC score and joint perception divided by all participants classified not at risk by the postoperative WOMAC score and joint perception (actual outcome).
- Positive predictive value: number of participants classified at risk by the PA and the postoperative WOMAC score and joint perception divided by all participants classified at risk by the PA (predicted outcome).
- Negative predictive value: number of participants classified not at risk by the PA and the postoperative WOMAC score and joint perception divided by all participants classified not at risk by the PA (predicted outcome).
- Positive likelihood ratio: sensitivity/ (1-specificity)
- Negative likelihood ratio: (1-sensitivity)/specificity.

Figure 1 Prediction algorithm to identify patients at risk of suboptimal outcomes after hip arthroplasty



INSTRUCTIONS: This survey asks for your view about your hip. This information will help us keep track of how you feel about your hip and how well you are able to do your usual activities.

Gender F ☐ M ☐ **Age** _____

Weight _____ kg **Height** _____ m **BMI** _____ kg/m²

Answer every question by checking the appropriate box. If you are unsure about how to answer a question, please give the best answer you can.

The following question concerns your physical function. By this we mean your ability to move around and to look after yourself. For each of the following activities, please indicate the degree of difficulty you have experienced in the **last 48 hours** due to your hip.

Taking off socks/stockings

None	Mild	Moderate	Severe	Extreme
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

The following questions concerns the amount of joint pain you have experienced during the **last 48 hours** in your hip. For each of the following activities, please indicate the degree of pain you have experienced in the **last 48 hours** due to your hip.

Walking on a flat surface

None	Mild	Moderate	Severe	Extreme
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

During the night

None	Mild	Moderate	Severe	Extreme
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

The patient is AT RISK in case of any of the following combination of answers:

Q1: Female
Q3: > 33.7 kg/m²

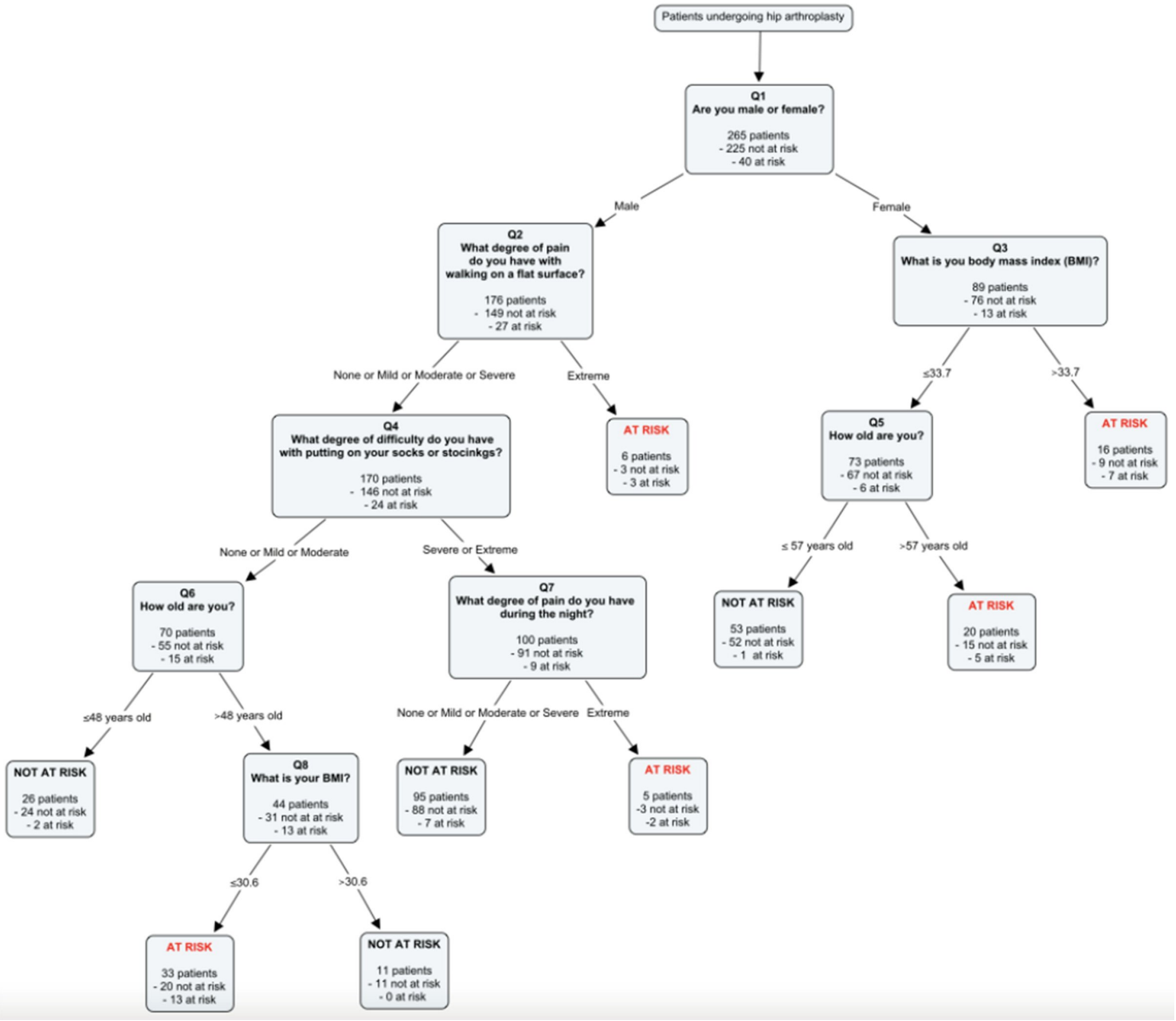
OR
Q1: Female
Q3: ≤ 33.7 kg/m²
Q5: > 57 years old

OR
Q1: Male
Q2: Extreme (5) pain walking on a flat surface

OR
Q1: Male
Q2: None (1) or Mild (2) or Moderate (3) or Severe (4) pain walking on a flat surface
Q4: Severe (4) or Extreme (5) difficulty putting socks/stockings
Q7: Extreme (5) pain during the night

OR
Q1: Male
Q2: None (1) or Mild (2) or Moderate (3) or Severe (4) pain walking on a flat surface
Q4: None (1) or Mild (2) or Moderate (3) difficulty putting socks/stockings
Q6: > 48 years old
Q8: ≤ 30.6 kg/m²

Figure 2 Graphical representation of the prediction algorithm identifying patients at risk of suboptimal surgical outcomes after hip arthroplasty



CHAPTER 5

Preoperative determinants of patient-reported pain and physical function levels following total knee arthroplasty: a systematic review

E. Lungu[1,2§], P-A. Vendittoli[2,3], F. Desmeules[2,4]

[1]Department of Biomedical Sciences, Faculty of Medicine, University of Montreal, Montreal, QC, Canada
[2]Centre de recherche de l'Hôpital Maisonneuve-Rosemont, 5415 Boul. L'Assomption, Montreal, Québec, Canada, H1T 2M4.
[3]Surgery Department, Maisonneuve-Rosemont Hospital, University of Montreal affiliated Research Center, 5415 Boul. L'Assomption, Montréal, Québec, Canada, H1T 2M4
[4]School of Rehabilitation, Faculty of Medicine, University of Montréal, Montreal, QC, Canada

Manuscript in first revision by *The Open Orthopaedics Journal*

1. Abstract

Background

A sound knowledge of the determinants of TKA outcomes could help in patient selection, preparation and education before they undergo surgery. We aimed to assess the current status of the literature evaluating preoperative determinants of early and medium term patient-reported pain and disability following TKA.

Method

A search in Medline, Pubmed, Embase and CINAHL until October 2014 was undertaken. Selection criteria included: 1- participants undergoing primary unilateral TKA with a follow-up up to 2 years, 2- validated disease-specific patient-reported outcome measures assessing pain and/or function used as outcome measure and 3- identification of preoperative determinants obtained via multivariate analyses. Risk of bias was assessed using a modified version of the *Methodology checklist for prognostic studies*.

Results

Thirty-four prognostic explanatory studies were included. Mean total score of the methodological quality was 80.0±12.7 %. Sociodemographic and psychosocial determinants included greater socioeconomic deprivation (two out of two studies), greater levels of depression and/or anxiety (seven out of 11 studies) and greater preoperative pain catastrophizing (all four studies). Significant clinical determinants included worse preoperative pain and disability (21 out of 23 studies), presence or greater levels of comorbidity (12 out of 23 studies), back pain (five out of six studies) and lower general health (all 11 studies).

Conclusions

Several significant determinants of short to medium-term pain and functional outcomes following TKA have been summarized by studies with moderate-to-high methodological quality. No conclusions can be reached regarding the strength of the associations between significant determinants and TKA results because of heterogeneity of study methodologies and results. Further high-quality research is required.

Keywords: Knee Osteoarthritis**,** Total Knee Arthroplasty**,** Pain**,** Functional limitation**,** Determinants

2. Résumé

Introduction

Une meilleure connaissance des déterminants des résultats de l'arthroplastie totale du genou (ATG) pourrait aider dans la sélection, la préparation et à l'éducation des candidats à la chirurgie. L'objectif de cette étude était d'évaluer la littérature identifiant les déterminants à court et moyen terme de la douleur et des incapacités fonctionnelles suivant l'ATG.

Méthodes

Une recherche dans Medline, Pubmed, Embase et CINAHL jusqu'en octobre 2014 a été réalisée. Les critères de sélection pour l'inclusion des études étaient : 1 – patients subissant une ATG primaire unilatérale avec un suivi post-opératoire d'au plus 2 ans, 2 – les études utilisaient des mesures de résultats auto-rapportés validés évaluant la douleur et la fonction et 3 – l'identification des déterminants était réalisée à l'aide d'analyses statistiques multivariées. La qualité méthodologique était évaluée à l'aide d'une version modifiée de *Methodology checklist for prognostic studies*.

Résultats

Trente-quatre études pronostiques ont été incluses. Le score total moyen de la qualité méthodologique des études était de 80.0%±12.7%. Les facteurs sociodémographiques et psychosociaux significativement associés avec des douleurs et des incapacités post-opératoires incluaient un niveau socioéconomique plus faible (deux études sur un total de deux), des symptômes de dépression et/ou d'anxiété (sept études sur 11) ainsi que la catastrophisation de la douleur (quatre études sur quatre). Les déterminants cliniques considérés comme significativement associés avec la douleur et les incapacités post-opératoires incluaient un niveau préopératoire de douleur et d'incapacité plus élevé

(21 études sur 23), la présence ou un plus grand nombre de comorbidités (12 études sur 23), la présence de lombalgie (cinq études sur six) et un état plus précaire de la santé générale (11 études sur 11).

Conclusions

Plusieurs déterminants de la douleur et des incapacités fonctionnelles à court et moyen terme suivant l'ATG ont été identifiés en se basant sur des études de qualité méthodologique modérée à élevée. Une hétérogénéité importante des méthodologies et des résultats empêche des conclusions formelles quant à la force d'association entre les déterminants significatifs et les résultats de l'ATG. Davantage de recherches de haute qualité sont nécessaires.

Mots-clés : Arthrose du genou, Arthroplastie totale du genou, Douleur, Incapacité fonctionnelle, Déterminants

3. Introduction

TKA is a common procedure intended at treating patients with knee OA suffering from pain and disability [118]. Its predominant success rendered it the second most common type of orthopaedic intervention [14]. This tendency will likely maintain, as projections suggest a six-fold increase in the number of primary TKAs performed in the next decades [14]. Although TKA is generally a successful intervention, leading to amelioration in pain levels and functional status, it yields suboptimal results in up to one third of patients [4, 5, 12, 16, 119]. Sound knowledge of determinants of TKA outcomes can help in patient selection, preparation and education, especially regarding possible risks and benefits of the procedure [18]. This is particularly relevant with respect to early and medium-term outcomes, as after a significant amelioration three to six months postoperatively, pain and physical function levels vary little subsequently until two years following surgery [19, 53]. During this time, patients are closely monitored by their surgeons, and the medical treatment and rehabilitation can be readily altered if progress is deemed unsatisfactory.

Previous systematic reviews attempted to summarize the determinants of TKA outcomes. Santaguida et al. (2008) identified older age and female gender to be associated with worse function following TKA [18]. However, their results are based on studies published until 2001. Van Jorbegen et al. (2014) focused on protective determinants of anterior knee pain following TKA, and their findings included mostly surgical factors, namely femoral components with a posterior centre of rotation, resection of Hoffa's pad, patellar rim electrocautery and preventing combined component internal rotation [21]. Vissers et al. (2012) focused their systematic review on psychosocial factors associated with TKA outcomes and identified pain catastrophizing and lower preoperative mental health as significant determinants of poor TKA outcomes [22]. Regardless of the evidence summarized by these systematic

reviews, no consensus exists concerning either the identity or the strength of association between TKA determinants and poor outcomes. Consequently, there is an evident necessity of a comprehensive review encompassing the highest quality of evidence, which can be achieved by focusing on studies employing validated PROMs of pain and function that also gauge the independent effect of determinants via multivariate analyses [91].

The purpose of this systematic review was to assess the current status of the literature evaluating the determinants of poor outcomes in terms of pain and functional levels following TKA. We also aimed to compare the determinants according to the approach of quantifying TKA results, i.e. as a measure of patients' *postoperative status* or of *postoperative change*. Finally, because some studies evaluate pain and function either separately, such as in the case of WOMAC pain and function subscales, or in a combined manner (total WOMAC score), we intended to parallel determinants according to this categorization.

4. Materials and methods

4.1. Literature search and study identification

A search in four databases (Medline, Pubmed, Embase and CINAHL) from their respective inception dates until October 2014 was undertaken using a combination of keywords and MESH terms (see Appendix D). Manual searches of previously published reviews and reference lists from relevant articles were also conducted. Two authors independently reviewed the titles, abstracts and full texts of the articles in order to evaluate their eligibility.

4.2. Study selection

The following selection criteria were applied:

1. Participants were primary unilateral TKA patients with ≤10% of the sample undergoing unicompartmental knee arthroplasty, bilateral TKA or revision TKA
2. ≥90% of the study sample was diagnosed with knee OA
3. Results are presented for a follow-up between 6 weeks and 2 years
4. The outcome measure was a disease-specific validated PROM assessing pain and/or function
5. Identification of determinants was obtained using multivariate analyses
6. Article is published in English or French

4.3. Data extraction

A standardized form was employed to extract data. Participants' characteristics (diagnosis, type of surgery, age and gender proportion), number of patients, follow-up period, outcome measures, statistical methods used and statistical adjustments, as well as significant and non-significant determinants reported by each study were recorded. Each article was extracted by one of the raters and verified by another in order to reduce the risk of extraction errors.

4.4. Methodological quality appraisal

Two trained reviewers independently performed the appraisal of the methodological quality of the included studies and results were discussed in order to reach consensus. In case of disagreement, a third reviewer was available for mediation of differences.

The risk of bias and the methodological quality of the included studies was assessed using a modified version of the *Methodology Checklist for Prognostic Studies* developed by Hayden et al. (2003) [74]. This tool includes six items: "Study participation", 'Study attrition", "Prognostic factor measurement", 'Outcome measurement", "Confounding measurement and account" and 'Analysis". Each item is evaluated according to its risk of potential bias: "yes" indicates a low risk of bias, "no" indicates a high risk of bias and "unclear" indicates an unclear or an unknown risk given the information available. For each item of the checklist, a score of 2 was given if a low risk of bias was present, a score of 1 if the risk was judged unclear and 0 if the risk was high. For the 'Study participation" item, a score of 1 was attributed if the study was retrospective in nature and that no information was available regarding patients not included in the study because of incomplete data. For the "Study attrition" item, a score of 0 was given automatically if the follow-up proportion at the relevant time-point was inferior to 80%. A score of 0 was given for the "Confounding measurement and account" item if confounding factors such as age, gender and BMI were not accounted for in the multivariate analysis.

4.5. Data synthesis

Determinants of TKA outcomes were summarized based on whether results were reported as postoperative change or postoperative status, and whether pain and function were assessed as separate or combined constructs. Given the nature of the study designs and the heterogeneity of included studies in terms of depended and independent variables' constructs and definitions, as well as variations in follow-up periods, only a qualitative synthesis of results was performed.

5. Results

5.1. Description of the included studies

Initial literature search yielded 139 full-text articles for assessment of eligibility. After further exclusion of 105 full-text articles for reasons presented in Figure 1, 34 manuscripts were included. Table 1 indicates relevant characteristics of the included studies. Results from two manuscripts are presented conjointly because of analyses performed on the same cohort [52, 53]. The WOMAC was the validated tool used to measure postoperative pain and/or function in 25 studies, whereas the Oxford Knee Score (OKS) was employed in 9 studies. Nine studies have employed the change in pain and/or function after the surgery as an outcome measure.

Postoperative raw scores at follow-up were considered as a measure of outcome in 25 studies. Seven studies had a sample size smaller than 100 and 13 had a sample size greater than 500 patients. Only six studies presented a power calculation or considered a way of estimating required sample sizes [120-124].

5.2. Methodological quality of the included studies

Table 2 indicates the methodological quality scores of the included studies after consensus. Mean total score for the methodological quality was 80.0% (SD 12.7%). No study received lower than 58.3% and four studies were graded 100% [121, 123, 125, 126]. Overall, these results indicate a moderate-to-high methodological quality.

Three domains of the methodology appraisal ("Prognostic factor measurement", "Outcome measurement" and "Analysis") scored on average the maximal possible grade. The domain with the

worse mean score (0.97, SD 1.02) was "Confounding measurement and account", with 17 studies not accounting for age, gender or BMI or other potential confounding factor in the multivariate analyses. A noteworthy number of studies (11 out of 34) reported a follow-up proportion inferior to 80%. This negatively impacted the study attrition domain.

5.3. Preoperative determinants of TKA pain and function outcomes

5.3.1. Demographic determinants

Sixteen studies investigated the association of age at the time of surgery and *postoperative status*. Neuburger et al. (2012) mention that being less than 60 years old is a significant determinant of poorer total OKS score at six months [127]. However, the same study reports that being older than 80 years old was also related to worse total OKS score at six months. Four more studies identified older age at the time of surgery as a factor associated with worse functional level following TKA [19, 128-130]. Nevertheless, 11 studies report no significant effect of age on postoperative pain and function status [52, 99, 123, 124, 128, 131-134]. Eleven studies did not report a significant relationship between gender and TKA outcomes [19, 52, 99, 123, 127, 128, 131, 134-136]. The three that found a significant association seem to yield more consistent results regarding the deleterious effect of female gender on TKA pain and function outcomes [129, 130, 133].

Only limited evidence can be extracted regarding demographic determinants of *postoperative change* in terms of pain or function. In regards to gender, one study identified male gender to be associated with a smaller change in the 12-month WOMAC function score [12]. Alzharani et al. (2011) report that male patients were 0.72 times more likely to not achieve the MCID for total OKS score one year after TKA compared to women, i.e. female gender is a determinant of unsatisfactory outcome [137].

Baker et al. (2012) suggest that younger age is associated with less improvement on the total OKS score recorded six to 12 months postoperatively, whereas Alzharani et al. (2011) indicate that older age is associated with lower odds of attaining the MCID of the total WOMAC score one year following TKA [120, 137].

5.3.2. Socioeconomic determinants

Although scarce, the evidence regarding socioeconomic factors seems to point to several significant findings only in the case of the outcomes measured as *postoperative status*. Greater social deprivation was identified in two studies as a determinant of worse pain and functional limitation when simultaneously controlling for multiple confounding factors [127, 129]. A lower income was linked to a worse WOMAC pain score at 12 months postoperatively [138]. A lower educational status has been associated with better pain levels at six months in a study by Lopez-Olivo et al. (2012). However, six studies report no significant effect of education on either pain or function following TKA [52, 53, 121, 122, 131, 138].

5.3.3. Psychosocial determinants

Several studies that were included in the review were dedicated to exploring the relationship between possible psychosocial determinants and TKA outcomes measured as *postoperative status*. Presence or higher levels of anxiety and/or depression have been consistently identified as significant determinants of worse TKA outcomes in six of the included studies [127, 129, 134-136, 139]. Three studies report that greater preoperative pain catastrophizing is linked to worse pain six weeks and 12 months after TKA and to higher levels of disability 12 months postoperatively [132, 133, 140]. Escobar et al. (2007) identified absence of social support to be related to worse six-month pain and function levels [128].

Other significant psychosocial variables associated to pain and function status following TKA are presented in Table 3.

In terms of *postoperative change,* Riddle et al. (2010) determined that greater pain catastrophizing was related to higher odds of not achieving an improvement of 50% in the pain domain of the WOMAC at 6 months as well as not attaining a change greater than 4 points out of 20 on the WOMAC pain score at six months [126]. A previous diagnosis of depression and higher levels of depression/anxiety as measured by the EuroQ5D questionnaire were related to a smaller change on the six to 12 month total OKS score [120].

5.3.4. Clinical determinants

The investigation of the association between clinical characteristics and TKA outcomes measured as *postoperative status* has received a great deal of attention. One of the most studied potential determinants of knee pain and function following TKA is the baseline, preoperative levels of the respective variables. Eighteen studies linked a poor preoperative status to a worse postoperative status in terms of pain and function [19, 52, 53, 103, 121-124, 127-129, 131, 132, 134-136, 139, 141].

Poor preoperative mental health, as measured by the SF-36 questionnaire, has been associated to worse outcomes in seven studies [122, 124, 128, 130, 135, 139, 141]. Even if TKA is performed predominantly for patients with primary OA, two studies seem to point to a diagnosis of primary gonarthrosis as a determinant of worse outcomes, when compared to rheumatoid arthritis or other diagnoses [127, 129]. Higher baseline BMI has been linked to poorer functional results as well as to worse outcomes of pain and function combined in four studies [92, 124, 129, 130]. Six studies

identified the presence of back pain before surgery to be related to substandard pain and function status after TKA [123, 128, 131, 133, 135, 139]. Regardless of whether comorbidity was measured as the influence of individual comorbidities, of the number of comorbidities per patient or when considering their severity and impact on patients' life, seven studies suggests it to be a significant determinant of worse outcomes in terms of pain and function following TKA [19, 99, 124, 127, 128, 130, 139]. More symptomatic joints, including ankle, feet, toes and neck were associated with greater level of pain and worse function 12 months after TKA in two studies [123, 134]. Other significant clinical determinants of TKA outcomes measured as postoperative status can be found in Table 3.

Regarding outcomes measured as *postoperative change*, four studies report that better baseline levels of function and/or pain are related to lower levels of improvement following TKA. Jones et al. (2001) reported that lower preoperative pain was associated with smaller changes in functional abilities six months after the surgery [19]. A better preoperative total OKS score was related to a smaller change six to 12 months following TKA in two studies [120, 142]. Better preoperative function level was found to be a significant determinant of lower gains in functional abilities [12].

Greater comorbidity was shown to significantly determine lower changes in pain and functional status [11, 106, 120, 142]. In particular, Kauppila et al. (2011) showed that presence of osteoporosis was associated with a smaller level of change in function and with decreased odds of attaining the OMERACT-OARSI set of responder criteria 12 months after surgery [12]. A study by Gandhi et al. (2013) revealed that a greater level of synovial fluid levels of three inflammatory markers (TNF- α, MMP-13 and IL-6) were related to poor gains in physical function two years after TKA as measured by the WOMAC function score [125]. Other miscellaneous clinical determinants identified in the

included studies were worse general health status (as measured by the American Society of Anaesthesiology grade), presence of self-reported disability and lower self-reported general health [120], greater preoperative bodily pain [11], worse mental health [142, 143] and presence of back pain [142].

5.3.5. Surgical determinants

Only one of the included studies identified a significant surgical determinant of poor *postoperative status* as measured by pain levels at six months: cruciate-retaining implant [121]. Sullivan et al. (2011) studied the effect of surgery duration and of the identity of the surgeon on the 12-month WOMAC pain and function scales, but their analysis yielded non-significant results [132].

In terms of *postoperative change,* findings by Jones et al. (2001) indicate that cementless prosthesis is associated with a lower change in the WOMAC pain score six months after TKA [11]. A British study by Baker et al. (2012) evaluated the effect of different types of prosthesis brands on the improvement of the total OKS score 12 months following the intervention. They found that the NexGen prosthesis brand is related to greater improvements when compared to all the other brands used in their study (PFC, Genesis 2, AGC and Triathlon) [120]. The same study evaluated the effect of the type of hospital where the surgery was performed. They showed that surgeries performed at a National Health Services hospital are more likely to be associated with poor improvement than surgeries performed at an independent hospital or an Independent Sector Treatment Centre.

6. Discussion

6.1 Highlight points

Because TKA clinical results are still suboptimal in a large percentage of patients, a better knowledge of determinants of pain and function following the intervention could help improve outcomes. The aim of our study was to systematically assess the literature reporting the determinants of pain and functional outcomes following primary unilateral TKA in patients with knee OA. Thirty-four studies with a moderate-to-high mean methodological quality (80.0%, SD 12.7%) were included. Even if several significant determinants of pain and functional outcomes following TKA have been summarized by studies, no conclusions can be reached regarding the strength of the associations between significant determinants and TKA results because of heterogeneity of study methodologies and results.

6.2. Strengths and limitations of the review

The main strength of the present systematic review is the rigorousness of the inclusion criteria ensuring high quality of evidence of determinants compiled from four important databases. Moreover, focus on all types of determinants provides a comprehensive overview of all relevant variables with a significant relationship to TKA outcomes.

The main limitation is the inability to pool the results into meta-analyses, resulting in the failure to conclude on the strength of association between patient factors and TKA outcomes due to the heterogeneity of the methodologies of the included studies. Moreover, the findings of the review do not necessarily apply to all patients undergoing TKA, namely those with a diagnosis different from OA, or undergoing bilateral or revision surgery. Also, the study does not review determinants of long-

term outcomes. Finally, two studies had to be excluded because they were published in languages not mastered by the reviewers.

6.3 Main findings

It is difficult to conclude to a significant association of any demographic determinant with TKA pain or functional outcomes based on the results of the included studies. Although female gender and older age were found significant in several studies, there is an overwhelming amount of evidence pointing to either an association in an opposite direction or to no relationship at all. These findings contrast the ones by Santaguida et al. (2008) in their systematic review. We therefore suggest that according to the available evidence, patients should not be denied surgery based on gender or age.

Regarding socioeconomic determinants, greater social deprivation achieved statistical significance in both studies evaluating its association with TKA outcomes among patients in the United Kingdom [127, 129]. Patients with greater social deprivation may experience worse TKA outcomes because of an inequality in the continuity of care following discharge compared to patients with less deprivation [127]. Caution should be warranted regarding the generalizability of these findings however, as they may not apply to other countries, although the impact of social deprivation in terms of pain and function on other musculoskeletal disorders is well established [108, 109].

Psychosocial determinants with considerable evidence include the presence or a greater level of depression and/or anxiety. The previous review by Vissers et al. (2012) did not find definite evidence that supports the significance of this association. However, all of the seven studies included in our review that conclude to such a relationship were published after the aforementioned systematic review.

The causes behind the significant association are not well understood; depressed patients may be less likely to participate actively in the rehabilitation process, thus experiencing worse outcomes [122]. Greater preoperative pain catastrophizing was also significantly associated to pain and functional outcomes after TKA, a finding consistent with the review by Vissers et al. (2012). It has been suggested that pain catastrophizing is linked with neurophysiological processes related to modulation of pain, and that greater levels of catastrophizing promote sensitization to pain [133].

The greatest amount of evidence is available for clinical determinants, the frontrunner being the relationship between worse or better preoperative levels in the respective dimensions (depending on the outcome being measured as postoperative patient state or change) and pain or functional outcomes. Although studies consistently refer to this relationship as a well-known fact, to our knowledge, this is the first systematic review underlining this fact. Our findings suggest that in the case of measuring outcome as a change in status, a higher preoperative status is related to a lower chance of improvement. In the case of measuring outcome as health status postoperatively, lower preoperative status is related to worse outcome. The importance of these concepts relies in the dilemma encountered when employing this information clinically: should intervention be undertaken in patients with worse preoperative state in order to obtain greater gains or should TKA be performed in patients as early as possible before they deteriorate considerably in order to guarantee better status after the surgery? Unfortunately, there does not seem to be a consensus regarding this predicament, and our review only emphasizes its importance, as other authors have done as well [110].

Presence or greater levels of comorbidities were also related to a worse outcome after TKA. The reasons behind such a relationship are unclear. Patients with other comorbidities may not meet the

demands of the intensive rehabilitation process following TKA, thus explaining their increased risk of poor surgical outcomes [130]. Several studies advocate that patients should receive appropriate counselling from their surgeon preoperatively according to the identity and number of their comorbidities [92, 130]. Of interest, presence of back pain was associated to poor TKA outcomes as well. The mechanism behind this association is however uncertain. Back pain may impede postoperative recuperation and rehabilitation or it may directly affect how patients rate their condition in terms of knee pain and function on the WOMAC, or on other outcome measures .

Worse measures of general health were significantly related to poor TKA outcomes in a surprisingly consistent manner. Among the included studies, general health was mainly measured with the SF-36 questionnaire, and a poorer mental health domain in particular was consistently related to poor pain and function after TKA. This may underline the importance of the overall health status, especially the extent of psychological distress, in selecting individuals for knee arthroplasty.

Limited evidence has been identified regarding surgical determinants of poor TKA outcomes. This may be due to the fact that surgical factors are traditionally investigated by studies employing a clinical trial methodology, whereas this review encompasses prognostic cohort studies. Association of surgical and technical factors with TKA outcomes is clearly a complex issue, and a different approach than the one employed by this review may be required to identify significant determinants.

No variable was consistently identified as non-significantly related to TKA outcome. The amount of evidence for certain significant determinants is nevertheless countered by numerous studies stating their non-significance and this inconsistency represents a limitation of the available literature. Several

reasons behind this discrepancy can emerge. Firstly, a low sample size can impact the ability to detect a truly present statistically significant relationship; type II errors may effectively limit the findings. Also, the duration of follow-up may lead to a disagreement as a significant relationship may arise at a critical time-point following the surgery. Finally, the methodological quality of the studies can lead to heterogeneous results.

It has previously been suggested that the determinants of pain and function after TKA are not the same [129]. We attempted to appraise this by capturing the results of the included studies based on whether determinants were assessed for pain and function as separate dependent variables or part of a combined construct (Tables 3 and 4). On several occasions, individual studies that evaluated significant determinants of pain and function separately concluded that they indeed presented different determinants. However, when viewing the overall picture, the determinants of pain and function seem to be similar regardless of the method of measurement. This is most probably due to the overlap between the findings of the studies.

As mentioned previously, outcomes after TKA are generally evaluated as a function of health change or of health state postoperatively [46-49, 144]. In our review, we identified fewer studies evaluating determinants based on postoperative change. Generally, from the available evidence, determinants are similar between the two approaches, with the exception of the preoperative status as discussed previously.

7. Conclusion

Moderate-to-high methodological quality of included studies suggests that preoperative determinants of pain and function outcomes following TKA include greater social deprivation, the presence or a greater level of depression and/or anxiety, greater preoperative pain catastrophizing, preoperative pain or function levels, presence or greater levels of comorbidity, presence of back pain and lower general health. Consensus is however limited by contradictory results regarding the importance of several determinants. The heterogeneity in the measurement of the outcome limits the ability to generalize the magnitude of association of determinants with TKA outcomes. Further high-quality research and a more standardized reporting of results is required in order to elucidate with greater precision the identity of determinants of pain and function following TKA in order to provide the best possible care for patients with severe knee OA.

8. Conflict of interest

Authors declare that they have no conflict of interest.

9. Acknowledgements

EL is supported by a training award for M.D.-M.Sc. students from the FRQS.

Figure 1 Flowchart of the literature search

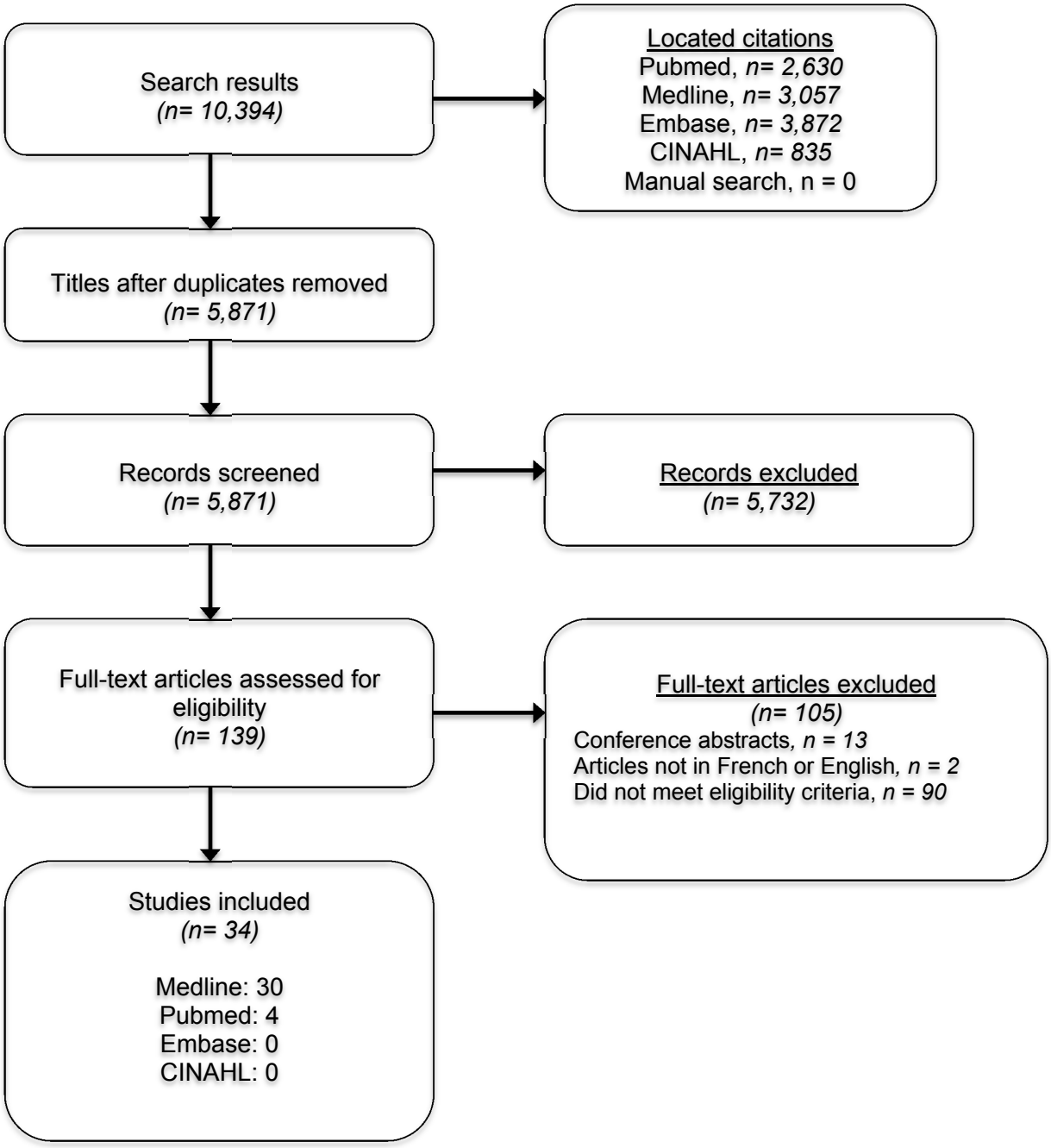


Table 1 Description of the included studies

Study	Participants				Number of patients	Follow-up period	Outcome Measure	Statistical method	Statistical adjustment	Results	
	Diagnosis	Type of Surgery	Mean Age (SD)	Gender (%female)						Significant Determinants	Non-significant Determinants
Alzahrani et al. (2011)	Primary or secondary OA	Primary unilateral TKA	Cohort A: 67.5 (9.6) Cohort B: 69.0 (9.1) Overall: 68.2 (9.4)	Cohort A: 62% Cohort B: 63% Overall: 63%	Cohort A: 457 Cohort B: 2720 Overall: 3177	1 year	No clinical improvement at 1 year: Minimal Clinically Important Difference WOMAC: ≤ 7.5/100 points OKS: ≤ 5.0/60 points	Logistic regression	Age, Gender BMI Comorbidity	Cohort A: **Increased age at time of surgery (-)**: OR 1.06 (95% CI 1.02-1.09) Cohort B: **Male gender (+)**: OR 0.72 (95% CI 0.57-0.92)	Cohort A **Gender** **BMI** **Comorbidity** Cohort B: **Age** **BMI** **Comorbidity**
Ayers et al. (2005)	OA	Primary unilateral TKA	68.0 (9.8)	62.4%	165	12 months	12 month WOMAC-Physical function score improvement (change score)	Blocked multiple regression	NONE	**Age (unclear)** **Gender (unclear)** **Worse preoperative physical function (WOMAC function) (-)** **Higher preoperative mental health (SF-36 mental component score) (+)**	NONE
Baker et al. (2012)	OA	Primary unilateral TKA	Not available	Not available	22691	6 - 12 months (median 199 days)	6 – 12 month OKS improvement (change score)	Stepwise multiple linear regression	NONE	**Higher age (+)**: estimate = 0.06 (95% CI 0.04 to 0.07) **Higher preoperative function and lower pain (-)**: estimate = -0.66 (95% CI -0.67 to -0.64) **Higher number of comorbidities (-)**: estimate = -0.25 (95% CI -0.39 to -0.12) **Presence of self-reported preoperative disability (-)** estimate = -1.49 (95% CI -1.75 to -1.23) **Very good self-reported preoperative general health (**vs. excellent) **(-):** estimate = -1.12 (95% CI -1.78 to -0.45) **Good self-reported preoperative general health (**vs. excellent) **(-):** estimate = -2.78 (95% CI -3.42 to -2.12) **Fair self-reported preoperative general health (**vs. excellent) **(-):** estimate = -5.23 (95% CI -5.93 to -4.53) **Poor self-reported preoperative general health (**vs. excellent) **(-):** estimate = - 8.13 (95% CI -9.09 to -7.16) **Presence of depression (-)**: estimate = -0.95 (95% CI -1.44 to 0.46) **Moderate anxiety/depression** (vs. no anxiety/depression) **(-)**: estimate = -1.17 (95% CI -1.45 to -0.90) **Severe anxiety/depression** (vs. no anxiety/depression) **(-)**: estimate = -2.78 (95% CI -3.48 to -2.07) **ASA Grade 3** (vs. Grade 1) **(-)**: estimate = -1.00 (95%CI -1.52 to -0.49) **PFC prosthesis brand** (vs NexGen) **(-)**: estimate = -0.98 (95%CI -1.35 to -0.62) **Genesis 2 prosthesis brand** (vs NexGen) **(-)**:	**ASA Grade 2**

										estimate = -1.50 (95%CI -2.02 to -0.98) **AGC prosthesis brand** (vs NexGen) **(-)**: estimate = -1.20 (95%CI -1.68 to -0.72) **Triathlon prosthesis brand** (vs NexGen) **(-)**: estimate = -1.74 (95%CI -2.16 to -1.36) **Independent hospital** (vs. NHS hospital) **(+)**: estimate = 0.83 (95%CI 0.35 to 1.31) **ISTC** (vs. NHS hospital) **(+)**: estimate = 1.84 (95%CI 1.23 to 2.45)	
Caracciolo et al. (2005)	OA	Primary TKA	71.6 (6.6)	81%	47	6 months	WOMAC function score at 6 months	Logistic regression	NONE	**Higher preoperative function (+)**, OR = 1.15, 95% CI = 1.04 to 1.28, compared to lower preoperative function (worst quartile of WOMAC function score)	**Preoperative osteoarthritis morbidity**: Charnley or Modified Charnley Class C
Clement et al. (2013)	Primary OA	Primary unilateral TKA	70.4 (9.4)	57.5	2392	1 year	1 year OKS score	Multivariate linear regression analysis	NONE	**Presence of back pain (-)**: β = - 2.41 (95%CI -3.18 to -1.64) **Presence of depression (-)**: β = - 4.17 (95%CI -5.42 to -2.92) **Better preoperative levels of pain and function (+)**: β = 0.45 (95%CI 0.39 to 0.51) **Higher preoperative mental health (+)**: β = 0.19 (95%CI 0.16 to 0.22)	**Gender** **Heart disease** **Hypertension** **Lung disease** **Vascular disease** **Neurological disease** **Diabetes mellitus** **Gastric ulceration** **Kidney disease** **Liver disease** **Anemia**
Clement et al. (2013)	N/A	Primary TKA	70.4	56.6	2389	1 year	1 year OKS score	Multivariate linear regression analysis	NONE	**Vascular comorbidity (-)**: β = -1.91, 95% CI -3.78 to -0.05 **Depression (-)**: β = -4.19, 95% CI -5.44 to -2.95 **Back pain (-)**: β = -2.38, 95% CI -3.14 to -1.61 **Better preoperative levels of pain and function (+)**: β = 0.45, 95% CI 0.39 to 0.51 **Higher preoperative mental health (+)**: β = 0.19 (95%CI 0.16 to 0.22)	**Heart disease** **High blood pressure** **Lung disease** **Neurological disease** **Diabetes** **Stomach ulcer** **Kidney Disease** **Liver disease** **Anemia** **Preoperative physical health**
Clement et al. (2013)	Primary OA	Primary TKA	70.6 (7.0)	57.6	966	1 year	Mean OKS improvement after 1 year	Multivariate linear regression analysis	NONE	**Presence of back pain (-)**: β = -2.53, 95% CI -3.75 –to -1.30 **More than 4 comorbidities (-)**: β = -3.78, 95% CI -6.11 to -1.45 **Higher preoperative function and lower pain (-)**: β = 0.58, 95% CI 0.50 to 0.87 **Higher preoperative mental health (+)**: β =0.16, 95% CI 0.11 to 0.22	NONE
Davis et al (2008)	Primary OA	Unilateral primary TKA	71.1 (49 to 85)*	51.0	974	24 months	3, 12, 24 month total WOMAC score and WOMAC pain and function scores	Multivariate linear regression analysis	Age, Gender Number of co-morbid conditions Country Center within country	WOMAC pain 12 months **Low income (-):** p = 0.014	WOMAC pain **Low income** at 3 and 24 months **Education status** at 3, 12 and 24 months WOMAC function **Low income** at 3, 12 and 24 months

								Preoperative status (WOMAC pain or function)			**Education status** at 3, 12 and 24 months
Desmeules et al. (2013)	OA (96%), RA (4%)	Primary unilateral TKA	67 (9.3)	66%	138	6 months	WOMAC pain and function scores at 6 months	Stepwise multiple regression analysis	NONE	WOMAC pain (r2 = 0.11) **Higher preoperative pain level (-)**: β = 0.25, 95% CI 0.08-0.41 **Cruciate retaining implant (-)**: β = -8.21, 95% CI -15.01 to -1.34	WOMAC pain **Household living status** **Initial diagnosis (OA/RA)** **BMI** **Burden of comorbidities** **Duration of the disease** **Use of walking aid** **Pain contralateral knee** **Formal education** **Employment status** **Household income** **Size of social network** **Psychological distress** **Surgical variables – bearing type, patella resurfacing** **Marital status** **Occupational status**
										WOMAC function (r2 = 0.16) **Higher preoperative function level (+)**: β = 0.35, 95% CI 0.16-0.54 **Marital status (single, separated, divorced or widowed) (-)**: β = -6.84, 95% CI -14.74 to -0.95 **Occupational status (unemployed or retired) (-)**: β = -7.77, 95% CI -14.70 to -0.87	WOMAC function **Household living status** **Initial diagnosis (OA/RA)** **BMI** **Burden of comorbidities** **Duration of the disease** **Use of walking aid** **Pain contralateral knee** **Formal education** **Employment status** **Household income** **Size of social network** **Psychological distress** **Surgical variables – bearing type, patella resurfacing** **Implant type**
Engel et al. (2004)	OA	TKA	67.1 (8.3)	49.3%	74	6 months	WOMAC pain and function scores at 6 months	Multiple hierarchical regression analysis	Control of other variable (efficacy variables vs. expectancy variables) Adjustment for preoperative WOMAC variables	WOMAC pain **Higher coping efficacy (+)**: β = -0.338, p < 0.01 and **High Arthritis Helplessness (-)**: β = 0.239, p < 0.05 adjusted r2 = 0.053 Expectancy variables**: greater pessimism, grater expected chance of recovery, greater expected change in QoL**: β not given, adjusted r2 = 0.067	WOMAC pain **None**
										WOMAC function **Higher coping efficacy (+)**: β = -0.337, p < 0.05, adjusted r2 = 0.032	WOMAC function **Arthritis Helplessness** Expectancy variables**: pessimism, expected chance of recovery, expected chance in QoL**
Escobar et al. (2007)	OA	Primary TKA	71.8 (6.7)	73.6%	640	6 months	WOMAC Pain an function score at 6 months	General linear models	NONE	WOMAC pain **Higher age (+)**: Diff β = -0.24, 95% CI -.045 to – 0.03 **Presence of social support (+)**: Diff β = -5.13, 95% CI -9.31 to – 0.95 **Absence of back pain (+)**: Diff β = -5.26, 95% CI -8.24 to -2.27	WOMAC pain **Gender** **Charlson Index 1**

									Charlson Index ≥ 2 (-): Diff β = 6.50, 95% CI 2.0 to 11.0 **Higher preoperative mental health (+):** Diff β = -0.10, 95% CI -0.17 to – 0.04 **Higher preoperative pain on WOMAC (-):** Diff β = 0.26, 0.18 to 0.34	
									WOMAC function **Presence of social support (+)**: Diff β = -7.25, 95% CI -9.31 to – 0.95 **Absence of back pain (+)**: Diff β = -5.26, 95% CI -11.83 to -2.67 **Charlson Index ≥ 2 (-)**: Diff β = 6.60, 95% CI 1.70 to 11.52 **Higher preoperative mental health (+):** Diff β = -0.10, 95% CI -0.17 to – 0.03 **Lower preoperative function on WOMAC (-):** Diff β = 0.29, 0.19 to 0.38	WOMAC function **Age** **Charlson Index 1**
Fortin et al. (1999) and Fortin et al. (2002)	OA	Primary TKA	68.1± 9.1	56 %	106 and 81	6 months and 2 years	WOMAC pain and function scores at 6 months and 2 years	Multiple linear regression	NONE	WOMAC pain at 6 months: **Higher preoperative pain (WOMAC pain score) (-):** β = 0.44 ± 0.11, r2 = 0.25	WOMAC pain at 6 months: **Age** **Gender** **Center** **Education** **Comorbidity**
										WOMAC function at 6 months: **Lower preoperative function (WOMAC function score) (-):** β = 0.61 ± 0.11, r2 = 0.36 Similar results for (2002); data not shown	WOMAC function at 6 months: **Age** **Gender** **Center** **Education** **Comorbidity**
Gandhi et al. (2010)	Primary and secondary OA	Primary unilateral TKA	66.5	61	889	1 year	WOMAC total score at 1 year	Linear regression modelling	Age Gender Baseline total WOMAC score Comorbidity (excluding hypertension, hypercholesterolemia and diabetes)	**Obesity (BMI > 30) (-):** β = 3.6, 95% CI 0.02 to 7.2	**Number of metabolic syndrome risk factors** **Hypertension** **Hypercholesterolemia** **Diabetes**
Gandhi et al. (2013)	OA	Unilateral TKA	68.5 (9.4)	57	28	2 years	Change in WOMAC pain score at 2 years	Linear regression modelling	Age Gender BMI Comorbidity count	**Greater synovial fluid TNF- α levels (-)**: p = 0.001 **Greater synovial fluid MMP-13 levels (-)**: p = 0.03 **Greater synovial fluid IL-6 levels (-)**: p = 0.001	**Serum levels of:** **IL-6** **IL-1β** **MMP-9** **MMP-13** **MIP-1β** **MCP-1** **Adiponectin** **Leptin** **TNF- α** **IFN-γ** **VCAM-1** **Synovial fluid levels of:** **IL-1β** **MMP-9** **MIP-1β** **MCP-1** **Adiponectin**

											Leptin **IFN-γ** **VCAM-1**
Hanusch et al. (2014)	OA	Primary TKA	71 (42 to 92)	45	100	6 weeks and 1 year	6 week and 1 year OKS score	Stepwise multiple linear regression	NONE	6 weeks Model 1 **Higher illness coherence** (stronger belief about patient's understanding of their illness) **(+)**: β = -0.224, p = 0.02	6 weeks Model 1 **Age** **Gender** **Preoperative OKS score** **Consequences** (patient's beliefs about impact of illness on their life **Emotional representation** (patient's negative emotions caused by their illness)
										Model 2 **Better preoperative function and lower pain (+)**: β = -0.257, p = 0.03	Model 2 **Age** **Gender** **Depression**
										1 year Model 1 **None**	1 year Model 1 **Age** **Gender** **Preoperative OKS score** **Consequences** (patient's beliefs about impact of illness on their life **Emotional representation** (patient's negative emotions caused by their illness)
										Model 2 **Better preoperative function and lower pain (+)**: β = -0.296, p = 0.008 **Higher anxiety (-)**: β = 0.270, p = 0.01	Model 2 **Age** **Gender**
										Model 3 **Better preoperative function and lower pain (+)**: β = -0.239, p = 0.04 **Higher anxiety (-)**: β = 0.296, p = 0.01	Model 3 **Age** **Gender**
Jones et al. (2001)	Osteoarthritis (93%)	Primary unilateral TKA	70.6	59	257	6 months	Change in WOMAC pain and function scores at 6 months	Multiple linear regression models	NONE	Change in pain **Higher preoperative bodily pain (SF-36) (-)**: β = -0.42, 95% CI -0.56 to -0.27 **Cementless prosthesis (-):** β = -9.48, 95% CI -16.20 to -2.77	Change in pain **Age** **Gender** **Waiting time** **Number of comorbid conditions**
										Change in function **Lower preoperative joint pain (WOMAC) (-)**: β = -0.43, 95% CI -0.57 to -0.28 **Higher number of comorbid conditions (-)**: β = -1.56, 95% CI -2.74 to -0.37 **Higher preoperative bodily pain (SF-36) (-):** β = -0.21, 95% CI -0.35 to -0.07	Change in function **Age** **Gender** **Waiting time** **BMI** **Contralateral joint involvement** **Living alone**
Jones et al. (2003)	OA (94%)	Primary TKA	69.2 (9.2)	59	273	6 months	WOMAC function score	Multiple linear regression	NONE	**Older age (+)**: β = 0.35, 95% CI 0.10 to 0.60 **Higher preoperative function (WOMAC) (+)**:	**Gender**

							at 6 months			β = 0.30, 95% CI 0.16 to 0.43 **Greater number of comorbid conditions (-)**: β = -1.62, 95% CI -2.75 to -0.49 **Use of walking devices preoperatively (-):** β = -4.15, 95% CI -7.23 to -1.06	
Judge et al (2012)	OA (93.7%)	Primary TKA (92%) UKA (8%)	71.7 (9.1)	61	1991	6 months	**Model 1:** Total OKS, OKS pain score and OKS function score at 6 months **Model 2:** PASS score for Total OKS, OKS pain score and OKS function score at 6 months	**Model 1:** Multiple linear regression **Model 2:** Logistic regression	**Model 1:** NONE **Model 2:** NONE	**Model 1:** Total OKS **Higher baseline OKS (+)**: multivariable coefficient = 1.70, 95% CI 1.43 to 1.96 **Female sex (-)**: multivariable coefficient = -0.88, 95% CI -1.68 to -0.08 **Higher preoperative BMI (-)**: multivariable coefficient = -0.44, 95% CI -0.86 to -0.01 **Greater social deprivation** (higher log of Index of Multiple Deprivation-IMD 2004) **(-)**: multivariable coefficient = -1.40, 95% CI -1.96 to -0.85 **RA diagnosis** (vs. Primary OA) **(+)**: multivariable coefficient = 2.90, 95% CI 0.42 to 5.37 **Moderately anxious/depressed** (vs. Not Anxious/Depressed) **(-)**: multivariable coefficient = -0.85, 95% CI -1.68 to -0.03 **Extremely anxious/depressed** (vs. Not Anxious/Depressed) **(-)**: multivariable coefficient = -2.21, 95% CI -4.34 to -0.09 OKS pain score **Higher baseline OKS (+)**: multivariable coefficient = 1.30, 95% CI 1.03 to 1.57 **Greater social deprivation** (higher log of IMD 2004)**(-)**: multivariable coefficient = -0.64, 95% CI -0.91 to -0.37 **RA diagnosis** (vs. Primary OA) **(+)**: multivariable coefficient = 1.75, 95% CI 0.61 to 2.89 **Moderately anxious/depressed** (vs. Not Anxious/Depressed) **(-)**: multivariable coefficient = -0.43, 95% CI -0.83 to -0.03 **Extremely anxious/depressed** (vs. Not Anxious/Depressed) **(-)**: multivariable coefficient = -1.19, 95% CI -2.19 to -0.18 OKS function score **Higher baseline OKS (+)**: multivariable coefficient = 1.82, 95% CI 1.58 to 2.06 **Higher age (-)**: multivariable coefficient = -0.21, 95% CI -0.34 to -0.08 **Female sex (-)**: multivariable coefficient = -0.79, 95% CI -1.25 to -0.33 **Higher preoperative BMI (-)**: multivariable coefficient = -0.33, 95% CI -0.57 to -0.09 **Greater social deprivation** (higher log of IMD 2004)**(-)**: multivariable coefficient = -0.79, 95% CI -1.11 to -0.46 **Model 2:** PASS total OKS score **Higher baseline OKS (+)**: OR = 1.52, 95% CI 1.40 to 1.66	Total OKS **Age** **Operated side** **Diagnosis other than OA or RA** **ASA grade** **Year surgery was performed** OKS pain score **Age** **Gender** **Preoperative BMI** **Operated side** **Diagnosis other than OA or RA** **ASA grade** **Year surgery was performed** OKS function score **Operated side** **Diagnosis other than OA or RA** **RA diagnosis** **ASA grade** **Anxiety/depression level** **Year surgery was performed** **Model 2:** PASS total OKS score **Age** **Gender**

										Greater social deprivation (higher log of Index of Multiple Deprivation 2004) **(-)**: OR = 0.73, 95% CI 0.62 to 0.87 **RA diagnosis (+)**: OR = 2.17, 95% CI 1.02 to 4.60	**BMI** **Operated Side** **Diagnosis other than OA or RA** **ASA grade** **Anxiety/depression level** **Year surgery was performed**
										PASS OKS pain score **Higher baseline OKS (+)**: OR = 1.81, 95% CI 1.52 to 2.17 **Greater social deprivation** (higher log of Index of Multiple Deprivation 2004) **(-)**: OR = 0.80, 95% CI 0.68 to 0.94 **RA diagnosis (+)**: OR = 2.33, 95% CI 1.03 to 5.29 **Moderately anxious/depressed** (vs. not anxious/depressed) **(-)**:OR = 0.67, 95% CI 0.54 to 0.84 **Extremely anxious/depressed** (vs. not anxious/depressed) **(-)**:OR = 0.51, 95% CI 0.31 to 0.84	PASS OKS pain score **Age** **Gender** **BMI** **Operated Side** **Diagnosis other than OA or RA** **ASA grade** **Year surgery was performed**
										PASS OKS function score **Higher baseline OKS (+)**: OR = 2.08, 95% CI 1.82 to 2.39 **Older age (-)**:OR = 0.93, 95% CI 0.87 to 0.99 **Greater social deprivation** (higher log of Index of Multiple Deprivation 2004) **(-)**: OR = 0.76, 95% CI 0.64 to 0.89 **Moderately anxious/depressed** (vs. not anxious/depressed) **(-)**:OR = 0.77, 95% CI 0.61 to 0.97	PASS OKS function score **Gender** **BMI** **Operated Side** **Diagnosis other than OA or RA** **ASA grade** **Extremely anxious/depressed** (vs. not anxious/depressed) **Year surgery was performed**
Kauppila et al. (2011)	OA	Primary TKA	70.7 (5.5)	75	88	12 months	12 month WOMAC function change score OMERACT-OARSI responder criteria	Multiple linear regression Multivariate logistic regression	NONE	Multiple linear regression **Male gender (-)**: β = -12.0, 95% CI -23.1 to -0.9 **Presence of osteoporosis (-)**: β = -17.5, 95% CI -32.9 to -2.1 **Higher preoperative function (-)**: β = 0.31, 95% CI 0.06 to 0.56 Multivariate logistic regression **Presence of osteoporosis (-)**: OR = 14.7, 95% CI 1.1 to 106.1	Multiple linear regression **Age** **Preoperative function of the opposite knee** Multivariate logistic regression Data not shown
Lingard et al. (2004)	OA	Primary TKA	69.9	59.2	860	12 and 24 months	WOMAC pain and function at 12 and 24 months	Hierarchical linear modelling	NONE	WOMAC pain At 12 months **Female gender (-)**: F = 7.06, p < 0.05, parameter estimate = -3.77, 95% CI -6.55 to -0.99 **Lower preoperative pain (WOMAC pain score) (-)**: F = 29.16, p < 0.0005, parameter estimate = 0.20, 95% CI 0.13 to 0.28 **Lower preoperative mental health (SF-36 mental health score) (-)**: F = 17.53, p < 0.0005, parameter estimate = 0.16, 95% CI 0.09 to 0.24 **More comorbid conditions (-)**: F = 5.85, p <	WOMAC pain At 12 months **Age** **Country**

										0.05, parameter estimate = -1.33, 95% CI -2.41 to -0.25 At 24 months **Female gender (-)**: F = 3.98, p < 0.05, parameter estimate = -2.98, 95% CI -5.91 to -0.05 **Lower preoperative pain (WOMAC pain score) (-)**: F = 25.13, p < 0.0005, parameter estimate = 0.20, 95% CI 0.12 to 0.28 **Lower preoperative mental health (SF-36 mental health score) (-)**: F = 9.53, p < 0.005, parameter estimate = 0.13, 95% CI 0.05 to 0.21 **More comorbid conditions (-)**: F = 4.59, p < 0.05, parameter estimate = -1.24, 95% CI -2.38 to -0.11 WOMAC function At 12 months **Higher age (-)**: F = 5.62, p < 0.05, parameter estimate = -0.19, 95% CI -0.35 to -0.03 **Lower preoperative function (WOMAC function score) (-)**: F = 51.58, p < 0.0005, parameter estimate = 0.30, 95% CI 0.22 to 0.38 **Lower preoperative mental health (SF-36 mental health score) (-)**: F = 17.04, p < 0.0005, parameter estimate = 0.17, 95% CI 0.09 to 0.25 **Higher BMI (-)**: F = 4.70, p < 0.05, parameter estimate = -0.30, 95% CI -0.57 to -0.03 **More comorbid conditions (-)**: F = 11.96, p < 0.005, parameter estimate = -1.95, 95% CI -3.05 to -0.84 At 24 months **Lower preoperative function (WOMAC function score) (+)**: F = 55.75, p < 0.0005, parameter estimate = 0.34, 95% CI 0.25 to 0.43 **Lower preoperative mental health (SF-36 mental health score) (-)**: F = 6.02, p < 0.05, parameter estimate = 0.11, 95% CI 0.02 to 0.20 **Restricted knee flexion (-)**: F = 6.04, p < 0.05, parameter estimate = 0.12, 95% CI 0.02 to 0.21 **More comorbid conditions (-)**: F = 13.96, p < 0.0005, parameter estimate = -2.26, 95% CI -3.45 to -1.07	At 24 months **Age** WOMAC function At 12 months **Gender** At 24 months **Age** **Gender**
Lingard et al. (2007)	OA	Primary TKA	70.8	60.3	952	3, 12 and 24 months	WOMAC pain and function at 3, 12 and 24 months	General linear models	Age Gender Number of comorbidities Country Center within country Preoperative scores	*With substitution of missing values* WOMAC pain **Higher preoperative mental health (SF-36) (+)** At 3 months: parameter estimate = 0.095, p = 0.0143 At 12 months: parameter estimate = 0.128, p = 0.0008 At 24 months: parameter estimate = 0.096, p = 0.0109 WOMAC function	

										Higher preoperative mental health (SF-36) (+) At 3 months: parameter estimate = 0.129, p = 0.0001 At 12 months: parameter estimate = 0.150, p = 0.0001 At 24 months: parameter estimate = 0.106, p = 0.0071	
Lopez-Olivo et al. (2011)	OA	Primary TKA	65 (9)	65	232	6 months	WOMAC pain and function at 6 months	Multiple regression modelling	NONE	WOMAC pain **More education (+)**: β = -0.17, p =0.01 **More comorbidities (-)**: β = 0.17, p = 0.008 **More problem solving-style coping (+)**: β = -0.14, p =0.03 **More dysfunctional coping (-)**: β = 0.13, p =0.04 **More internal belief of control over health (+)**: β = -0.14, p =0.02 WOMAC function **More frequent availability of tangible support (+)**: β = - 0.15, p =0.01 **Worse depressive state (-)**: β = 0.15, p = 0.02 **More problem solving-style coping (+)**: β = -0.20, p =0.001 **Lower baseline function level (-)**: β = 0.25, p =0.0001	WOMAC pain **BMI** **Baseline pain level** WOMAC function **BMI** **Comorbidities**
Neuburger et al. (2013)	OA (90%)	Primary TKA (95%) Revision TKA (5%)	N/A	57%	62,303	6 months	Total OKS score at 6 months	Logistic regression analysis	**Model 1:** Age Sex Ethnicity Self-reported comorbid conditions Self-reported general health Primary OA Primary TKA or revision TKA Hospital	**Model 1:** **More social deprivation** (2nd quintile vs. 1st quintile) **(-)**: β = -0.7, 95% -0.9 to -0.5 **More social deprivation** (3rd quintile vs. 1st quintile) **(-)**: β = -1.1, 95% -1.3 to -0.9 **More social deprivation** (4th quintile vs. 1st quintile) **(-)**: β = -2.2, 95% -2.4 to -2.0 **More social deprivation** (5th quintile vs. 1st quintile) **(-)**: β = -3.5, 95% -3.8 to -3.3	**Model 1:** NONE
									Model 2: Age Sex Ethnicity Self-reported comorbid conditions Self-reported general health Primary OA Primary TKA or revision TKA Hospital Preoperative OKS Longstanding problems	**Model 2:** **More social deprivation** (2nd quintile vs. 1st quintile) **(-)**: β = -0.4, 95% -0.6 to -0.2 **More social deprivation** (3rd quintile vs. 1st quintile) **(-)**: β = -0.6, 95% -0.8 to -0.4 **More social deprivation** (4th quintile vs. 1st quintile) **(-)**: β = -1.5, 95% -1.8 to -1.3 **More social deprivation** (5th quintile vs. 1st quintile) **(-)**: β = -2.4, 95% -2.7 to -2.2	**Model 2:** NONE
									Model 3: NONE	**Model 3:** **Age < 51 years** (vs. 71-80 years) **(-)**: β = -2.9, 95% -3.4 to -2.4 **Age 51-60 years** (vs. 71-80 years) **(-)**: β = -1.6, 95% -1.8 to -1.3	**Model 3:** **Gender** **Age 61-70 years** (vs. 71-80 years) **Lung disease** **Cancer**

										Age > 80 years (vs. 71-80 years) **(-)**: β = -0.5, 95% -0.7 to -0.2 **South-Asian, black or other ethnicity** (vs. white ethnicity) **(-)**: β = -2.5, 95% -2.9 to -2.2 **Heart disease (-)**: β = -0.6, 95% -0.8 to -0.3 **High blood pressure (+)**: β = 0.3, 95% 0.2 to 0.5 **Stroke (-)**: β = -0.9, 95% -1.5 to -0.3 **Poor circulation (-)**: β = - 2.3, 95% -2.6 to -2.0 **Diabetes (-)**: β = -0.7, 95% -1.0 to -0.5 **Depression (-)**: β = - 1.8, 95% -2.1 to -1.5 **Very good general health** (vs. excellent) **(-)**: β = -1.2, 95% -1.6 to -0.9 **Good general health** (vs. excellent) **(-)**: β = -3.6, 95% -4.1 to -3.3 **Fair general health** (vs. excellent) **(-)**: β = -7.3, 95% -7.7 to -6.9 **Poor general health** vs. excellent) **(-)**: β = -11.0, 95% -11.6 to -10.4 **Diagnosis of OA (-)**: β = -0.5, 95% -0.9 to -0.3 **Revision operation (-)**: β = -6.3, 95% -6.7 to -5.8 **Longstanding problems (+)**: β = 0.4, 95% 0.3 to 0.6 **Better preoperative pain/function** (2nd decile of preoperative OKS score vs. 1st decile – lowest) **(+):** β = 2.5, 95% 2.0 to 2.9 **Better preoperative pain/function** (3rd decile of preoperative OKS score vs. 1st decile – lowest) (vs. 1st decile - lowest **(+):** β = 3.9, 95% 3.5 to 4.2 **Better preoperative pain/function** (4th decile of preoperative OKS score vs. 1st decile – lowest) **(+):** β = 4.8, 95% 4.5 to 5.2 **Better preoperative pain/function** (5th decile of preoperative OKS score vs. 1st decile – lowest) **(+):** β = 5.6, 95% 5.3 to 6.0 **Better preoperative pain/function** (6th decile of preoperative OKS score vs. 1st decile – lowest) **(+):** β = 6.7, 95% 6.4 to 7.0 **Better preoperative pain/function** (7th decile of preoperative OKS score vs. 1st decile – lowest) **(+):** β = 7.3, 95% 7.0 to 7.7 **Better preoperative pain/function** (8th decile of preoperative OKS score vs. 1st decile – lowest) **(+):** β = 8.2, 95% 7.9 to 8.5 **Better preoperative pain/function** (9th decile of preoperative OKS score vs. 1st decile – lowest) **(+):** β = 9.2, 95% 8.8 to 9.5 **Better preoperative pain/function** (10th decile of preoperative OKS score vs. 1st decile – lowest) **(+):** β = 10.5, 95% 10.2 to 10.8	
Papakostidu et al. (2012)	OA (96%)	Primary TKA	69.17 (6.69)	79.4	204	12 months	WOMAC pain and function at 12 months	General linear modelling	NONE	WOMAC pain **Higher pre-intervention pain (-)**: Diff = 0.10, 95% CI 0.02 to 2.29 WOMAC function	WOMAC pain **Gender** **Age** **BMI** **Education** **Social support** **Residence** WOMAC function

									Lower pre-intervention function (-): Diff = 0.17, 95% CI 0.06 to 0.28	**Gender** **Age** **BMI** **Education** **Social support** **Residence**
Perruccio et al. (2012)	OA	Primary unilateral TKA	65	65	494	12 months	WOMAC pain and function at 12 months	Multiple linear regression	NONE	Pain **Symptomatic ankles/feet/toes (-):** β = 1.24, 95% CI 0.48 to 2.00 **Symptomatic neck (-):** β = 1.07, 95% CI 0.17 to 1.98 **Higher pre-surgery knee pain (-):** β = 0.34, 95% CI 0.24 to 0.45	Pain **Age** **Gender** **Education** **BMI** **Comorbidity count** **Symptomatic contralateral knee** **Symptomatic hips** **Symptomatic elbows/wrists/hands** **Symptomatic shoulder** **Symptomatic spine/lower back**
										Physical function **Symptomatic ankles/feet/toes (-):** β = 3.14, 95% CI 0.69 to 5.59 **Symptomatic neck (-):** β = 3.46, 95% CI 0.54 to 6.38 **Higher Pre-surgery knee function (+):** β = 0.41, 95% CI 0.31 to 0.50	Physical function **Age** **Gender** **Education** **Overweight** **BMI** **Comorbidity count** **Symptomatic contralateral knee** **Symptomatic hips** **Symptomatic elbows/wrists/hands** **Symptomatic shoulder** **Symptomatic spine/lower back**
Rajgopal et al. (2008)	OA	Primary TKA (7.1% with history of contralateral TKA)	N/A	59.3	550	1 year	Total WOMAC score at 1 year	Multiple linear regression	NONE	**Higher baseline mental health (+):** β = 0.210, 95% CI 0.063 to 357 **Charnley Class C (-):** β = -4.897, 95% CI -8.701 to -1.093 **Higher baseline WOMAC score (+):** β = 0.301, 95% CI 0.202 to 0.399 **BMI ≥40 (-)**: β = -5.188, 95% CI -9.771 to -0.606	**Age** **Gender** **Prior contralateral TKA** **BMI**
Ramaesh et al. (2013)	Arthrosis	TKA	70.5	58	205	1 year	Oxford Knee Score at 1 year	Multiple linear regression	NONE	**More comorbidity (-)**: B = -1.77, 95% CI -2.35 to -1.19 **Better preoperative function/pain level (+)**: B = 0.26, 95% CI 0.10 to 0.43	**Age** **Gender** **Personality type**
Riddle et al. (2010)	OA	Primary TKA	63.7	70.7	157	6 months	WOMAC pain and function change scores at 6 months	Logistic regression	Age Gender BMI Comorbidity Rheumatoid arthritis status Race/ethnicity Preoperative WOMAC pain score	WOMAC pain score Model 1: change by <50% **Greater pain catastrophizing** (PCS score ≥ 16) **(-)**: OR = 2.67, 95% CI 1.2 to 6.1 Model 2: change ≤ 4 points **Greater pain catastrophizing** (PCS score ≥ 16) **(-)**: OR = 6.04, 95% CI 1.75 to 20.82 WOMAC function score	WOMAC pain score Model 1 **None** Model 2 **Self-efficacy** **Kinesiophobia** WOMAC function score

									Model 1: change by <50% **None** Model 2: change ≤ 15 points **None**	Model 1: change by <50% **None** Model 2: change ≤ 15 points **None**
Smith et al. (2004)	OA	Primary TKA	67.2 (8.3)	52	64	6 months	WOMAC pain and function at 6 months	Multiple linear regression analysis	Gender Education Pre-surgery health measure	**NONE**	WOMAC Pain: **Optimism** **Pessimism** **Emotionality** **Purpose in life** WOMAC Function: **Optimism** **Pessimism** **Emotionality** **Purpose in life**
Sullivan et al. (2009)	OA	Primary TKA	68.6	61.3	75	6 weeks	WOMAC pain and function at 6 weeks	Multiple linear regression	NONE	WOMAC pain **Female sex (-)**: β = 0.20, p < 0.05 **Greater preoperative pain (-)**: β = 0.28, p < 0.05 **Greater preoperative pain catastrophizing (-)**: β = 0.28, p < 0.05 WOMAC function **Comorbid back pain (-)**: β = 0.23, p < 0.05, r2 = 0.05	WOMAC pain **Age** **Comorbid back pain** **Kinesiophobia** **Depression** WOMAC function **Age** **Sex** **Preoperative function** **Preoperative pain catastrophizing** **Kinesiophobia** **Depression**
Sullivan et al. (2011)	OA	Primary TKA	67	60.8	120	12 months	WOMAC pain and function at 12 months	Hierarchical regression analysis	NONE	WOMAC pain **Greater preoperative pain catastrophizing (-)**: β = 0.27, p < 0.05 WOMAC function **Greater preoperative pain catastrophizing (-)**: β = 0.34, p < 0.01	WOMAC pain **Preoperative pain** **Preoperative function** **Age** **Sex** **BMI** **Comorbidites** **Surgery Duration** **Surgeon** **Kinesiophobia** **Depression** WOMAC function **Preoperative pain** **Preoperative function** **Age** **Sex** **BMI** **Comorbidites** **Surgery Duration** **Surgeon** **Kinesiophobia** **Depression**
Wylde et al. (2012)	OA	Primary TKA	70(9)	62	220	1 year	WOMAC pain and function at 1 year	Ordinary least square regression	NONE	WOMAC pain **Higher preoperative anxiety (-)**: Unstandardized regression coefficient = 1.082,	WOMAC pain **Age** **Gender**

										95% CI 0.283 to 1.881 **Higher preoperative pain severity (-)**: Unstandardized regression coefficient = 0.183, 95% CI 0.034 to 0.331 WOMAC function **Worse self efficacy (-)**: Unstandardized regression coefficient = -0.256, 95% CI -0.478 to -0.034 **More painful joints elsewhere (-)**: Unstandardized regression coefficient = 1.928, 95% CI 0.634 to 3.222 **Higher preoperative anxiety (-)**: Unstandardized regression coefficient = 0.867, 95% CI 0.128 to 1.623 **Worse preoperative function level (-)**: Unstandardized regression coefficient = 0.289, 95% CI 0.134 to 0.444	**Other painful joints** **Number of comorbidities** **Depression** **Self efficacy** WOMAC function **Age** **Gender** **Depression**
Yakobov et al. (2014)	OA	Primary TKA	67 (range 50 to 85)	61	116	1 year	WOMAC pain and function at 1 year	Hierarchical regression analysis	NONE	WOMAC pain **Higher perceived injustice** (-): $\beta = 0.29$, $p < 0.01$ WOMAC function **Greater preoperative pain catastrophizing (-)**: $\beta = 0.26$, $p < 0.01$	WOMAC pain **Age** **Sex** **BMI** **Illness duration** **Preoperative pain** **Number of comorbid health conditions** **Pain catastrophizing** **Kinesiophobia** WOMAC function **Age** **Sex** **BMI** **Illness duration** **Preoperative pain** **Number of comorbid health conditions** **Kinesiophobia** Perceived injustice

(+): determinant of successful outcome; (-): determinant of poor outcome; ASA –American Society of Anaestesiologists; ISTC – Independent Sector Treatment Centre; NHS - National Health Services; OR – Odds Ratio; QoL – quality of life; RA – Rheumatoid arthritis; UKA – unicompartmental knee arthroplasty; β – regression coefficient

Table 2 Methodological appraisal of the included studies

Included studies (n= 34)	Study participation	Study attrition	Prognostic factor measurement	Outcome measurement	Confounding measurement and account	Analysis	Total score /12
Alzharani et al. (2011)	2	1	2	2	2	2	11
Ayers et al. (2005)	2	1	2	2	0	2	9
Baker et al. (2012)	1	0	2	2	0	2	7
Caracciolo et al. (2005)	2	1	2	2	0	2	9
Clement et al. (2013)	1	1	2	2	0	2	8
Clement et al. (2013)	2	2	2	2	0	2	10
Clement et al. (2013)	1	1	2	2	0	2	8
Davis et al. (2008)	2	2	2	2	0	2	10
Desmeules et al. (2013)	2	2	2	2	2	2	12
Engel et al. (2004)	1	2	2	2	0	2	9
Escobar et al. (2007)	2	0	2	2	2	2	10
Fortin et al. (1999)9 & (2002)	2	0	2	2	0	2	8
Gandhi et al. (2010)	2	1	2	2	2	2	10
Gandhi et al. (2013)	2	2	2	2	2	2	12
Hanusch et al. (2013)	2	2	2	2	0	2	10
Jones et al. (2001)	2	1	2	2	2	2	11
Jones et al. (2003)	2	0	2	2	2	2	10
Judge et al. (2012)	2	0	2	2	2	2	10
Kauppila et al. (2011)	2	2	2	2	0	2	10
Lingard et al. (2004)	2	0	2	2	2	2	10
Lingard et al. (2007)	2	0	2	2	0	2	8

Lopez-Olivo et al. (2011)	2	1	2	2	2	2	11
Neuburger et al. (2013)	1	0	2	2	0	2	7
Papakostidou et al. (2012)	2	1	2	2	2	2	11
Perruccio et al. (2012)	2	2	2	2	2	2	12
Rajgopal et al. (2008)	2	1	2	2	2	2	11
Ramaesh et al. (2013)	1	2	2	2	0	2	9
Riddle et al. (2010)	2	2	2	2	2	2	12
Smith et al. (2004)	1	0	2	2	0	2	7
Sullivan et al. (2009)	1	0	2	2	0	2	7
Sullivan et al. (2011)	2	0	2	2	2	2	10
Wylde et al. (2012)	1	2	2	2	0	2	9
Yakobov et al. (2014)	1	0	2	2	2	2	9
TOTAL (mean ± SD) /12	1.70 ± 0.47	0.97 ± 0.84	2.00 ± 0.0	2.00 ± 0.0	0.97 ± 1.02	2.00 ± 0.0	9.61 ± 1.52
TOTAL (mean ± SD) /100	**85.0 ± 23.5**	**48.5 ± 42.0**	**100 ± 0.0**	**100 ± 0.0**	**48.5 ± 51.0**	**100 ± 0.0**	**80.0 ± 12.7**

Study participation: The study sample represents the population of interest with regard to key characteristics, sufficient to limit potential bias to the results
Study attrition: Loss to follow-up is unrelated to key characteristics (that is, the study data adequately represent the sample), sufficient to limit potential bias
Prognostic factor measurement: The prognostic factor of interest is adequately measured in study participants, sufficient to limit potential bias
Outcome measurement: The outcome of interest is adequately measured in study participants, sufficient to limit bias
Confounding measurement and account: Important potential confounders are appropriately accounted for, limiting potential bias with respect to the prognostic factor of interest
Analysis: The statistical analysis is appropriate for the design of the study, limiting potential for the presentation of invalid results
SD: standard deviation

Table 3 Significant preoperative determinants of poor outcomes as measured by pain and/or function status at 6 weeks to 2 years following TKA

Determinant type	**PAIN**	*Studies*	**FUNCTION**	*Studies*	**PAIN & FUNCTION combined**	*Studies*
Demographic	**Younger age**	[128]	**Younger age**	[19]	**Female gender**	[127, 129]
	Female gender	[130, 133]	**Older age**	[129, 130]	**South-Asian, black or other non-white ethnicity**	[127]
			Female gender	[129]	**Younger age**	[127]
			Single, separated or divorced	[121]	**Older age**	[127]
Socioeconomic	**Low income**	[138]	**Unemployed or retired**	[121]	**Greater social deprivation**	[127, 129]
	Greater social deprivation	[129]	**Greater social deprivation**	[129]		
	Lower education level	[122]				
Psychosocial	**Lower coping efficacy**	[145]	**Lower coping efficacy**	[145]	**Presence of depression**	[127, 129, 135, 139]
	High arthritis helplessness	[145]	**Absence of social support**	[128]	**Lower illness coherence**	[136]
	Higher pessimism	[145]	**Higher anxiety**	[129, 134]	**Higher anxiety level**	[129, 136]
	Lower expected chance of recovery	[145]	**Higher depression level**	[122, 129]		
	Lower expected change in quality of life	[145]	**Less frequent availability of tangible support**	[122]		
	Absence of social support	[128]	**Less problem-solving coping style**	[122]		
	Higher anxiety level	[129, 134]	**Greater pain catastrophizing**	[132, 140]		
	Higher depression level	[129]	**Worse self-efficacy**	[134]		
	Less problem-solving coping style	[122]				
	More dysfunctional coping	[122]				
	Less internal belief of control over health	[122]				
	Greater pain catastrophizing	[132, 133]				
	Greater perceived injustice	[140]				
Clinical	**Worse pain level**	[52, 53, 99, 121, 128-131, 133, 134]	**Worse function level**	[19, 52, 53, 103, 121-123, 128-131, 133, 134]	**Presence of back pain**	[135, 139]
	Presence of back pain	[128-130]			**Worse pain/function levels**	[99, 124, 127, 129, 136, 139]
	Greater comorbidity	[122, 128-130, 141]	**Presence of back pain**	[128, 133]	**Worse mental health**	[124, 135, 139]
	Worse mental health	[129]	**Greater comorbidity**	[19, 103, 128, 130]	**Worse general health status**	[127]
	OA diagnosis	[123]	**Worse mental health**	[19, 128, 130, 141]	**Vascular comorbidity**	[127, 130]
	Symptomatic ankles/feet/toes	[123]	**Use of walking devices**	[19]	**Obesity**	[92]
	Symptomatic neck	[123]	**Higher BMI**	[129, 130]	**Higher BMI**	[124, 129]
			Symptomatic ankles/feet/toes	[123]	**OA diagnosis**	[127, 129]
			Symptomatic neck	[123]	**Greater comorbidity**	[99, 124]
			More painful joints elsewhere	[134]	**Heart disease**	[127]
			Restricted knee flexion	[130]	**Absence of high blood pressure**	[127]
					Stroke	[127]
					Diabetes	[127]
					Revision surgery	[127]
					Shorter duration of symptoms	[127]
Surgical	**Cruciate-retaining implant**	[121]	NONE	NONE	NONE	NONE
Healthcare-related	NONE	NONE	NONE	NONE	NONE	NONE

Table 4 Significant preoperative determinants of poor outcomes as measured by change in pain and/or function status 6 weeks to 2 years following TKA

Determinant type	**PAIN**	*Studies*	**FUNCTION**	*Studies*	**PAIN & FUNCTION combined**	*Studies*
Demographic	NONE	NONE	**Male gender**	[12]	**Younger age** **Older age** **Female gender**	[120] [137] [137]
Socioeconomic	NONE	NONE	NONE	NONE	NONE	NONE
Psychosocial	**Greater pain catastrophizing**	[126]	NONE	NONE	**Presence of depression** **Greater depression level**	[120] [120]
Clinical	**Greater levels of inflammatory markers** • **Greater synovial fluid TNF- α levels** • **Greater synovial fluid MMP-13 levels** • **Greater synovial fluid IL-6 levels** **Higher bodily pain levels**	[125] [11]	**Better pain level** **Greater comorbidity** **Higher bodily pain levels** **Worse mental health** **Presence of osteoporosis** **Better function level**	[11] [11] [11] [143] [12] [12]	**Better pain/function levels** **Worse general health status** **Greater comorbidity** **Presence of self-reported disability** **Lower self-reported general health** **Presence of back pain** **Worse mental health** **Presence of osteoporosis**	[120, 142] [120] [120, 142] [120] [120] [142] [12] [12]
Surgical	**Cementless prosthesis**	[11]	NONE	NONE	**PFC prosthesis brand** **Genesis 2 prosthesis brand** **AGC prosthesis brand** **Triathlon prosthesis brand**	[120] [120] [120] [120]
Healthcare-related	NONE	NONE	NONE	NONE	**Surgery performed at a National Health Services Hospital (United Kingdom)**	[120]

CHAPTER 6

Prediction of poor outcomes six months following total knee arthroplasty in patients awaiting surgery

E. Lungu[1,2], F. Desmeules[2,*], C. E. Dionne[4,5], É.L. Belzile[6], P-A. Vendittoli[2,7]

Published in BMC Musculoskeletal Disorders. 2014 September 8; 15(1):299.

[1]Department of Biomedical Sciences, Faculty of Medicine, University of Montreal, Montreal, QC, Canada
[2]Orthopaedic Clinical Research Unit - Maisonneuve-Rosemont Hospital University of Montreal affiliated Research Center, CP 6128 Succursale Centre-Ville, Montréal, Québec, Canada, H3C 3J7
[3]School of Rehabilitation, Faculty of Medicine, University of Montréal, Montreal, QC, Canada
[4]URESP, Centre hospitalier universitaire (CHU) de Québec; Quebec, QC, Canada; [5]Department of Rehabilitation, Faculty of Medicine, Laval University, Quebec, QC, Canada
[6]Centre hospitalier universitaire (CHU) de Québec, Quebec, QC, Canada
[7]Department of Surgery, Faculty of Medicine, University of Montreal, Montreal, QC, Canada

1. Abstract

Background

Identification of patients experiencing poor outcomes following TKA before the intervention could allow better case selection, patient preparation and, likely, improved outcomes. The objective was to develop a preliminary prediction rule (PR) to identify patients enrolled on surgical wait lists who are at the greatest risk of poor outcomes six months after TKA.

Methods

141 patients scheduled for TKA were recruited prospectively from the wait lists of three hospitals in Quebec City, Canada. Knee pain, stiffness and function were measured six months after TKA with the WOMAC and participants in the lowest quintile for the WOMAC total score were considered to have a poor outcome. Several variables measured at enrolment on the wait lists (baseline) were considered potential predictors: demographic, socioeconomic, psychosocial, and clinical factors including pain, stiffness and functional status measured with the WOMAC. The PR was built with recursive partitioning.

Results

The best prediction was provided by five items of the baseline WOMAC. The rule had a sensitivity of 82.1% (95% CI: 66.7-95.8), a specificity of 71.7% (95% CI: 62.8-79.8), a positive predictive value of 41.8% (95% CI: 29.7-55.0), a negative predictive value of 94.2% (95% CI: 87.1-97.5) and positive and negative likelihood ratios of 2.9 (95% CI: 1.8-4.7) and 0.3 (95% CI: 0.1-0.6) respectively.

Conclusions

The developed PR is a promising tool to identify patients at risk of worse outcomes six months after TKA as it could help improve the management of these patients. Further validation of this rule is however warranted before clinical use.

Keywords: Total Knee Arthroplasty**.** Osteoarthritis**,** Prediction rule, Determinants

2. Résumé

Introduction

L'identification préopératoire des patients à risque de résultats chirurgicaux insatisfaisants suivant une arthroplastie totale du genou (ATG) pourrait permettre une meilleure sélection des cas, une meilleure préparation des patients et, vraisemblablement, des meilleurs résultats suivant la chirurgie. L'objectif de cette étude était de développer une règle de prédiction clinique (RPC) permettant d'identifier chez des patients en attente d'un ATG, les candidats à risque de douleurs et d'incapacité plus sévères suivant l'intervention**.**

Méthodes

141 patients en attente d'une ATG ont été recrutés des listes d'attente de trois hôpitaux à Québec, Canada. La douleur, la raideur du genou ainsi que la fonction étaient mesurées six mois suivant l'ATG avec le Western Ontario and McMaster Osteoarthritis Index (WOMAC) et les participants dans le quintile le plus faible (niveau de douleurs et d'incapacités élevées) étaient considérés comme ayant un résultat défavorable. Un ensemble de variables mesurées à l'inscription sur la liste d'attente (démographiques, socio-économiques, psychosociales et cliniques) étaient considérées comme des prédicteurs potentiels. La partition récursive a été utilisée pour bâtir la RPC.

Résultats

Le meilleur modèle prédictif comprenait cinq items du WOMAC préopératoire. La règle présente une sensibilité de 82.1% (IC 95%: 66.7-95.8), une spécificité de 71.7% (IC 95%: 62.8-79.8), une valeur prédictive positive de 41.8% (IC 95%: 29.7-55.0), une valeur prédictive négative de 94.2% (IC 95%: 87.1-97.5) et des rapports de vraisemblance positifs et négatifs de 2.9 (IC 95%: 1.8-4.7) et de 0.3 (IC

95%: 0.1-0.6) respectivement.

Conclusions

La RPC est un outil prometteur pour identifier des patients à risque de résultats défavorables six mois suivant l'ATG. Cette règle pourrait améliorer la prise en charge de cette population. La validation de cet outil est nécessaire avant une implémentation et une utilisation clinique.

Mots-clés : Arthroplastie totale du genou, Arthrose, Règle de prédiction, Déterminants

3. Background

TKA surgery is widely regarded as the treatment of choice for patients suffering from knee OA once the options for conservative treatment have been exhausted [118]. TKA is the second most popular type of orthopaedic surgery and projective data suggests a six-fold increase in the number of primary TKAs in the following decades in North America [14]. TKA is an effective procedure and the majority of patients will show important improvements in pain, disabilities and HRQoL [146]. However, a growing body of evidence suggests that 10-30% of patients undergoing TKA have very poor or no improvement following surgery [4, 5, 16, 119]. Several factors that are associated to such negative outcomes have been identified; inappropriate expectations, contralateral knee pain, higher psychological distress, high body mass index, use of a walking aid, advanced age, female gender, lower OA grade and thyroid disease have all been found to be significantly associated to worse physical function following TKA [4, 11, 18, 19, 52-54, 121, 128, 133, 139, 147-152]. Nonetheless, these findings are often not consistent across studies and the exact strength of the associations between these factors and the outcomes remain elusive. It thus remains a challenge to identify which TKA candidates will likely do well, or do poorly following TKA [121].

The fact that the surgery might not be successful for many patients prompts the necessity of identifying those who are at the greatest risk of having poor outcomes following TKA. Their identification could orient both clinicians and patients regarding the decision of undertaking the procedure [18]. Moreover, medical or rehabilitation interventions could be initiated preoperatively or postoperatively [19]. Successful identification of patients at risk of adverse outcomes after TKA could not only benefit patients, but also clinicians and policy makers in more efficiently allocating necessary healthcare resources required by the condition of these patients [19, 20].

By developing an accurate and easy-to-use prediction tool, better case management of patients enrolled on a wait list for TKA could be achieved. To our knowledge, no such tool has ever been developed for this population. Few clinical predictive rules have been built and validated to allow for better case management of other types of musculoskeletal complaints. The Ottawa Ankle and Knee Rules are used in order to identify the need for roentgenographic investigation following acute ankle and knee injuries respectively [71, 72]. The Cassandra Rule has been developed to identify patients with non-specific back pain that are most likely to develop or sustain long-term functional limitations [73]. Therefore, the objective of the present study was to develop a PR that would allow a better identification of patients at the greatest risk of poor outcome six months after TKA upon enrolment on a wait list for surgery.

4. Methods

4.1. Study design

This study employed a prospective longitudinal design with repeated measures. It was part of a broader study targeted at measuring the effects of wait time on patients undergoing TKA [148, 153]. It adheres to the STROBE guidelines for observational cohort studies (see Appendix M).

4.2. Settings

From 02/2006 to 09/2007, patients newly included on the waiting lists of the departments of orthopaedic surgery of three teaching hospitals in Quebec City, Canada (CHUL, HSFA and HDQ) were recruited. Follow-up of participants ended in 09/2010 because of the extensive wait times in the participating hospitals. All seven orthopaedic surgeons performing TKA in these three hospitals collaborated in the study.

4.3. Participants

Every week, patients newly enrolled on the surgical wait lists of the three hospitals were contacted by a research nurse by phone. Eligible subjects had to meet the following inclusion criteria: (1) age $\geq$ 40 years old; (2) scheduled for primary unilateral TKA; 3) understands, reads and speaks French. Patients were excluded if they were suffering from a severe cardiac condition, a severe degenerative disease (other than OA) such as Parkinson's disease, Alzheimer's disease, any type of dystrophies or other type of sclerosis with the potential to interfere with patient recovery following TKA or any severe mental disorder (severe depression, bipolar disorder, schizophrenia or dementia) that could interfere with the ability to answer the protocol questionnaires. Subjects with a previous joint arthroplasty (hip or knee) were also excluded. Those who suffered a major trauma to the knee in the previous year or underwent surgery urgently within 30 days of registration on the waiting list were further excluded.

4.4. Data collection

Data were collected via a review of the patients' medical files and structured 45 minutes phone interviews conducted by three trained interviewers. The interviews were performed a few days after enrolment on the wait lists (mean ± SD: 12.6 ± 4.7 days) and six months after the TKA (mean ± SD: 188.7 ± 5.4 days). Patients were also interviewed before surgery; these results have been reported previously [148].

4.5. Dependent variables

Pain, stiffness and function at enrolment and six months after surgery were measured with the WOMAC, a 24-question tool [116]. The WOMAC has been found to have very good reliability, convergent construct validity and responsiveness, and has been used extensively with similar

populations [76, 77, 154]. The WOMAC score was transformed in order to obtain a score that varied from 0 to 100, 0 indicating no pain, no functional limitations nor knee joint stiffness. As there is no universal agreement on what is considered poor outcome following TKA surgery, it was defined as the last quintile of the six-month postoperative WOMAC score (i.e. WOMAC score >40.4); a satisfactory outcome was defined by a WOMAC score in the first four other quintiles of the distribution (i.e. score $\leq$ 40.4).

4.6. Independent variables

Independent variables collected to be considered as potential predictors in the final predictive model included known important determinants of TKA outcomes reported in the literature [4, 11, 18, 19, 52-54, 121, 128, 133, 147-152]. Variables were collected at the time of the patients' enrolment on the wait list and six months after TKA.

4.6.1. Potential predictors at enrolment on surgical wait list

Initial diagnosis, anthropometric data and comorbidities were recorded from the subjects' medical files. The burden of comorbidities was assessed using the CIRS [86]. At the initial interview, questions drawn from the questionnaire of the 1998 Quebec Health Survey were used to measure formal education, employment status, and household income. Social support was also measured with questions from the Quebec Health Survey [87]. Marital status, household living status, and clinical variables such as duration of disease symptoms were also noted during the initial interview. Psychological distress was recorded with a modified version of the PSI. The modified PSI includes 13 questions that measure depression and anxiety during the past week (range: 0-42) [88]. We also considered individual questions from validated questionnaires (i.e.: social support tool, PSI and

WOMAC) to build the rule. This was done in an effort to simplify the number of items to include in the final PR.

4.6.2. Other variables

Several surgical variables such as type of implant, bearing type, implant fixation, patella resurfacing and the number and type of in-hospital complications (wound infection, dislocation, knee ankylosis and manipulation, cardiovascular/ pulmonary/ circulatory complications, peripheral/ central nervous system involvement, urinary infection, acute confusion, tendon and ligament rupture, blood transfusion) following TKA were recorded by reviewing the subjects' medical files. The same procedure was used to document hospital length of stay and discharge to a rehabilitation or recovery facility. The pre-surgery wait times were calculated from the data extracted from the wait list database of each hospital. Six months following the surgery, patients were asked about walking aid usage and the number of community physiotherapy treatment hours received since discharge from the hospital.

4.7. Statistical analysis

Less than 2% of the data of the WOMAC questionnaire was missing, and it was handled according to the recommendations of the tool's guidelines [116]. Recursive partitioning analyses were used to build the PR. One of the most effective algorithm is CART [82]. It relies on considering all combinations of the predictors in order to maximize homogeneity within nodes. The Gini heterogeneity coefficient was used as a criterion to build the models [83]. Since the sample size was relatively small, we used all data in the training set. An automatic approach was first used to build PRs. Then, a set of eligible candidate predictors was created by manual adjustment based on statistical, clinical and ease of use considerations. For each resulting PR, sensitivity, specificity, area Uuder the receiver operating

characteristics – ROC – Curve (AUC), predictive value of positive and negative tests, as well as positive and negative likelihood ratios were calculated with their 95% confidence intervals [84]. The simplest rule demonstrating the highest sensitivity with acceptable level of specificity was selected as the final tool. The accuracy of the proposed model using 1,000 bootstrap resamples was then calculated for internal validation [85]. All analyses were carried out using SPSS Answer Tree 3.1 (SPSS Inc., Chicago) and SAS statistical suite software version 9.2 (SAS Institute Inc., Cary, NC, U.S.A.).

4.8 Ethics

All participants signed an informed consent form. The study was approved annually by the Research Ethics Boards of all three hospitals (CHUL, HSFA and HSFA).

5. Results

5.1. Participants

Figure 1 shows the flow of subjects through the duration of the study. A total of 588 patients enrolled on the wait list of the three hospitals. Thirty-two patients could not be reached within three weeks of inclusion on the wait list and 45 declined participation. Out of 511 patients whose eligibility was assessed, 220 patients met the eligibility criteria. Following 23 further declinations, 197 patients were interviewed at the time of enrolment on the wait list. A further six patients withdrew from the study. Six others could not be reached before surgery. Thirteen had surgery performed in a different institution. Eleven decided not to undergo surgery. The surgeries of seven patients were cancelled due to medical reasons. One patient passed away while waiting for TKA and one after the surgery was performed (both deaths unrelated to TKA). Therefore, 153 patients underwent TKA. Of these, three

withdrew from the study, one could not be reached six months after the surgery and seven underwent contralateral knee arthroplasty within six months. A total of 141 patients where thus interviewed six months after TKA. The overall eligibility proportion was calculated as (220 / 511) = 0.43; the participation proportion was calculated as (197 / (220 + ((45 + 32) x 0.43))) = 77.8%, and the follow up proportion was calculated as (141 / (197 – 44)) = 92.2%.

Participants had a mean age of 66 (SD: ± 9.5) years. The majority of patients were women (66%) and suffered from contralateral knee pain (72%). The mean wait time of the participants was 184 (SD: ± 120.8) days and median wait time was 148 days (range: 32-692). The majority of TKA implants were postero-stabilized (82%) and cemented (96%). Mean hospital length of stay was 7.5 days (SD: ± 3.0) (Table 1).

Six months following TKA, participants showed a significant improvement in terms of both pain (-30.6, SD: ± 21.8, 95% CI -26.9 to -34.2), stiffness (-26.0, SD: ± 20.4, 95% CI -21.2 to -30.8), and function mean scores (-25.4, SD: ± 20.5, 95% CI - 22.0 to -28.8), as well as in overall WOMAC mean score (-27.3, SD: ± 15.8, 95% CI -23.6 to -31.0) (Table 2).

5.2. Final prediction rule

Overall, out of the 141 participants who completed this study, 28 (20%, corresponding to the first quintile of the distribution) scored $\geq$ 40.4% (total score) on the WOMAC questionnaire, thus being identified as patients with the worst outcomes. From all potential predictors measured at enrolment, the final PR included the answers to five questions drawn from the WOMAC at baseline: preoperative difficulty of taking off socks, getting on/off toilet, performing light domestic duties and rising from

bed as well as degree of morning stiffness after the first wakening (Figure 3). The patients answered these questions in a sequential manner by attributing a degree of difficulty (none, mild, moderate, severe or extreme) to the items described in the questions. Depending on the pattern of their answers, the patients could be classified as either at risk or not at risk for poor outcomes (Figure 2).

The final PR correctly identified 23 of the 28 patients with the worst outcomes and 81 of the 113 patients with the best outcomes (Table 3). Therefore, it had a sensitivity of 82.1% (95% CI: 64.4 to 92.1), a specificity of 71.7% (95% CI: 62.8 to 79.2) and a positive likelihood ratio of 2.901 (95% CI: 2.064 to 4.077).

Presented in the appendix are other prediction models developed that were also considered (see Appendix N).

5.3. Internal validation

The accuracy of the rule was confirmed using 1,000 bootstrap re-samples. For each and every measure of predictive validity, the estimate obtained with the bootstrap was very close to the original estimate (Table 4).

6. Discussion

6.1. Main Results

In this study, a cohort of 141 patients scheduled for primary TKA were followed from the moment of their enrolment on the waiting list until six months after the surgery. The objective of the study was to develop a prediction tool that would allow the early identification of patients at risk of poor outcome

following primary TKA.

Important determinants of TKA outcomes measured at enrolment on the pre-surgery wait list were considered in the process of building the prediction rule. While the choice of the final predictive model could have been made from several criteria, we decided that the rule demonstrating the best sensitivity and an acceptable level of specificity would be the most appropriate because such a tool could identify patients at risk with fewer false negatives. Consequently, a model with a sensitivity of 82.1% and a specificity of 71.7% was chosen. Compared to published standards in clinical epidemiology, the model presented a somewhat weak positive likelihood ratio of 2.90 (95% CI: 2.06 to 4.08) [155]. Nevertheless, this rule presented the best overall predictive validity and is comparable to other PR found to be valid in the literature. For example, the positive likelihood ratio of the final PR is higher than the value reported for the Ottawa Knee Rule (2.18, 95% CI 2.04 to 2.33) designed to identify the necessity for use of radiography in the emergency room in cases of acute knee injuries [72] or for the five-item Cassandra rule allowing the identification of patients at risk of long-term back-related functional limitations (1.95, 95% CI: 1.75 to 2.17) [73].

To our knowledge, no such tool has ever been built for candidates waiting for TKA or any other type of total joint replacement. A priority-setting tool for TKA and THA has been developed and validated by the Western Canada Waiting List project partnership [23, 70]. This tool allows for the quantification of the level of urgency of the status of patients enrolled on a waiting list for total knee or hip replacement, but its predictive capabilities have not been investigated.

The best predictive model developed in this study incorporates five items from the baseline WOMAC

questionnaire, specifically questions regarding baseline function and stiffness, although an extensive set of known determinants were considered as potential predictors. The fact that four questions are related to preoperative function is consistent with the literature where preoperative function is a major determinant of postoperative function [52]. It is noteworthy to mention however that the current model does not include items regarding the level of pain. This can be due to the fact that patients undergoing TKA generally experience a notable relief in their pain level following the surgery but may still experience important disabilities [4, 5, 16, 119]. It must be remembered, however, that the statistical approach used in building the PR does not allow for the interpretation of relationships as causal [117].

When building the PR, we intended to develop an applicable tool. Orthopedic surgeons and staff who assess the patients' status upon placing them on TKA wait lists could be the main users of the PR. However, the PR could eventually also be used by other healthcare professionals, namely physician assistants, physiotherapists, occupational therapists, and nurses. In this way, any healthcare professional who takes charge of the case can take into account the results of the PR when determining the most appropriate course of action for the patient's care. The PR shows promising practical implications, as it is relatively simple and easy to use in a clinical setting.

At times, the interpretation process may seem counterintuitive. Question 2 of the rule indicates that a patient may be categorized at risk if they experience mild difficulty when getting on or off the toilet, yet, depending on their answer to the Question 4, they may be classified as not at risk if they encounter severe difficulty performing the same task. Since the PR includes four items from the function subscale of the WOMAC, it would be expected that it classifies at risk those patients whose preoperative function is severely affected, considering that preoperative function has been consistently

identified as a determinant of functional outcome [52]. Nevertheless, a predictor is not necessarily a determinant; its purpose is to predict the desired outcome and the development with recursive partitioning determined that it is this answer pattern that shows the best predictive value.

6.2. Strengths of the study

This study followed a prospective longitudinal cohort design. It had high participation and follow-up proportions. There does not seem to be a selection bias, as there were no significant differences between participants and eligible non participants on age and gender as well as no significant differences in terms of pain, functional limitations at enrolment on the pre-surgery wait lists between subjects lost to follow-up, subjects who did not undergo surgery and participants who completed the interview six months after surgery (data not shown). Baseline measurements of the dependent variables were made as soon as the patients were enrolled on the pre-surgery wait list (mean ± SD: 12.6 ± 4.7 days).

6.3. Limitations of the study

The sample of 141 patients that was used to derive the PR was small. This may diminish the applicability of the PR to the general population. Moreover, the population under study was patients undergoing primary TKA. This effectively may disregard patients with revision or bilateral TKA. Furthermore, the clinical outcomes of TKA were assessed using the WOMAC questionnaire, a self-reported measure. Performance-based measures such as the Timed Up and Go Test or the Six Minute Walking Test could have been used in order to complement the information recorded by the self-reported measure [156]. An assessment of the patients' status in a more comprehensive manner could have thusly been achieved. In addition, we considered the patients having postoperative WOMAC

scores in the first quintile to have a poor outcome. Since there is no consensus on what constitutes an appropriate measure of poor outcome following TKA, we decided that this method could be applied. It is important to point out that this PR allows for the identification of patients at risk of poor outcomes in the short term following TKA and was not tested to predict long-term outcomes. The time point of six months after surgery was chosen to identify patients with poor outcomes, as it is a critical time in the patients' rehabilitation period when they are often seen by surgeons to monitor progress and where the rehabilitation protocol and conservative treatment options may be easily modified if recovery is not optimal. Finally, the final PR has not yet been validated with a different sample of patients, its predictive validity has not been compared to clinical judgment alone, and the clinical and financial impacts of its use have not been yet assessed. Until these further research steps are completed, the PR should be used with caution.

7. Conclusion

The PR developed in the current study has the potential to identify patients at risk of poor surgical outcomes following TKA. Such patients could then be assigned to an appropriate course of action, such as prehabilitation, conservative management, wait list priority or intensive postoperative rehabilitation. These conducts may diminish the extent of deterioration of patients waiting for TKA and could decrease the socioeconomic burden of TKA. A further validation in an external cohort is needed. Impact analysis determining the usefulness of the rule in the clinical setting regarding cost-benefit, time and resource allocation as well as patient satisfaction is equally required.

8. Competing interests

The authors declare the following competing interests:

EL declares that he has no competing interest.

FD declares that he has no competing interest.

CED declares that he has no competing interest.

ELB performs paid presentations for Ergoresearch Inc., Pfizer, Bayer Inc. and Sanofi Aventis. He is an educational consultant for Zimmer Inc. As a principal investigator, he receives research support from CIHR, MSSS-QC, Stryker and Depuy-Synthes.

PAV performs paid presentations for Stryker and Wright Medical. He is also a consultant for Wright Medical, Emovi and Stryker. As a principal investigator, he receives research support from Stryker, Biomet, Smith and Nephew and Zimmer.

9. Author contributions

EL participated in the interpretation of results and statistical analysis. He led the writing of the manuscript.

FD participated in the design, coordination, and collection of data as well as the statistical analysis and the writing of the manuscript. He led the interpretation of results.

CED participated in the design, coordination, and interpretation of results and writing of the manuscript.

EB participated in the design, coordination and writing of the manuscript.

PAV participated in the design, coordination and writing of the manuscript.

10. Acknowledgements

EL is supported by a training award from the FRQS.

Figure 1 Flowchart of patients' recruitment. *Eligibility status unknown (considered in calculation of participation proportion). TKA, total knee arthroplasty.

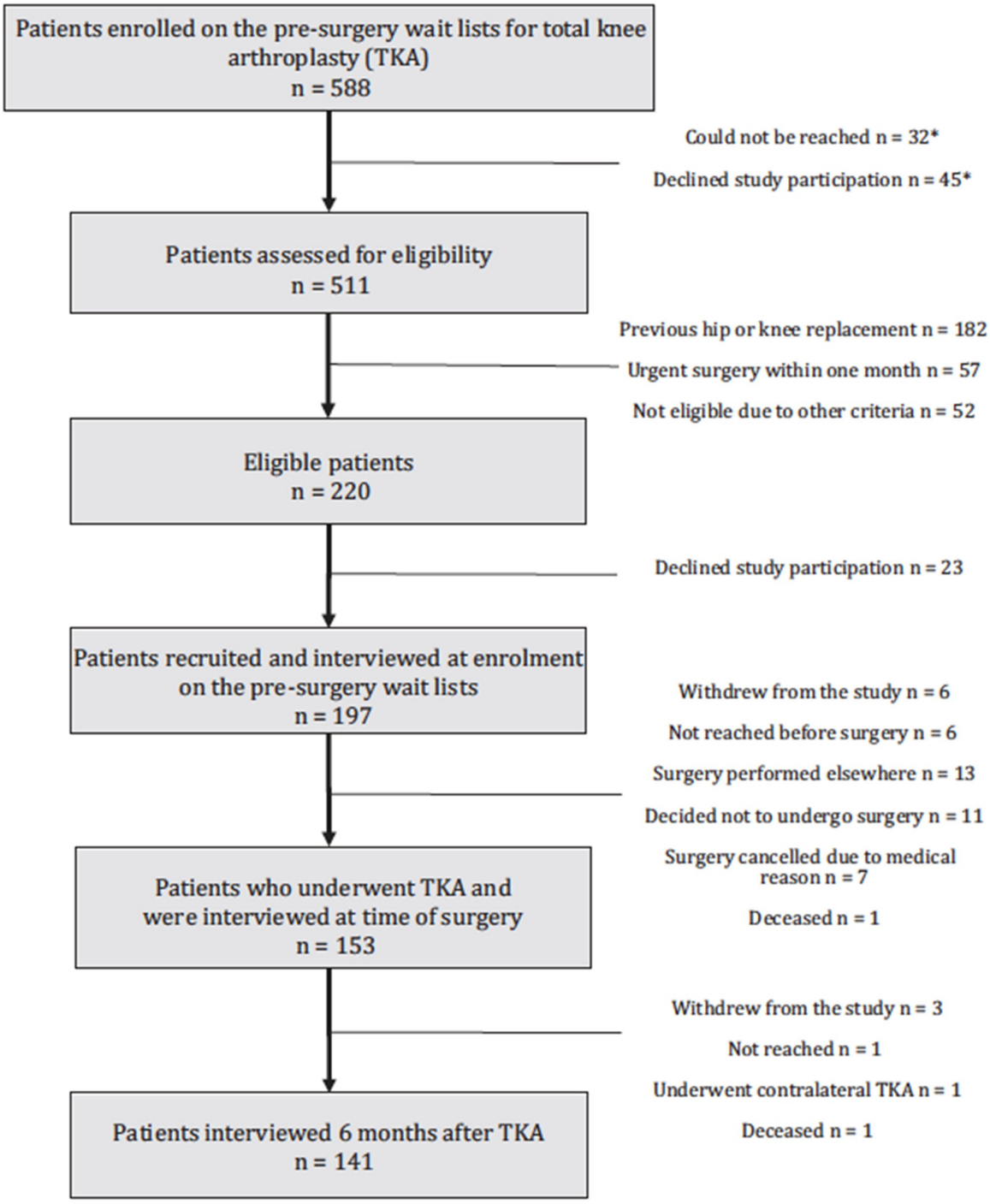
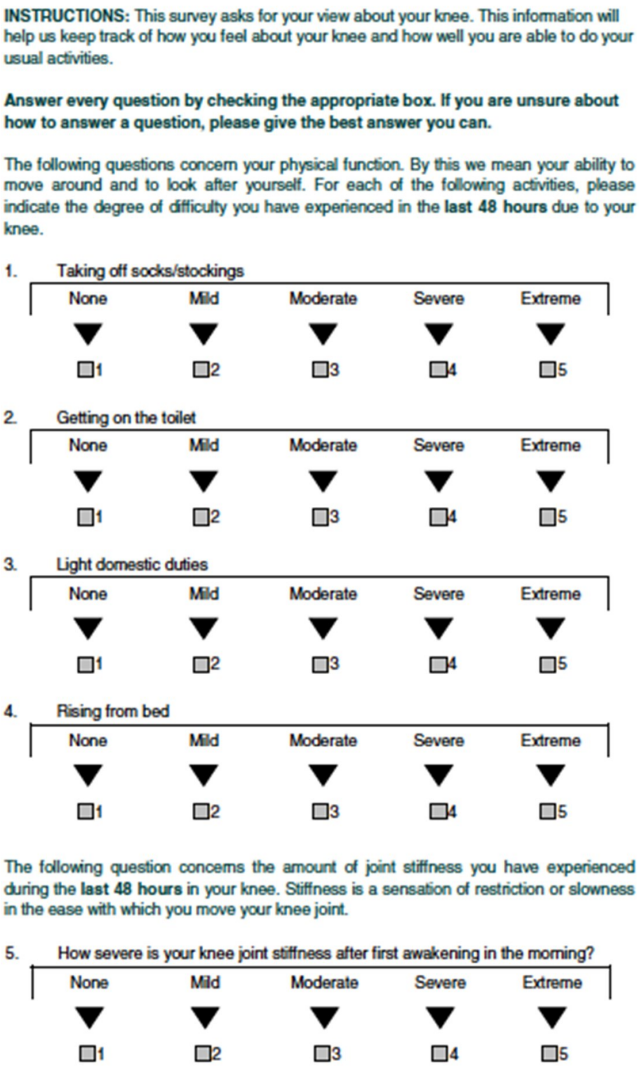


Figure 2 Prediction algorithm to identify patients at risk of poor outcome following TKA



INSTRUCTIONS: This survey asks for your view about your knee. This information will help us keep track of how you feel about your knee and how well you are able to do your usual activities.

Answer every question by checking the appropriate box. If you are unsure about how to answer a question, please give the best answer you can.

The following questions concern your physical function. By this we mean your ability to move around and to look after yourself. For each of the following activities, please indicate the degree of difficulty you have experienced in the **last 48 hours** due to your knee.

1. Taking off socks/stockings

None	Mild	Moderate	Severe	Extreme
☐1	☐2	☐3	☐4	☐5

2. Getting on the toilet

None	Mild	Moderate	Severe	Extreme
☐1	☐2	☐3	☐4	☐5

3. Light domestic duties

None	Mild	Moderate	Severe	Extreme
☐1	☐2	☐3	☐4	☐5

4. Rising from bed

None	Mild	Moderate	Severe	Extreme
☐1	☐2	☐3	☐4	☐5

The following question concerns the amount of joint stiffness you have experienced during the **last 48 hours** in your knee. Stiffness is a sensation of restriction or slowness in the ease with which you move your knee joint.

5. How severe is your knee joint stiffness after first awakening in the morning?

None	Mild	Moderate	Severe	Extreme
☐1	☐2	☐3	☐4	☐5

INTERPRETATION:

The patient is **AT RISK** in case of any of the following combination of answers:

Q1: Moderate (3) or Severe (4) or Extreme (5)
Q2: Moderate (3) or Severe (4) or Extreme (5)
Q4: Severe (4) Extreme (5)

OR

Q1: Moderate (3) or Severe (4) or Extreme (5)
Q2: None (1) or Mild (2)

OR

Q1: None (1) or Mild (2)
Q3: Moderate (3) or Severe (4) or Extreme (5)
Q5: Severe (4) or Extreme (5)

Figure 3 Graphical representation of the PR and its interpretation

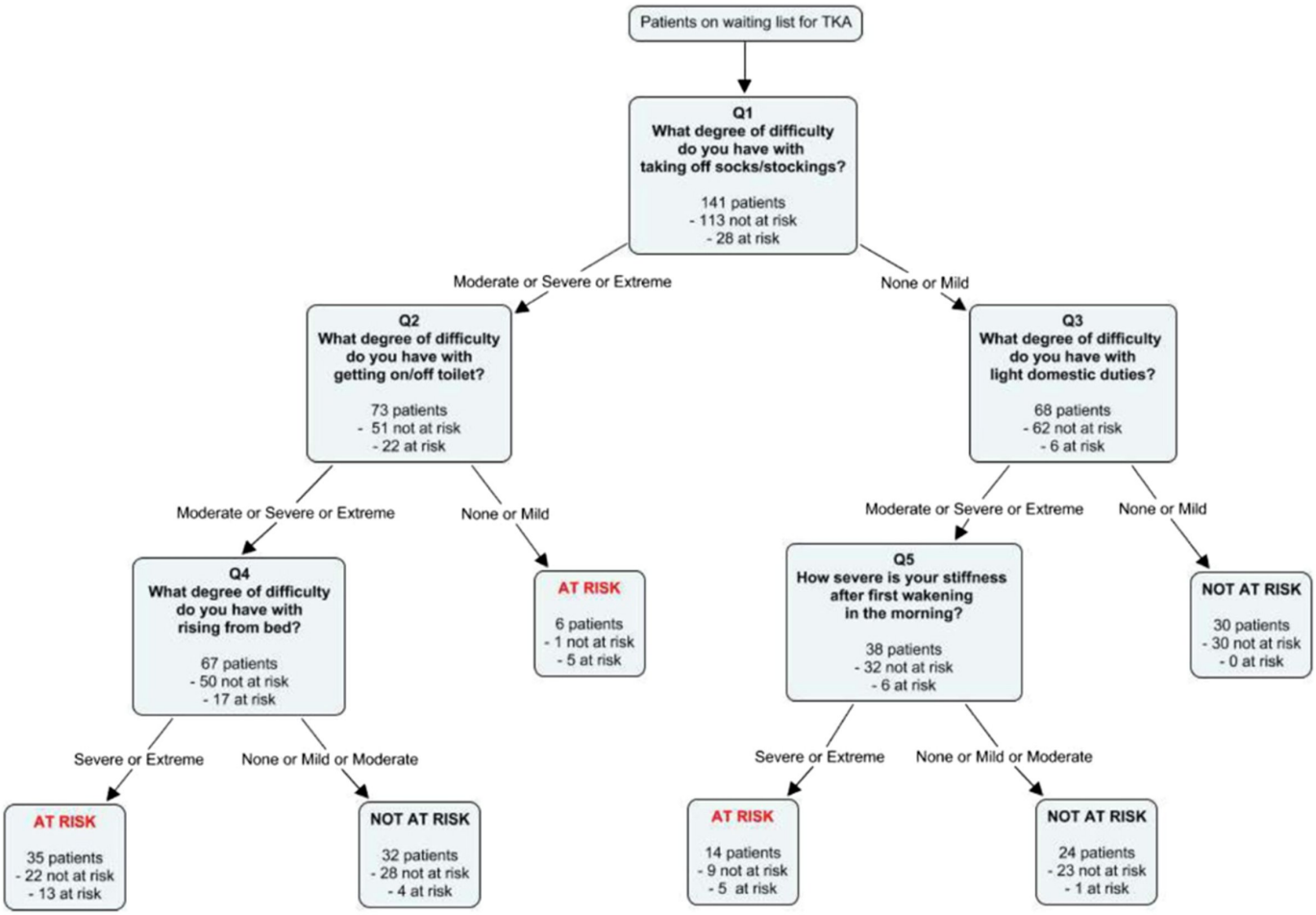


Table 1 Selected characteristics of the participants who underwent primary unilateral total knee arthroplasty surgery (n = 141)

Variables considered for PR development	n (%)	Mean (SD)	Other collected variables	n (%)	Mean (SD)
Demographics			**Pre-surgery wait**		
Age (years)		66 (9.5)	**Time between enrolment on wait list and surgery** (days)[*]		184 (120.8)
Female	93 (66)		**Categories of wait time**		
Marital status			≤3 months	30 (21)	62.5 (16.9)
Single, separated, divorced or widowed	51 (36)		>3-6 months	53 (38)	130.2 (28.7)
Married or common law	90 (64)		>6-9 months	31 (22)	216.8 (25.1)
Living alone	34 (24)		>9 months	27 (19)	386.3 (56.7)
Socioeconomic characteristics			**Surgery postponed for personal reasons**	9 (6)	
Educational level (part or complete)			**Surgical characteristics**		
High school or less	79 (56)		**Implant type**		
College or University	62 (44)		Postero-stabilized	115 (82)	
Employment status			Cruciate retaining	26 (18)	
Unemployed or retired	108 (77)		**Implant fixation**		
Employed	33 (23)		Cementless	4 (3)	
Household income [**]			Hybrid	2 (1)	
< $30 000 / year	48 (34)		Cemented	135 (96)	
$30 000 - $59 999/ year	43 (31)		**Implant bearing type**		
≥ $60 000/ year	34 (24)		Mobile	4 (3)	
Missing data	16 (11)		Fixed	137 (97)	
Psychosocial characteristics			**Patella resurfacing**	132 (93)	
Psychological distress (/42)		7.2 (7.0)	**In-hospital complications**[‡]		
Social support[°]			0	109 (77)	
Low	67 (48)		1	23 (16)	
High	74 (52)		≥2	10 (7)	
Clinical characteristics			**Health services utilization**		
Diagnosis			**Hospital length of stay** (days)		7.5 (3.0)
Osteoarthritis	136 (96)		**Discharged directly home**	123 (87)	
Rheumatoid arthritis	5 (4)		**Post-surgery community physiotherapy** (hours)		14.7 (18.7)
BMI [¯] (kg/m^2)		31.2 (6.2)			
Comorbidities (/56)		6.5 (2.2)			
Duration of knee symptoms before		7.9 (8.1)			
Contralateral knee pain [§]	101 (72)				
Use of a walking aid					
At enrolment on wait list	55 (39)				

SD: standard deviation

[*] Median (range): 148 days (32-692)

[**] n=125 – CND $

[°] Social support was dichotomized around the median score: Low (≤ 80) and High (>80)

[¯] Body mass index

[†] n=138

[§] WOMAC pain score at enrolment on pre-surgery wait list dichotomized into presence or absence of contralateral knee pain

[‡] In-hospital complications including: wound infection, dislocation, knee ankylosis and manipulation, cardiovascular/pulmonary/circulatory complications, peripheral/central nervous system involvement, urinary infection, acute confusion, tendon and ligament rupture or blood transfusion

Table 2 Overall changes in WOMAC scores of the participants between enrolment on the pre-surgery waiting lists and 6 months after TKA (n = 141)

	Mean score at enrolment † (SD)	Mean score 6 months after TKA † (SD)	Change in score ‡ (SD)	95% CI	Comparison between time points (*p* value)
WOMAC					
Pain	53.1 (17.9)	22.5 (17.1)	- 30.6 (21.8)	- 26.9 to - 34.2	<0.001*
Stiffness	59.3 (19.7)	33.3 (21.1)	- 26.0 (20.4)	- 21.2 to - 30.8	<0.001*
Function	53.5 (14.3)	28.1 (17.9)	- 25.4 (20.5)	- 22.0 to - 28.8	<0.001*
Total score	55.3 (15.2)	28.0 (16.3)	- 27.3 (15.8)	- 23.6 to - 31.0	<0.001*

SD: standard deviation

CI: confidence interval

† Scores presented as standardised scores. Lower scores sign a better condition

‡ Negative changes in score sign an improvement of the condition

* $p < 0.05$

Table 3 Two by Two table of predicted versus actual outcomes of the final PR

	Actual Outcome	
Predicted outcome	**AT RISK (Postoperative WOMAC > 40.4)**	**NOT AT RISK (Postoperative WOMAC ≤ 40.4)**
AT RISK	23	32
NOT AT RISK	5	81
TOTAL	*28*	*113*

Table 4 Validity measures of the predictive rule

Measure	Estimates in training sample	Estimates with 1,000 bootstrap resamples
Sensitivity % (95% CI)	82.1 (64.4-92.1)	82.1 (66.7-95.8*)
Specificity % (95% CI)	71.7 (62.8-79.2)	71.7 (62.8-79.8*)
Positive predictive value % (95% CI)	41.8 (29.7-55.0)	41.8 (29.1-55.8*)
Negative predictive value % (95% CI)	94.2 (87.1-97.5)	94.2 (88.8-98.8*)
Positive likelihood ratio (95% CI)	2.90 (2.06-4.08)	2.90 (1.81-4.74*)
Negative likelihood ratio (95% CI)	0.25 (0.11-0.57)	0.25 (0.11-0.58*)
Area under ROC curve (95% CI)	0.77 (0.69-0.85)	0.77 (0.69-0.85*)

- * 95% asymptotic confidence intervals
- Sensitivity: number of participants classified at risk both by the PR and the postoperative WOMAC score divided by all participants classified at risk by the postoperative WOMAC score (actual outcome).
- Specificity: number of participants classified not at risk by the PR and the postoperative WOMAC score divided by all participants classified not at risk by the postoperative WOMAC score (actual outcome).
- Positive predictive value: number of participants classified at risk by the PR and the postoperative WOMAC score divided by all participants classified at risk by the PR (predicted outcome).
- Negative predictive value: number of participants classified not at risk by the PR and the postoperative WOMAC score divided by all participants classified not at risk by the PR (predicted outcome).
- Positive likelihood ratio: sensitivity/ (1-specificity)
- Negative likelihood ratio: (1-sensitivity)/specificity.
- Area under the ROC curve is defined as the area under the sensitivity vs. 1-specificity curve

CHAPTER 7

DISCUSSION

The objective of this memoir was twofold: firstly, we aimed to identify the determinants of pain and disability following TJA with the greatest level of evidence by systematically reviewing the literature. Secondly, we aimed to develop prediction models allowing the preoperative identification of patients at risk of poor outcomes after TJA. In the following sections, we discuss the main results of the systematic reviews as well as of the prediction models. Subsequently, we address the main strengths and limitations of this memoir, and conclude by discussing possible practical applications and future avenues.

1. Determinants of high levels of pain and functional disabilities following THA

In this study, we identified 22 articles with a mean methodological score of 81.0%, SD 10.3% that describe with a moderate-to-high level of evidence the determinants of short- and medium-term pain and functional levels following THA. The results of this study allowed to conclude that significant preoperative determinants of poor pain and function following THA with the highest level and amount of evidence are a lower educational level, the preoperative level of pain and function, higher BMI, presence or a greater level of comorbidities, worse general health and lower radiographic OA severity. Considering that a significant proportion of patients undergoing THA experience unfavourable outcomes in terms of pain and disability, the identification of these determinants may help target patients at risk in order to ultimately ameliorate their outcomes.

The main finding of this study is the significant relationship between the preoperative and the postoperative levels of pain and function. As a matter of fact, out of 12 studies that investigated this relationship, a significant association was found in nine studies. Although the literature consistently

refers to this relationship as a well-known fact, to our knowledge, our systematic review is the first underlining it with an appropriate extent of precision and validity. The direction of such an association is a function of the method of measurement of the outcome. Indeed, when the outcome is measured as postoperative state, i.e. the level of pain and/or function after the surgery regardless of the preoperative levels of the respective domains, a worse preoperative status is significantly associated with a worse postoperative status. When the studies investigate outcome as a measure of domain change, i.e. the difference between preoperative and postoperative levels, a lower preoperative status is associated with a higher chance of achieving an improvement on the respective domain. The clinical implication of these findings is challenging, as two opposite situations may emerge. THA may be performed early in the course of the disease when the patients' pain and function are still relatively preserved, such as to lead to appropriate levels postoperatively. The patients' pain and function will then achieve a certain plateau, and their ultimate outcome will be limited by the longevity of the prosthesis. If THA is performed later in the disease process when the patients are very limited by OA, although their postoperative status will be worse compared to if they had THA earlier in the course of their disease, the amount of improvement they would experience would be greater. In this situation, patients would achieve a notable relief, and will likely be very satisfied with their intervention. The caveat is their long-term levels of pain and function, which would be worse compared to their peers undergoing surgery earlier. Although this logic seems to favour performing THA for patients early in their disease process, it is realistically difficult to do so for all surgical candidates. It is possible that patients' symptoms can be managed with less invasive procedures incurring fewer risks and smaller costs, such as analgesic medication, intra-articular corticosteroid injections and rehabilitation, among others, that can ultimately halt disease progression [6]. Unfortunately, there does not seem to be a consensus regarding this predicament, and our findings only complement the uncertainty that these concepts

instigate.

An interesting finding of this review was the amount of conflicting evidence regarding the association of demographic variables with THA outcomes. For example, only four out of the 12 included studies that investigated this relationship showed that older age was associated with worse pain and disability following THA. Similarly, two out of ten included studies proposed gender as a significant determinant of THA outcomes. We can therefore conclude that these demographic variables are most likely not associated to THA surgical results, which is in stark contrast with the previously published systematic reviews [10, 18]. Although our study has some advantages over the aforementioned systematic reviews, and may conversely be affected by the limitations of the included prognostic studies that will be discussed shortly, this conclusion has potential important clinical implications. For example, a prehabilitation program may equally target THA candidates of all ages as well as men and women. Moreover, surgery should not be denied based on these variables, at least with respect to pain and functional outcomes.

Another original finding of our systematic review is the association between a higher BMI and poor THA outcomes (six out of ten studies). A longstanding paradigm, the effect of body mass on THA results in terms of early and long-term complications, operative times, length of hospital stay and clinical results has been vastly debated, yet no consensus has been reached [157]. Although the design of the prognostic studies included in the review cannot suggest a causative link between BMI and THA results, this result may find clinical utility. Considering that obesity is a well-known risk factor for development of hip OA and that the average BMI of THA candidates has been increasing over time [158], patients with higher BMI may require particular attention, such as closer follow-up and possibly weight loss interventions. Nevertheless, some of the included studies suggest that the magnitude of

association between higher BMI and poor surgical outcomes in terms of pain and function is smaller compared to other determinants, and may not be clinically important. However, due to the current review's inability to conclude on the strength of the association between determinants and THA outcomes, this issue may require further clarification.

The consistent association between THA outcomes and the degree of general health either measured as presence or level of comorbidities, back pain, physical and mental health, or involvement of the contralateral hip, is another important finding of this review. The nature of the relationship between these variables and THA outcomes is intricate, and further research is indicated to elucidate this. Nevertheless, these variables could represent components that should be targeted in a prehabilitation program aimed at improving pain and function following THA. Moreover, similar to the practice of assessing fitness for surgery using the American Society of Anaesthesiologists Scale because of its association with procedural complications, employing measures of health could potentially assess the merit of undergoing the procedure.

Lastly, a lower radiographic OA severity was unanimously associated with smaller changes in pain after THA in three studies. Radiographic OA severity is traditionally assessed using the Kellgren-Lawrence scale, which takes into account radiographic aspects of OA such as joint space narrowing, osteophytes, subchondral bone sclerosis, cysts and deformity of the femoral head and acetabulum [159]. It has been previously suggested that less severe disease is appreciably associated to clinical manifestations such as pain and function levels [111]. The current finding likely parallels the significant association between preoperative and postoperative health state, with the implications previously described. It is however possible that in patients with low radiographic OA severity, the hip

pain actually represents referred pain from other regions, such as the lumbar spine. In this situation, the surgical indication is questionable, thus possibly explaining poor outcomes in terms of pain and function. Moreover, patients with low radiographic OA severity have low pain threshold and possible inadequate expectations regarding the results of the intervention, also eventually contributing to suboptimal outcomes.

2. Preliminary prediction algorithm identifying patients at risk of suboptimal outcomes following hip arthroplasty

This study entailed the development of a preliminary PA identifying patients at risk of poor outcomes following hip arthroplasty in a retrospective cohort of 265 patients collected from the HMR's arthroplasty database. Forty patients (15.1% were in the worst quartile of the postoperative WOMAC score and perceived their operated hip as artificial with minimal or major limitations, and were therefore considered to have had a poor outcome. The final PA included age, gender, BMI as well as three items of the preoperative WOMAC. This algorithm shows favourable predictive capabilities, with a sensitivity of 75.0% (95% CI 59.8.4-85.8), a specificity of 77.8% (95% CI 71.9-82.7) and a positive likelihood ratio of 3.38 (95% CI 2.49-4.57). The internal validity of the PA was successfully determined using 1,000 bootstrap resamples.

In order to develop a clinical prediction algorithm, candidate variables can be selected based on clinical and theoretical understanding, as well as on previous similar attempts reported in the literature [160]. The literature of predictors of hip arthroplasty outcome is however very limited, with only one study reporting an analogous prediction algorithm to our knowledge [69]. Consequently, in order to

select potential predictor variables, the pool of significant determinants of hip arthroplasty outcomes can be surveyed. Out of the six predictors included in this PA, BMI and the preoperative WOMAC levels are also significant determinants of THA outcomes as reported by the systematic review in Chapter 3. Additionally, age, gender and BMI can also be found in the prediction model by Slaven et al. These encouraging findings are however limited by the fact that several significant determinants as identified by the systematic review were not considered as potential predictor variables when developing the PA due to the limitations associated with using a retrospective database. For example, education levels, general health and radiographic OA severity were either incompletely or not available in the database. This represents a limitation of the current study, and a future prospective study should collect data on all important determinants. Interestingly, age and gender were not identified as significant determinants in the systematic reviews, but were included in the final predictive model. This situation underlines the difference between a predictor and a determinant; the former is able to predict an outcome, whereas the latter merely indicates that there is a significant association with an outcome.

Two constructs (postoperative total WOMAC score and hip perception) were used to define patients as being at risk in order to maximize the likelihood of them being truly subject to poor surgical outcomes. Conceptually, this classification is different from the MCID and other measures of change in status described in Appendices A and B. Our purpose was to develop a prediction model that would help identify the patients with poor postoperative functional status, severe pain and poor hip perception regardless of the effectiveness of their intervention. We did not intend to target the prediction of a treatment effect, such as would have been the case if we used the MCID or other similar measure as an outcome measure. The latter equally present numerous potential limitations, including multiple

possibilities of calculation methods, different definitions of "change" and derivation based on non-representative populations, ultimately yielding values of doubtful validity [110]. Some authors do not recommend their use in a clinical setting [110].

The proposed PA was developed on a sample of patients undergoing THA or HR, although their indications can be different. We decided to include both types of procedures in order to develop a tool capable of successfully performing in a heterogeneous population. In fact, in a clinical setting, an orthopaedist may consider both types of interventions for a patient, and this tool may be employed regardless of the final choice of the type of joint replacement. Moreover, in the case of our sample, there were no statistically significant differences between the two procedures in terms of age, gender and number of comorbidities ($p > 0.05$); BMI was higher for the THA group (mean ± SD: 28.8±5.3 kg/m^2 vs. 27.0±4.5 kg/m^2, $p = 0.01$). In terms of the patients that were identified at risk ($n = 40$), 8 underwent HR and 32 underwent THA; the observed and the expected counts were not statistically different ($p = 0.171$). Furthermore, the variable "Type of surgery" was not identified as a significant predictor. While these interventions may have different indications and outcomes such as failure rates [40], based on the similarities between the two groups, it is acceptable to evaluate them conjointly, albeit uniquely for the purpose of assessing pain, disability and joint perception.

We intended to develop a clinically pertinent tool, whereby the prediction process is logical and applicable. Ideally, the direction of the link between a variable with the outcome is similar, whether it is considered a predictor or a determinant; for example, a higher BMI as a determinant is associated with a worse outcome, and as a predictor, a higher BMI predicts a worse outcome. Nevertheless, at one instance, the interpretation process may be counterintuitive. For example, it is possible, in an extreme

scenario, for a 49 year-old male patient with a BMI of 22 kg/m^2 and with no pain when walking on a flat surface as well as with no difficulty with putting on socks or stockings to be classified as at risk of suboptimal surgical outcome. Development with recursive partitioning indicated that this pattern of answers has the best predictive capabilities, again emphasizing the difference between a predictor variable and a determinant.

3. Determinants of poor pain and function following total knee arthroplasty

In this study, we systematically reviewed the literature of determinants of poor short and medium-term outcomes in terms of pain and function following TKA. We identified 34 eligible studies, which allowed us to pinpoint demographic, socioeconomic, clinical, surgical and healthcare-related determinants with an excellent level of evidence (mean methodological score 80.0%, SD 12.7%). This in turn allowed us to select a list of determinants associated with poor pain and function with the greatest amount of high-quality evidence.

The most evidence was found for the significant relationship between the preoperative and the postoperative levels of pain and function (21 out of 23 studies). The logic and implications of such an association is identical to the similar finding for THA, as discussed in section 1 of this chapter. A greater level of preoperative health, whether measured as the total number of comorbidities, presence of specific comorbidities, the effect exerted by the comorbidities on one's life or level of general health has been significantly associated to poor outcome in terms of pain and function. The effect of comorbidities on mortality and operative complications associated to TKA is well studied and understood [161, 162]. The association between comorbidities and outcomes in terms of pain and function is less known, and our review strongly emphasizes it. However, the rationale behind such an

association is poorly understood, and because it is presumed that greater levels of comorbidity is detrimental to rehabilitation, appropriate preoperative counselling may be recommended [92, 128, 130]. In line with the previously described importance of comorbidities, greater amount of evidence regarding psychosocial variables associated with poor outcomes after TKA has started to emerge. Among these, pain catastrophizing likely alters the neurophysiological processes related to pain modulation, leading to a greater sensitization to pain [133]. Combined with the potential impact of preoperative depression and anxiety, this finding underlies the importance of preoperative mental health in determining outcome of TKA surgery. Therefore, interventions targeted at ameliorating surgical outcomes should also consider tackling psychological health. Finally, a lower socioeconomic status was associated with worse pain and function after TKA in two large cohorts. This may suggest that socially deprived patients are more in need of interventions such as prehabilitation, and would be more likely to benefit from them. Finally, regarding surgical determinants, very few studies identified variables such as cementless and cruciate-retaining implants as associated with pain and function after TKA, and none of these are consistent between the studies. This may suggest that the effect of this type of factors is generally investigated using a clinical trial methodology [163, 164], whereas our review focused solely on prognostic studies.

This study underlines that several variables that are traditionally thought to be associated with worse TKA outcomes in terms of pain and function are not significant determinants when studies with high methodological quality are overviewed. For example, neither age, gender, nor BMI were found to be related to pain or function after TKA in a consistently significant manner. Several studies identified older patients, females and patients with higher BMI as at risk of poor outcomes in terms of pain and function, yet when viewing the general picture, an overwhelming amount of evidence points to an

opposite direction of association or to no association at all. Therefore, according to these findings, we suggest that TKA should not be systematically refused to patients based on their age, gender or BMI with respect to pain and functional outcomes.

When comparing the findings of this review to the similar study on determinants of THA outcomes, some interesting observations can be made. For example, higher BMI was found to be a significant determinant of pain and function after THA but not TKA. Although obesity is a significant risk factor for developing OA for both joints [158, 165], the disparity between its association with TJA outcomes is difficult to explain.

Similarly to THA, TKA outcomes are not associated with preoperative age and gender, further suggesting the relevance of targeting patients for outcome improvement irrespective of these variables. The important significant association between preoperative and postoperative levels of pain and function seems to be equally valid for TKA and THA, and advocates for the amelioration of pain and function prior to surgery by means of targeted interventions such as prehabilitation. Finally, although psychosocial variables were not found to be associated to THA outcomes like in the case of TKA, it is possible that a smaller attention to this variable in the THA literature contributes to this finding.

4. Clinical prediction tool identifying patients awaiting surgery who are at risk of greater pain and disability following TKA

In this study, 141 patients were followed from their inclusion on the waiting lists for a primary unilateral TKA until six months following surgery in order to develop a CPR allowing the

identification of patients at the greatest risk of poor pain and function postoperatively. The developed CPR encompasses five items of the preoperative WOMAC scale measured at the time of the inclusion on the waiting list, namely difficulty with taking off socks, getting on/off toilet, performing light domestic duties and rising from bed as well as degree of morning stiffness after the first wakening. The CPR has good predictive capabilities, allowing it to potentially serve as an appropriate screening tool.

This CPR was developed by considering an extensive set of variables acquired at the time of enrolment on the waitlist that were carefully selected based on the literature of potential determinants of TKA results at the moment of the inception of the prospective cohort (2006). When comparing to the significant determinants of TKA outcomes as identified by the systematic review, only back pain and extent of preoperative pain catastrophizing were not collected in this study. Moreover, the choice of the final tool was based on selecting a simple model with appropriate sensitivity and specificity. In this way, we ensured the development of a prediction rule that is both evidence-based as well as clinically applicable. The end-result is an applicable screening clinical tool with that is also user-friendly.

To our knowledge, this CPR is the first predictive tool developed to identify patients at risk of poor outcomes after TKA or any other type of joint replacement specifically in patients awaiting surgery. The main implication of this approach is the ability of the users of this tool to identify the patients at risk early in their management process. In this way, targeted prehabilitation during the waiting time can potentially lead to diminishing their risk. Moreover, postoperative rehabilitation of the identified patients could additionally improve their outcomes but has yet to be formally proven.

The fact that the CPR is formed by items of the preoperative WOMAC scale is not surprising, given

that the preoperative level of pain and function is the most important determinant of postoperative pain and function, as evidenced by the findings of our systematic review. The strength of this association is indirectly emphasized by the fact that such an extensive set of possible TKA determinants were considered for the development of the CPR.

5. Strengths of the current memoir

One of the strengths of the systematic reviews resides in the application of rigorous eligibility criteria, ensuring the inclusion of studies with the greatest level of evidence for determinants of TJA outcomes from four prominent databases. This bears advantage over the previously published reviews, whose conclusions were based mainly on poor-quality studies. Moreover, the focus on all types of determinants of surgical outcomes ensures a thorough survey of all variables that were identified significantly related to joint arthroplasty outcomes. This process allowed the elaboration of a list of determinants with the highest level of evidence.

The PA identifying the patients at risk of suboptimal results after hip arthroplasty is, to our knowledge, the first model that has the ability to discriminate hip arthroplasty outcomes based on more than one parameter, namely patients' functional, stiffness and pain levels as well as their perception of the replaced hip joint. This approach increases the likelihood of correct classification, especially in a context where an accepted definition of poor outcome is still controversial. Moreover, this classification identified 15% of the patients as having unsuccessful outcomes, well in line with the published proportions of what can be considered a suboptimal outcome [3]. Finally, the rigorous statistical analysis employed in the development of the PA underlines the stringency of our approach.

The development of the CPR identifying the patients at risk of poor outcomes following TKA was based on a prospective cohort with high participation and follow-up rates (77.8% and 92.2% respectively). The patients' baseline measures were collected very soon after they were included on the waitlist (mean ± SD: 12.6 ± 4.7 days). The risk of selection bias is very low, as there was no statistically significant difference between participants and eligible non participants on age and gender as well as no significant differences in terms of pain, functional limitations at enrolment on the pre-surgery wait lists between subjects lost to follow-up, subjects who did not undergo surgery and participants who completed the interview six months after surgery.

6. Limitations of the current memoir

As mentioned previously, although we aimed to consider manuscripts with the highest methodological qualities, the prognostic studies that were included in the systematic reviews have several important limitations with respect to our aim of summarizing the determinants of TJA outcomes and especially evaluating the strength of association. For example, employing validated PROMs ensures a higher quality of evidence, yet when different tools are used, pooling of results becomes impossible. Moreover, the prognostic studies employ different definitions of surgical outcome (either change in status or postoperative status). Although we conclude that determinants are similar between the two definitions, pooling of regression coefficients or odds ratios is challenging. Also, while multivariate analyses are more robust than their univariate counterparts, linear and logistic regressions are distinct methods, a fact that limits the potential for pooling the strength of associations between determinants and TJA outcomes. Finally, the method of measurement of independent variables is inconsistent among the included studies. Ideally, future prognostic studies should focus on standardizing reporting of associations between determinants and TJA outcomes. However, considering the lack of consensus

regarding the different aspects, this may prove extremely challenging.

Because of similar methodologies, the PA and the CPR have comparable limitations. The populations under study were patients undergoing primary unilateral hip arthroplasty or TKA, which limits the generalizability of these findings to patients with revision and bilateral interventions. Furthermore, the surgical outcomes were based on self-reported measures (WOMAC and joint perception). Use of performance-based measures such as the Timed Up and Go Test and the Six Minute Walking test could have complemented these assessments, and potentially provided a more holistic overview of the patients' faring [156]. In terms of the CPR, the prospective cohort entailed a small sample of 141 patients. The CART algorithm that was employed to develop the prediction models has several disadvantages as discussed in Chapter 2, and these may limit the current findings. The retrospective nature of the study that led to the development of the PA holds an increased risk of selection bias, considering that when comparing to subjects excluded due to missing data, the included participants were on average younger, had a greater number of comorbidities and a greater proportion were male. Because of their preliminary status, these prediction tools should not be implemented clinically as of yet, as external validation with different samples should be first undertaken. Finally, the approach described in this memoir, which involves implementing interventions for patients at risk of poor outcomes as defined, by pain, function and satisfaction, is limited by the fact that surgical results encompass multiple aspects. For example, other outcome variables that may or may not be associated with pain and disability include surgery-associated complications, implant revision or activity levels. These may need to be taken into account when assessing the requirements for prehabilitation, rehabilitation, and ultimately the decision to undergo TJA. Developing clinical tools and guidelines assessing all of these aspects would prove intricate, and likely of limited clinical applicability.

Ultimately, as described in the introduction, patient-reported outcomes are an important aspect of surgical results, and we believe that focusing our research on them may prove the most advantageous clinically, especially considering the higher occurrence of poor results in terms of pain, function and satisfaction compared to the other types of outcomes.

7. Impacts of current findings and future directions

In this memoir, we provided a summary of preoperative determinants of pain and physical function levels up to 2 years following TKA and THA. Moreover, we developed two predictive models allowing the identification of patients at risk of poor surgical outcomes following TJA. The main implication of these findings is the assistance they could bring in developing and implementing clinical and public health measures in order to ameliorate the outcomes of patients undergoing TJA. A possible intervention is the prioritization of patients or groups of patients who are suspected of being at risk of less successful surgical outcomes. It has been shown that a longer waiting time is associated with deterioration in pain and function preoperatively, and could even impact the postoperative outcome [148, 166]. With this in mind, earlier intervention can have the potential of improving outcomes, and this hypothesis should be evaluated in future studies. From a clinical point of view, knowledge of determinants of unsuccessful outcomes as well as the ability to predict which patients will fare worse can lead to the implementation of intensive prehabilitation and postoperative rehabilitation. Targeting of modifiable risk factors, such as weight loss, amelioration of back pain or management of comorbidities for example, may possibly diminish the risk profile. Ultimately, if the developed prediction models prove valid in subsequent external validation studies, they could be employed as tools evaluating the necessity or the merit of undergoing the procedure. Finally, our findings may aid stakeholders in assessing which patients or groups would benefit the most from these prevention

strategies.

A limitation of the literature that the systematic reviews identified resides in the conflicting evidence regarding the significance of the association of certain variables with the outcome of joint arthroplasties. This limitation underlines the importance of conducting higher quality studies in order to elucidate this association. A structured, consistent approach of future cohort studies may also significantly facilitate the pooling of results in ensuing systematic reviews and meta-analyses.

Subsequent steps in the development of the prediction models consist in their external validation as well as impact assessment, as discussed in Chapter 2. Further studies should employ a prospective design, and maximize the number of collected determinants of TJA outcomes in order to optimize the selection of predictors. If successful, these tools could significantly ameliorate the management of patients suffering from incapacitating knee and hip OA.

8. Conclusion

TJAs are successful and cost-effective interventions, yet their shortcomings aren't negligible, especially in terms of postoperative pain and function. The first step in improving management is to elaborate methods of identifying which patients or groups of patients are most vulnerable to unsuccessful outcomes, and would ultimately benefit from further attention. A systematic review of the literature identified 22 manuscripts with moderate-to-excellent methodological quality suggesting that significant determinants of poor pain and function following THA include a lower educational level, the preoperative level of pain and function, higher BMI, presence or a greater level of comorbidities, worse general health and lower radiographic OA severity. A similar methodology

identified 34 eligible studies concluding that significant determinants of poor outcomes following TKA with the greatest amount and quality of evidence are a greater social deprivation, the presence or a greater level of depression and/or anxiety, greater preoperative level of pain catastrophizing, preoperative pain or function levels, presence or greater levels of comorbidity, presence of back pain and lower general health. Using data from a retrospective cohort of 265 patients undergoing THA and HR, we developed a PA consisting of patients' age, gender, BMI and three items of the preoperative WOMAC that identified patients at risk of poor outcomes on average 15 months postoperatively with a sensitivity of 75.0% (95% CI 59.8.4-85.8), a specificity of 77.8% (95% CI 71.9-82.7) and a positive likelihood ratio of 3.38 (95% CI 2.49-4.57). Finally, in a prospective cohort of 141 patients undergoing TKA, we developed a CPR consisting of five WOMAC measured preoperatively that identifies patients at risk of poor outcomes six months following TKA with a sensitivity of 82.1% (95% CI 66.7-95.8), a specificity of 71.7% (95% CI 62.8-79.8), and a positive likelihood of 2.9 (95% CI 1.8-4.7). By systematically evaluating the literature of all determinants of TJA results and developing clinical prediction tools, we laid groundwork for future investigations of methods aimed at improving surgical outcomes. This work has the potential of finding applications at every level of care, and ultimately could improve the faring of patients with hip and knee OA.

BIBLIOGRAPHY

1. Buckwalter JA, Stanish WD, Rosier RN, Schenck Jr RC, Dennis DA, Coutts RD. The increasing need for nonoperative treatment of patients with osteoarthritis. Clin Orthop Relat Res. 2001; 385:36-45.

2. MacDonald KV, Sanmartin C, Langlois K, Marshall DA. Symptom onset, diagnosis and management of osteoarthritis. Health Reports. 2014, 25(9):10

3. Beswick AD, Wylde V, Gooberman-Hill R, Blom A, Dieppe P. What proportion of patients report long-term pain after total hip or knee replacement for osteoarthritis? A systematic review of prospective studies in unselected patients. BMJ open. 2012; 2(1).

4. Jones CA, Voaklander DC, Johnston DW, Suarez-Almazor ME. Health related quality of life outcomes after total hip and knee arthroplasties in a community based population. J Rheum. 2000; 27(7):1745-1752.

5. Wylde V, Hewlett S, Learmonth ID, Dieppe P. Persistent pain after joint replacement: prevalence, sensory qualities, and postoperative determinants. Pain. 2011; 152(3):566-572.

6. McAlindon T, Bannuru R, Sullivan M, Arden N, Berenbaum F, Bierma-Zeinstra S, et al. OARSI guidelines for the non-surgical management of knee osteoarthritis. Osteoarthritis Cartilage. 2014; 22(3):363-388.

7. Zhang W, Doherty M, Arden N, Bannwarth B, Bijlsma J, Gunther K-P, et al. EULAR evidence based recommendations for the management of hip osteoarthritis: report of a task force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). Ann Rheum Dis. 2005; 64(5):669-681.

8. Zhang W, Moskowitz R, Nuki G, Abramson S, Altman R, Arden N, et al. OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. Osteoarthritis Cartilage. 2008; 16(2):137-162.

9. Brander VA, Stulberg SD, Adams AD, Harden RN, Bruehl S, Stanos SP, et al. Predicting total knee replacement pain: a prospective, observational study. Clin Orthop & Relat Res. 2003(416):27-36.

10. Ethgen O, Bruyere O, Richy F, Dardennes C, Reginster J-Y. Health-related quality of life in total hip and total knee arthroplastyA qualitative and systematic review of the literature. The Journal of Bone & Joint Surgery. 2004; 86(5):963-974.

11. Jones CA, Voaklander DC, Johnston DWC, Suarez-Almazor ME. The effect of age on pain, function, and quality of life after total hip and knee arthroplasty. Arch Intern Med. 2001; 161(3):454.

12. Kauppila A-M, Kyllonen E, Ohtonen P, Leppilahti J, Sintonen H, Arokoski JP. Outcomes of primary total knee arthroplasty: the impact of patient-relevant factors on self-reported function and quality of life. Disabil Rehabil. 2011; 33(17-18):1659-1667.

13. Vissers MM, Bussmann JB, Verhaar JA, Arends LR, Furlan AD, Reijman M. Recovery of physical functioning after total hip arthroplasty: systematic review and meta-analysis of the literature. Phys Ther. 2011; 91(5):615-629.

14. 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):780-785.

15. Hip and Knee Replacements in Canada: Canadian Joint Replacement Registry 2014 Annual Report. Canadian Institute of Health Information 2014.

16. Singh JA, O'Byrne M, Harmsen S, Lewallen D. Predictors of moderate-severe functional limitation after primary Total Knee Arthroplasty (TKA): 4701 TKAs at 2-years and 2935 TKAs at 5-years. Osteoarthritis Cartilage. 2010; 18(4):515-521.

17. Anakwe RE, Jenkins PJ, Moran M. Predicting dissatisfaction after total hip arthroplasty: a study of 850 patients. J Arthroplasty. 2011; 26(2):209-213.

18. Santaguida PL, Hawker GA, Hudak PL, Glazier R, Mahomed NN, Kreder HJ, et al. Patient characteristics affecting the prognosis of total hip and knee joint arthroplasty: a systematic review. Can J Surg. 2008; 51(6):428.

19. Jones CA, Voaklander DC, Suarez-Alma ME. Determinants of function after total knee arthroplasty. Phys Ther. 2003; 83(8):696-706.

20. Jones CA, Beaupre LA, Johnston D, Suarez-Almazor ME. Total joint arthroplasties: current concepts of patient outcomes after surgery. Rhem Dis Clin North Am. 2007; 33(1):71-86.

21. van Jonbergen H-PW, Reuver JM, Mutsaerts EL, Poolman RW. Determinants of anterior knee pain following total knee replacement: a systematic review. Knee Surg Sports Traumatol Arthrosc. 2014; 22(3):478-499.

22. Vissers MM, Bussmann JB, Verhaar JA, Busschbach JJ, Bierma-Zeinstra S, Reijman M. Psychological factors affecting the outcome of total hip and knee arthroplasty: a systematic review. Semin Arthritis Rheum; 2012: 576-588.

23. Conner-Spady B, Arnett G, McGurran J, Noseworthy T. Prioritization of patients on scheduled waiting lists: validation of a scoring system for hip and knee arthroplasty. Can J Surg. 2004; 47(1):39 - 46.

24. Postl BD. Final report of the federal advisor on wait times: Health Canada Ottawa, 2006.

25. Hunter DJ, McDougall JJ, Keefe FJ. The symptoms of osteoarthritis and the genesis of pain. Med Clin North Am. 2009;93(1)-83-100.

26. Kopec J, Sayre E, Flanagan W, Fines P, Cibere J, Rahman MM, et al. Development of a population-based microsimulation model of osteoarthritis in Canada. Osteoarthritis Cartilage. 2010; 18(3):303-311.

27. Badley EM, Wang PP. Arthritis and the aging population: projections of arthritis prevalence in Canada 1991 to 2031. J Rheum. 1998; 25(1):138-144.

28. March LM, Bachmeier CJ. 10 Economics of osteoarthritis: a global perspective. Baillieres Clin Rheumatol. 1997; 11(4):817-834.

29. Doherty M, Dougados M. Evidence-based management of osteoarthritis: practical issues relating to the data. Best Pract Res Clin Rheum. 2001; 15(4):517-525.

30. Hip and Knee Replacements in Canada: Canadian Joint Replacement Registry 2010 Annual Report. Canadian Institute of Health Information 2010.

31. Hip and Knee Replacements in Canada: Canadian Joint Replacement Registry 2007. Canadian Institute of Health Information 2007.

32. Forster M. Survival analysis of primary cemented total knee arthroplasty. J Arthroplasty. 2003; 18(3):265-270.

33. Borus T, Thornhill T. Unicompartmental knee arthroplasty. J Am Acad Orthop Surg. 2008; 16(1):9-18.

34. Kozinn SC, Scott R. Unicondylar knee arthroplasty. J Bone Joint Surg. 1989; 71(1):145-150.

35. Carr AJ, Robertsson O, Graves S, Price AJ, Arden NK, Judge A, et al. Knee replacement. Lancet. 2012; 379(9823):1331-1340.

36. Amin AK, Patton JT, Cook RE, Gaston M, Brenkel IJ. Unicompartmental or total knee arthroplasty?: Results from a matched study. Clin Orthop Relat Res. 2006; 451:101-106.

37. Newman JH, Ackroyd CE, Shah NA. Unicompartmental or total knee replacement? Five-year results of a prospective, randomised trial of 102 osteoarthritic knees with unicompartmental arthritis. J Bone Joint Surg Br. 1998; 80(5):862-865.

38. Burnett RSJ. Total hip arthroplasty: techniques and results. B C Med J. 2010; 52(9):455-464.

39. Jarrett CA, Ranawat AS, Bruzzone M, Blum YC, Rodriguez JA, Ranawat CS. The squeaking hip: a phenomenon of ceramic-on-ceramic total hip arthroplasty. J Bone Joint Surg. 2009; 91(6):1344-1349.

40. Hutt J, Lavigne M, Vendittoli P. Total Hip Arthroplasty versus Hip Resurfacing: Evidence Based Review and Current Indications. Orthop Muscul Syst. 2014; 3(173):2161-0533.1000173.

41. Sedrakyan A, Normand S-LT, Dabic S, Jacobs S, Graves S, Marinac-Dabic D. Comparative assessment of implantable hip devices with different bearing surfaces: systematic appraisal of evidence. BMJ. 2011; 343.

42. Abdulkarim A, Ellanti P, Motterlini N, Fahey T, O'Byrne JM. Cemented versus uncemented fixation in total hip replacement: a systematic review and meta-analysis of randomized controlled trials. Orthop Rev. 2013; 5(1).

43. Beaulé PE, Antoniades J. Patient selection and surgical technique for surface arthroplasty of the hip. Orthop Clin North Am. 2005; 36(2):177-185.

44. Bourne RB. Measuring tools for functional outcomes in total knee arthroplasty. Clin Orthop Relat Res. 2008; 466(11):2634-2638.

45. Kennedy DM, Stratford PW, Robarts S, Gollish JD. Using outcome measure results to facilitate clinical decisions the first year after total hip arthroplasty. J Orthopa Sports Phys Ther. 2011; 41(4):232-239.

46. Dougados M. It's Good to Feel Better But It's Better to Feel Good. J Rheumatol 2015; 32(1): 1-2

47. Chesworth BM, Mahomed NN, Bourne RB, Davis AM, Group OS. Willingness to go through surgery again validated the WOMAC clinically important difference from THR/TKR surgery. J Clin Epidemiol. 2008; 61(9):907-918.

48. Escobar A, Gonzalez M, Quintana JM, Vrotsou K, Bilbao A, Herrera-Espineira C, et al. Patient acceptable symptom state and OMERACT-OARSI set of responder criteria in joint replacement. Identification of cut-off values. Osteoarthritis Cartilage. 2012; 20(2):87-92.

49. Escobar A, Quintana JM, Bilbao A, Arostegui I, Lafuente I, Vidaurreta I. Responsiveness and clinically important differences for the WOMAC and SF-36 after total knee replacement. Osteoarthritis Cartilage. 2007; 15(3):273-280.

50. Quintana JM, Aguirre U, Barrio I, Orive M, Garcia S, Escobar A. Outcomes after total hip replacement based on patients' baseline status: what results can be expected? Arthritis Care Res. 2012; 64(4):563-572.

51. Quintana JM, Escobar A, Bilbao A, Arostegui I, Lafuente I, Vidaurreta I. Responsiveness and clinically important differences for the WOMAC and SF-36 after hip joint replacement. Osteoarthritis Cartilage. 2005; 13(12):1076-1083.

52. 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 Rheum. 1999; 42(8):1722-1728.

53. Fortin PR, Penrod JR, Clarke AE, St-Pierre Y, Joseph L, Belisle P, et al. Timing of total joint replacement affects clinical outcomes among patients with osteoarthritis of the hip or knee. Arthritis Rheum. 2002; 46(12):3327-3330.

54. Mahomed NN, Liang MH, Cook EF, Daltroy LH, Fortin PR, Fossel AH, et al. The importance of patient expectations in predicting functional outcomes after total joint arthroplasty. J Rheumatol. 2002; 29(6):1273-1279.

55. Hernigou P, Flouzat-Lachianette C-H, Jalil R, Batista SU, Guissou I, Poignard A. Treatment of infected hip arthroplasty. Open Orthop J. 2010; 4:126.

56. Leopold SS. Minimally invasive total knee arthroplasty for osteoarthritis. New Engl J Med. 2009; 360(17):1749-1758.

57. DeHart MM, Riley L. Nerve injuries in total hip arthroplasty. J Am Acad Orthop Surg. 1999; 7(2):101-111.

58. Schinsky MF, Macaulay W, Parks ML, Kiernan H, Nercessian OA. Nerve injury after primary total knee arthroplasty. J Arthroplasty. 2001; 16(8):1048-1054.

59. Douketis JD, Eikelboom JW, Quinlan DJ, Willan AR, Crowther MA. Short-duration prophylaxis against venous thromboembolism after total hip or knee replacement: a meta-analysis of prospective studies investigating symptomatic outcomes. Arch Int Med. 2002; 162(13):1465.

60. White RH, Romano PS, Zhou H, Rodrigo J, Bargar W. Incidence and time course of thromboembolic outcomes following total hip or knee arthroplasty. Arch Intern Med. 1998; 158(14):1525-1531.

61. Ulrich SD, Seyler TM, Bennett D, Delanois RE, Saleh KJ, Thongtrangan I, et al. Total hip arthroplasties: what are the reasons for revision? Int Orthop. 2008; 32(5):597-604.

62. Vessely MB, Whaley AL, Harmsen WS, Schleck CD, Berry DJ. The Chitranjan Ranawat Award: Long-term survivorship and failure modes of 1000 cemented condylar total knee arthroplasties. Clin Orthop Relat Res. 2006; 452:28-34. 97.

63. Laupacis A, Sekar N. Clinical prediction rules: a review and suggested modifications of methodological standards. JAMA. 1997; 277(6):488-494.

64. McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS, et al. Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. JAMA. 2000; 284(1):79-84.

65. Adams ST, Leveson SH. Clinical prediction rules. BMJ. 2012; 344.

66. Zhang H, Singer B. Recursive partitioning and applications: Springer Science & Business

Media, 2010.

67. Strobl C, Malley J, Tutz G. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. Psychol Methods. 2009; 14(4):323.

68. Steyerberg EW, Harrell FE, Borsboom GJ, Eijkemans M, Vergouwe Y, Habbema JDF. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol. 2001; 54(8):774-781.

69. Slaven EJ. Prediction of functional outcome at six months following total hip arthroplasty. Phys Ther. 2012; 92(11):1386-1394.

70. 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):290-296.

71. Stiell IG, Greenberg GH, McKnight RD, Nair RC, McDowell I, Worthington JR. A study to develop clinical decision rules for the use of radiography in acute ankle injuries. Ann Emerg Med. 1992; 21(4):384-390.

72. Stiell IG GG, Wells GA, McKnight RD, Cwinn AA, Cacciotti T, McDowell I, Smith NA. Derivation of a decision rule for the use of radiography in acute knee injuries. Ann Emerg Med. 26(4):405-413.

73. Dionne CE, Le Sage N, Franche RL, Dorval M, Bombardier C, Deyo RA. Five questions predicted long-term, severe, back-related functional limitations: evidence from three large prospective studies. J Clin Epidemiol. 2011; 64(1):54-66.

74. Hayden JA, Côté P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. Ann Int Med. 2006; 144(6):427-437.

75. Bellamy N. Validation study of WOMAC: a health status instrument for measuring clinically-important patient-relevant outcomes following total hip or knee arthroplasty in osteoarthritis. J Orthop Rheumatol. 1988; 1:95-108.

76. McConnell S, Kolopack P, Davis AM. The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC): a review of its utility and measurement properties. Arthritis Care Res. 2001; 45(5):453-461.

77. Bellamy N. WOMAC: a 20-year experimental review of a patient-centered self-reported health status questionnaire. J Rheumatol. 2002; 29(12):2473-2476.

78. Bellamy N. WOMAC osteoarthritis index: user guide IX: Nicholas Bellamy, 2008.

79. Collins M, Lavigne M, Girard J, Vendittoli P-A. Joint perception after hip or knee replacement surgery. Orthop Traumatol Surg Res. 2012; 98(3):275-280

80. Charnley J, Halley DK. Rate of wear in total hip replacement. Clin Orthop and Relat Res. 1975; 112:170-179.

81. Söderman P, Malchau H, Herberts P, Zügner R, Regnér H, Garellick G. Outcome after total hip arthroplasty: Part II. Disease-specific follow-up and the Swedish National Total Hip Arthroplasty Register. Acta Orthop. 2001; 72(2):113-119.

82. Breinman L FJ, Olshen R, Stome C. Classification and regression trees: Pacific Grove Wadsworth, 1984.

83. Strobl C, Boulesteix A-L, Augustin T. Unbiased split selection for classification trees based on the Gini index. Comput Stat Data An. 2007; 52(1):483-501.

84. Simel DL, Samsa GP, Matchar DB. Likelihood ratios with confidence: sample size estimation for diagnostic test studies. J Clin Epidemiol. 1991; 44(8):763-770.

85. Efron B TR. An introduction to the bootstrap. New York, NY: Chapman & Hall, 1993.

86. Miller MD, Paradis CF, Houck PR, Mazumdar S, Stack JA, Rifai AH, et al. Rating chronic medical illness burden in geropsychiatric practice and research: application of the Cumulative Illness Rating Scale. Psychiatry Res. 1992; 41(3):237-248.

87. Daveluy C, L. Pica, N. Audet, R. Courtemanche, F. Lapointe. Enquête sociale et de santé 1998, 2e édition. Québec: Institut de la statistique du Québec 2000.

88. Ilfeld FW, Jr. Further validation of a psychiatric symptom index in a normal population. Psychol Rep. 1976; 39:1215-1228.

89. Dowsey MM, Nikpour M, Dieppe P, Choong PF. Associations between preoperative radiographic osteoarthritis severity and pain and function after total hip replacement : Radiographic OA severity predicts function after THR. Clin Rheumatol. 2016; 35(1):183-189.

90. Holstege MS, Lindeboom R, Lucas C. Preoperative quadriceps strength as a predictor for short-term functional outcome after total hip replacement. Arch Phys Med Rehabil. 2011; 92(2):236-241.

91. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Prev Med. 2007; 45(4):247-251.

92. Gandhi R, Razak F, Davey JR, Mahomed NN. Metabolic syndrome and the functional outcomes of hip and knee arthroplasty. J Rheumatol. 2010; 37(9):1917-1922.

93. Kessler S, Kafer W. Overweight and obesity: two predictors for worse early outcome in total

hip replacement? Obesity. 2007; 15(11):2840-2845. 90.

94. Quintana JM, Escobar A, Aguirre U, Lafuente I, Arenaza JC. Predictors of health-related quality-of-life change after total hip arthroplasty. Clin Orthop Relat Res. 2009; 467(11):2886-2894.

95. Stevens M, Paans N, Wagenmakers R, van Beveren J, van Raay JJAM, van der Meer K, et al. The influence of overweight/obesity on patient-perceived physical functioning and health-related quality of life after primary total hip arthroplasty. Obesity Surg. 2012; 22(4):523-529. 85.

96. Braeken AM, Lochhaas-Gerlach JA, Gollish JD, Myles JD, Mackenzie TA. Determinants of 6-12 month postoperative functional status and pain after elective total hip replacement. Intern J Qual Health Care. 1997; 9(6):413-418.

97. Judge A, Batra RN, Thomas GE, Beard D, Javaid MK, Murray DW, et al. Body mass index is not a clinically meaningful predictor of patient reported outcomes of primary hip replacement surgery: prospective cohort study. Osteoarthritis Cartilage. 2014; 22(3):431-439.

98. Judge A, Cooper C, Arden NK, Williams S, Hobbs N, Dixon D, et al. Preoperative expectation predicts 12-month postoperative outcome among patients undergoing primary total hip replacement in European orthopaedic centres. Osteoarthritis Cartilage. 2011; 19(6):659-667.

99. Ramaesh R, Jenkins P, Lane JV, Knight S, Macdonald D, Howie C. Personality, function and satisfaction in patients undergoing total hip or knee replacement. J Orthop Sci. 2013.

100. Judge A, Cooper C, Williams S, Dreinhoefer K, Dieppe P. Patient-reported outcomes one year after primary hip replacement in a European Collaborative Cohort. Arthritis Care Res. 2010; 62(4):480-488.

101. Jenkins PJ, Perry PRW, Yew Ng C, Ballantyne JA. Deprivation influences the functional outcome from total hip arthroplasty. Surg J R Coll Surg E. 2009; 7(6):351-356.

102. Vergara I, Bilbao A, Gonzalez N, Escobar A, Quintana JM. Factors and consequences of waiting times for total hip arthroplasty. Clin Orthop Relat Res. 2011; 469(5):1413-1420.
103. Caracciolo B, Giaquinto S. Determinants of the subjective functional outcome of total joint arthroplasty. Arch Gerontol Geriatr. 2005; 41(2):169-176.

104. Dowsey MM, Nikpour M, Dieppe P, Choong PF. Associations between preoperative radiographic changes and outcomes after total knee joint replacement for osteoarthritis. Osteoarthritis Cartilage. 2012;29(10):1095-1102.

105. Garbuz DS, Xu M, Duncan CP, Masri BA, Sobolev B. Delays worsen quality of life outcome of primary total hip arthroplasty. Clin Orthop Relat Res. 2006; 447:79-84.

106. Peter WF, Dekker J, Tilbury C, Tordoir RL, Verdegaal SH, Onstenk R, et al. The association between comorbidities and pain, physical function and quality of life following hip and knee arthroplasty. Rheumatol Intern. 2015;35(7):1233-1241

107. Wylde V, Sayers A, Lenguerrand E, Gooberman-Hill R, Pyke M, Beswick AD, et al. Preoperative widespread pain sensitization and chronic pain after hip and knee replacement: a cohort analysis. Pain. 2015; 156(1):47-54. 128.

108. Carr JL, Moffett JAK. Review Paper: The impact of social deprivation on chronic back pain outcomes. Chronic Illn. 2005; 1(2):121-129.

109. Symmons DP. Environmental factors and the outcome of rheumatoid arthritis. Best Practe Res Clin Rheumatol. 2003; 17(5):717-727.

110. Katz NP, Paillard FC, Ekman E. Determining the clinical importance of treatment benefits for interventions for painful orthopedic conditions. J Orthop Surg Res. 2015; 10(1):24.

111. Kinds M, Welsing P, Vignon E, Bijlsma J, Viergever M, Marijnissen A, et al. A systematic review of the association between radiographic and clinical osteoarthritis of hip and knee. Osteoarthritis Cartilage. 2011; 19(7):768-778.

112. Garbuz DS, Tanzer M, Greidanus NV, Masri BA, Duncan CP. The John Charnley Award: Metal-on-metal hip resurfacing versus large-diameter head metal-on-metal total hip arthroplasty: a randomized clinical trial. Clin Orthop Relat Res. 2010; 468(2):318-325.

113. Girard J, Vendittoli P, Roy A, Lavigne M. [Femoral offset restauration and clinical function after total hip arthroplasty and surface replacement of the hip: a randomized study]. Rev Chir Orthop Reparatrice Appar Mot. 2008; 94(4):376-381.

114. Shimmin A, Baré J. Comparison of functional results of hip resurfacing and total hip replacement: a review of the literature. Orthope Clin North Am. 2011; 42(2):143.

115. Vendittoli P, Rivière C, Roy A, Barry J, Lusignan D, Lavigne M. Metal-on-metal hip resurfacing compared with 28-mm diameter metal-on-metal total hip replacement A randomised study with six to nine years' follow-up. Bone Joint J. 2013; 95(11):1464-1473. 144.

116. Bellamy N, Buchanan W. A preliminary evaluation of the dimensionality and clinical importance of pain and disability in osteoarthritis of the hip and knee. Clin Rheumatol. 1986; 5(2):231-241. 188.

117. Kleinbaum DG. Applied regression analysis and multivariable methods: CengageBrain. com, 2007.

118. NIH Consensus Statement on total knee replacement. NIH Consens State Sci Statements. 2003; 20(1):1-34. 146.

119. Liang MH, Cullen KE, Larson MG, Thompson MS, Schwartz JA, Fossel AH, et al. Cost-effectiveness of total joint arthroplasty in osteoarthritis. Arthritis Rheum. 1986; 29(8):937-943. 149.

120. 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):1058-1066.

121. Desmeules F, Dionne CE, Belzile ÉL, Bourbonnais R, Champagne F, Frémont P. Determinants of pain, functional limitations and health-related quality of life six months after total knee arthroplasty: results from a prospective cohort study. BMC Sports Sci Med Rehabil. 2013; 5(1):2.

122. Lopez-Olivo MA, Landon GC, Siff SJ, Edelstein D, Pak C, Kallen MA, et al. Psychosocial determinants of outcomes in knee replacement. Ann Rheumat Dis. 2011; 70(10):1775-1781.

123. Perruccio AV, Power JD, Evans HMK, Mahomed SR, Gandhi R, Mahomed NN, et al. Multiple joint involvement in total knee replacement for osteoarthritis: Effects on patient-reported outcomes. Arthritis Care Res. 2012; 64(6):838-846.

124. Rajgopal V, Bourne RB, Chesworth BM, MacDonald SJ, McCalden RW, Rorabeck CH. The impact of morbid obesity on patient outcomes after total knee arthroplasty. J Arthroplasty. 2008; 23(6):795-800.

125. Gandhi R, Santone D, Takahashi M, Dessouki O, Mahomed NN. Inflammatory predictors of ongoing pain 2 years following knee replacement surgery. Knee. 2013; 20(5):316-318.

126. Riddle DL, Wade JB, Jiranek WA, Kong X. Preoperative pain catastrophizing predicts pain outcome after knee arthroplasty. Clin Orthop Relat Res. 2010; 468(3):798-806.

127. Neuburger J, Hutchings A, Black N, van der Meulen JH. Socioeconomic differences in patient-reported outcomes after a hip or knee replacement in the English National Health Service. J Public Health. 2013; 35(1):115-124.

128. Escobar A, Quintana JM, Bilbao A, Azkarate J, Güenaga J, Arenaza J, et al. Effect of patient characteristics on reported outcomes after total knee replacement. Rheumatology. 2007; 46(1):112-119.

129. Judge A, Arden NK, Cooper C, Kassim Javaid M, Carr AJ, Field RE, et al. Predictors of outcomes of total knee replacement surgery. Rheumatology. 2012; 51(10):1804-1813.

130. Lingard EA, Katz JN, Wright EA, Sledge CB, Kinemax Outcomes G. Predicting the outcome of total knee arthroplasty. J Bone Joint Surg Am. 2004; 86-A(10):2179-2186.

131. Papakostidou I, Dailiana ZH, Papapolychroniou T, Liaropoulos L, Zintzaras E, Karachalios

TS, et al. Factors affecting the quality of life after total knee arthroplasties: a prospective study. BMC Musculoskelet Disord. 2012; 13:116.

132. Sullivan M, Tanzer M, Reardon G, Amirault D, Dunbar M, Stanish W. The role of presurgical expectancies in predicting pain and function one year following total knee arthroplasty. Pain. 2011; 152(10):2287-2293.

133. Sullivan M, Tanzer M, Stanish W, Fallaha M, Keefe FJ, Simmonds M, et al. Psychological determinants of problematic outcomes following Total Knee Arthroplasty. Pain. 2009; 143(1-2):123-129.

134. Wylde V, Dixon S, Blom AW. The role of preoperative self-efficacy in predicting outcome after total knee replacement. Musculoskelet. 2012; 10(2):110-118.

135. Clement ND, MacDonald D, Simpson AHRW, Burnett R. Total knee replacement in patients with concomitant back pain results in a worse functional outcome and a lower rate of satisfaction. Bone Joint J. 2013; 95-B(12):1632-1639.

136. Hanusch BC, O'Connor DB, Ions P, Scott A, Gregg PJ. Effects of psychological distress and perceptions of illness on recovery from total knee replacement. Bone Joint J. 2014; 96-B(2):210-216.

137. Alzahrani K, Gandhi R, Debeer J, Petruccelli D, Mahomed N. Prevalence of clinically significant improvement following total knee replacement. J Rheumatol. 2011; 38(4):753-759.

138. Davis ET, Lingard EA, Schemitsch EH, Waddell JP. Effects of socioeconomic status on patients' outcome after total knee arthroplasty. Intern J Qual Health Care. 2008; 20(1):40-46.

139. Clement ND, MacDonald D, Burnett R, Breusch SJ. Diabetes does not influence the early outcome of total knee replacement: a prospective study assessing the Oxford knee score, short form 12, and patient satisfaction. Knee. 2013; 20(6):437-441.

140. Yakobov E, Scott W, Stanish W, Dunbar M, Richardson G, Sullivan M. The role of perceived injustice in the prediction of pain and function after total knee arthroplasty. Pain. 2014; 155(10):2040-2046.

141. Lingard EA, Riddle DL. Impact of psychological distress on pain and function following knee arthroplasty. J Bone Joint Surg Am. 2007; 89(6):1161-1169.

142. 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):52-58.

143. Ayers DC, Franklin PD, Ploutz-Snyder R, Boisvert CB. Total knee replacement outcome and coexisting physical and emotional illness. Clin Orthop Relat Res. 2005; 440:157-161.

144. Clement N, Macdonald D, Simpson A. The minimal clinically important difference in the Oxford knee score and Short Form 12 score after total knee arthroplasty. Knee Surg Sports Traumatol Arthrosc. 2013:1-7.

145. Engel C, Hamilton NA, Potter PT, Zautra AJ. Impact of two types of expectancy on recovery from total knee replacement surgery (TKR) in adults with osteoarthritis. Behav Med. 2004; 30(3):113-123.

146. Kauppila AM, Sintonen H, Aronen P, Ohtonen P, Kyllonen E, Arokoski JP. Economic evaluation of multidisciplinary rehabilitation after primary total knee arthroplasty based on a randomized controlled trial. Arthritis Care Res (Hoboken). 2011; 63(3):335-341.

147. Benjamin J, Johnson R, Porter S. Knee scores change with length of follow-up after total knee arthroplasty. J Arthroplasty. 2003; 18(7):867-871.

148. Desmeules F, Dionne CE, Belzile E, Bourbonnais R, Fremont P. The burden of wait for knee replacement surgery: effects on pain, function and health-related quality of life at the time of surgery. Rheumatology (Oxford). 2010; 49(5):945-954.

149. Mainard D, Guillemin F, Cuny C, Mejat-Adler E, Galois L, Delagoutte J. Quality of life assessment one year after total hip or knee arthroplasty]. Rev Chir Orthop Réparatrice App Moteur. 2000; 86(5):464.

150. Parent É, Moffet H. Preoperative predictors of locomotor ability two months after total knee arthroplasty for severe osteoarthritis. Arthritis Care Res. 2003; 49(1):36-50.

151. Singh JA, Lewallen D. The impact of gender, age, and preoperative pain severity on pain after TKA. Clin Orthop Relat Res. 2008; 466(11):2717-2723.

152. Valdes AM, Doherty SA, Zhang W, Muir KR, Maciewicz RA, Doherty M. Inverse relationship between preoperative radiographic severity and postoperative pain in patients with osteoarthritis who have undergone total joint arthroplasty. Semin Arthritis Rheum; 2012: 568-575.

153. Desmeules F, Dionne CE, Belzile EL, Bourbonnais R, Fremont P. The impacts of pre-surgery wait for total knee replacement on pain, function and health-related quality of life six months after surgery. J Eval Clin Pract. 2010.

154. Faucher M, Poiraudeau S, Lefevre-Colau MM, Rannou F, Fermanian J, Revel M. Assessment of the test-retest reliability and construct validity of a modified WOMAC index in knee osteoarthritis. Joint Bone Spine. 2004; 71(2):121-127.

155. Akobeng AK. Understanding diagnostic tests 1: sensitivity, specificity and predictive values. Acta Paediatrica. 2007; 96(3):338-341.

156. Kennedy DM, Stratford PW, Hanna SE, Wessel J, Gollish JD. Modeling early recovery of physical function following hip and knee arthroplasty. BMC Musculoskelet Disord. 2006; 7(1):100.

157. Fehring TK, Odum SM, Griffin WL, Mason JB, McCoy TH. The obesity epidemic: its effect on total joint arthroplasty. J Arthroplasty. 2007; 22(6):71-76.

158. Spahn G, Kaiser M, Gantz S, Gantz MS, Gantz BH, Schiele R, et al. Risk factors for hip osteoarthritis. Results of a systematic review and meta-analysis. ASU Intern. 2013(11).

159. Kellgren J, Lawrence J. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957, 16(4):494. 192.

160. Bouwmeester W, Zuithoff NP, Mallett S, Geerlings MI, Vergouwe Y, Steyerberg EW, et al. Reporting and methods in clinical prediction research: a systematic review. PLoS Med. 2012; 9(5):641.

161. Gill GS, Mills D, Joshi AB. Mortality following primary total knee arthroplasty. J Bone Joint Surg. 2003; 85(3):432-435.

162. SooHoo NF, Lieberman JR, Ko CY, Zingmond DS. Factors predicting complication rates following total knee replacement. J Bone Joint Surg 2006; 88(3):480-485.

163. Hofstede SN, Nouta KA, Jacobs W, van Hooff ML, Wymenga AB, Pijls BG, et al. Mobile bearing vs fixed bearing prostheses for posterior cruciate retaining total knee arthroplasty for postoperative functional status in patients with osteoarthritis and rheumatoid arthritis. Cochrane Libr. 2015.

164. Verra WC, van den Boom L, Jacobs W, Clement DJ, Wymenga A, Nelissen R. Retention versus sacrifice of the posterior cruciate ligament in total knee arthroplasty for treating osteoarthritis. Cochrane Database Syst Rev. 2013; 10.

165. Blagojevic M, Jinks C, Jeffery A, Jordan K. Risk factors for onset of osteoarthritis of the knee in older adults: a systematic review and meta-analysis. Osteoarthritis cartilage. 2010; 18(1):24-33.

166. Desmeules F, Dionne CE, Belzile É, Bourbonnais R, Frémont P. Waiting for total knee replacement surgery: factors associated with pain, stiffness, function and quality of life. BMC Musculoskelet Disord. 2009; 10(1):52.

APPENDIX A - Anchor items and respective cut-off responses employed to determine the estimates of appropriate outcomes following TKA and THA

Study	Method of estimation of appropriate improvement	Anchor item	Cut-off responses
Quintana et al. (2012)	PASS (75th percentile of postoperative score)	''If you had to live the rest of your life with the hip symptoms you have now, how would you feel?''	Very satisfied or somewhat satisfied
	ROC		Satisfied vs. Not satisfied (Satisfied : having answered Very satisfied or Somewhat satisfied to the anchor item Not Satisfied: remainder of answers)
	Mean change (Positive MCID)	''Compared with your status before being placed the hip prosthesis, how would you rate the status of your hip right now?''	Somewhat better
Quintana et al. (2005)	Mean change (Positive MCID)	Not explicitly stated: ''Transitional question about […] joint improvement after the intervention''.	Somewhat better
Escobar et al. (2007)	Mean change (Positive MCID)	Not explicitly stated: '' All patients had to answer question about its (sic) improvement in the their knee at 6 months and 2 years''	Somewhat better
Escobar et al (2012) [	PASS (75th percentile of postoperative score)	''If you had to live the rest of your life with the hip symptoms you have now, how would you feel?''	Very satisfied or somewhat satisfied
	ROC		Satisfied vs. Not satisfied (Satisfied: having answered Very satisfied or Somewhat satisfied to the anchor item Not Satisfied: remainder of answers)
	OMERACT-OARSI (75th percentile of postoperative score)	Responder if: 1) Improvement in pain or function (WOMAC pain and function score respectively) ≥ 50% and absolute change (total WOMAC score) ≥ 20 OR 2) At least two of the following: a. Pain ≥ 20% and absolute change ≥ 10 b. Function ≥20% and absolute change ≥ 10 c. Patients global assessment of the disease ≥ 20 % (1 point change on the Likert scare = 20% global assessment change) and absolute change ≥ 10	

Chesworth et al. (2007)	CID (ROC)	**Derivation CID**: Not explicitly stated; ''[…] compared to when they (patients) went on the wait list for surgery, (were) they better, worse or the same (?)''	**Derivation CID** **Positive CID** If initial response is ''Better'', analysis was done on: ''a good deal better'' or ''a great deal better'' or ''a very great deal better'' **vs.** ''a very great deal worse'' and ''a great deal worse'' and ''a good deal worse'' and ''moderately worse'' and ''somewhat worse'' and ''a little worse'' and '' almost the same-hardly any worse'' and ''same'' and ''hardly any better'' and ''a little better'' and ''somewhat better'' and ''moderately better'' **Negative CID** If initial response is ''Worse'', analysis was done on: ''a very great deal worse'' and ''a great deal worse'' and ''a good deal worse'' and ''moderately worse'' and ''somewhat worse'' and ''a little worse'' and '' almost the same-hardly any worse'' **vs**. ''same'' and ''hardly any better'' and ''a little better'' and ''somewhat better'' and ''moderately better'' and ''a good deal better'' or ''a great deal better'' or ''a very great deal better''

		Validation CID: '' knowing what your hip or knee replacement surgery did for you, if you could go back in time, would you still have undergone this surgery (?)''	**Validation CID** **Positive CID** ''Yes'' vs. ''No'' or ''Uncertain'' **Negative CID** ''No'' or ''Uncertain'' vs. 'Yes''

WOMAC – Western Ontario and McMaster Universities Osteoarthritis Index; PASS – Patient Acceptable Symptom Scale; MCID – Minimal Clinically Important Difference; ROC – Receiver Operating Characteristics; OMERACT-OARSI – Outcome Measures in Rheumatology-Osteoarthritis Research Society International responder criteria; CID – Clinically Important Difference

APPENDIX B - Estimates of appropriate and inappropriate outcomes as measured with the WOMAC following TKA and THA

Study	Type of surgery	Time point after intervention	WOMAC domain	Estimate of outcome
Quintana et al. (2012)	Primary THA	6 months	Pain	**Derivation Cohort** **PASS** 1st quartile of preoperative total WOMAC scores: **20** 2nd quartile of preoperative total WOMAC scores: **25** 3rd tertile of preoperative total WOMAC scores: **25** **Positive MCID** 1st quartile of preoperative total WOMAC scores: **14.66 ± 12.55** 2nd quartile of preoperative total WOMAC scores: **23.29 ± 16.99** 3rd tertitle of preoperative total WOMAC scores: **23.29 ± 16.99** **ROC** 1st quartile of preoperative total WOMAC scores: **19** 2nd quartile of preoperative total WOMAC scores: **25** 3rd tertitle of preoperative total WOMAC scores: **25** **Validation Cohort** **PASS** 1st quartile of preoperative total WOMAC scores: **20** 2nd quartile of preoperative total WOMAC scores: **25** 3rd tertitle of preoperative total WOMAC scores: **25** **ROC** 1st quartile of preoperative total WOMAC scores: **20** 2nd quartile of preoperative total WOMAC scores: **25** 3rd tertitle of preoperative total WOMAC scores: **25**
		6 months	Function	**Derivation Cohort** **PASS** 1st quartile of preoperative total WOMAC scores: **28** 2nd quartile of preoperative total WOMAC scores: **35** 3rd tertitle of preoperative total WOMAC scores: **42** **Positive MCID** 1st quartile of preoperative total WOMAC scores: **9** 2nd quartile of preoperative total WOMAC scores: **22** 3rd tertitle of preoperative total WOMAC scores: **1** **ROC** 1st quartile of preoperative total WOMAC scores: **26.4**

				2nd quartile of preoperative total WOMAC scores: **39** 3rd tertitle of preoperative total WOMAC scores: **40** **Validation Cohort** **PASS** 1st quartile of preoperative total WOMAC scores: **32** 2nd quartile of preoperative total WOMAC scores: **32** 3rd tertitle of preoperative total WOMAC scores: **40** **ROC** 1st quartile of preoperative total WOMAC scores: **25** 2nd quartile of preoperative total WOMAC scores: **39** 3rd tertitle of preoperative total WOMAC scores: **40**
Quintana et al. (2005)	Primary THA	6 months	Pain	**Positive MCID** : 29.26 (16.90)*
			Function	**Positive MCID** : 26.54 (17.79)*
			Stiffness	**Positive MCID** : 25.91 (22.94)*
		2 years	Pain	**Positive MCID** : 33.13 (19.04)*
			Function	**Positive MCID** : 25.93 (18.04)*
			Stiffness	**Positive MCID** : 33.20 (18.68)*
Escobar et al. (2007)	Primary TKA	6 months	Pain	**Positive MCID** : 22.87 (18.13)*
			Function	**Positive MCID** : 19.01 (17.48)*
			Stiffness	**Positive MCID** : 14.53 (26.50)*
		2 years	Pain	**Positive MCID** : 27.98 (19.44)*
			Function	**Positive MCID** : 20.84 (18.20)*
			Stiffness	**Positive MCID** : 21.35 (27.22)*
Escobar et al. (2012)	Primary THA	1 year	Pain	**PASS**:** Worst quartile of preoperative total WOMAC scores: **25.0 (15.0 – 30.0)***** Medium quartile of preoperative total WOMAC scores: **20.0 (15.0 – 25.0)***** Worst quartile of preoperative total WOMAC scores: **20.0 (15.0 – 25.0)***** Full sample: **20.0 (15.0 – 25.0)***** **OMERACT-OARSI**:** Worst quartile of preoperative total WOMAC scores: **25.0 (15.0 – 30.0)***** Medium quartile of preoperative total WOMAC scores: **20.0 (15.0 – 25.0)***** Worst quartile of preoperative total WOMAC scores: **15.0 (15.0 – 25.0)***** Full sample: **20.0 (15.0 – 25.0)***** **ROC****: **30.0 (29.4 – 30.6)*****

			Function	**PASS**:** Worst quartile of preoperative total WOMAC scores: **39.9 (32.8 – 42.7**)*** Medium quartile of preoperative total WOMAC scores: **27.9 (23.5 – 36.8)***** Worst quartile of preoperative total WOMAC scores: **25.0 (20.6 – 29.4)***** Full sample: **30.9 (27.9 – 32.4)***** **OMERACT-OARSI**:** Worst quartile of preoperative total WOMAC scores: **41.2 (37.5 – 45.6)***** Medium quartile of preoperative total WOMAC scores: **28.0 (25.0 – 36.8**)*** Worst quartile of preoperative total WOMAC scores: **26.5 (23.5 – 30.9)***** Full sample: **32.4 (31.2 – 35.3)***** **ROC****: **31.2 (30.7 – 31.6)*****
	Primary TKA	1 year	Pain	**PASS**:** Worst quartile of preoperative total WOMAC scores: **30.0 (25.0 – 35.0)***** Medium quartile of preoperative total WOMAC scores: **25.0 (20.0 – 30.0)***** Worst quartile of preoperative total WOMAC scores: **20.0 (15.0 – 25.0)***** Full sample: **25.0 (20.0 – 30.0)***** **OMERACT-OARSI**:** Worst quartile of preoperative total WOMAC scores: **30.0 (25.0 – 35.0**)*** Medium quartile of preoperative total WOMAC scores: **25.0 (20.0– 30.0)***** Worst quartile of preoperative total WOMAC scores: **15.0 (12.5 – 20.0)***** Full sample: **25.0 (20.0 – 30.0)***** **ROC****: **28.6 (28.2 – 28.9**)***
			Function	**PASS**:** Worst quartile of preoperative total WOMAC scores: **39.7 (33.8 – 48.4)***** Medium quartile of preoperative total WOMAC scores: **32.3 (27.9 – 35.3)***** Worst quartile of preoperative total WOMAC scores: **27.2 (25.0 – 30.9)***** Full sample: **32.3 (29.4 – 35.3)***** **OMERACT-OARSI**:** Worst quartile of preoperative total WOMAC scores: **41.2 (36.8 – 47.1)***** Medium quartile of preoperative total WOMAC scores: **33.3 (30.9– 36.8)***** Worst quartile of preoperative total WOMAC scores: **25.4 (23.5 – 29.4)***** Full sample: **33.8 (30.9 – 35.3)***** **ROC****: **36.7 (36.4 – 37.1)*****

Chesworth et al. (2007)	Primary THA	1 year	Pain	**<u>Derivation:</u>** **Positive CID: 41** **Negative CID: 35** **<u>Validation</u>** **Positive CID: 36** **Negative CID: 20**
			Function	**<u>Derivation:</u>** **Positive CID: 34** **Negative CID: 33** **<u>Validation</u>** **Positive CID: 34** **Negative CID: 33**
	Primary TKA		Pain	**<u>Derivation:</u>** **Positive CID: 36** **Negative CID: 30** **<u>Validation</u>** **Positive CID: 31** **Negative CID: 25**
			Function	**<u>Derivation:</u>** **Positive CID: 33** **Negative CID: 25** **<u>Validation</u>** **Positive CID: 26** **Negative CID: 20**

* Results reported as mean (standard deviation)
** WOMAC values are transformed on a standardized 0-100 scale where 0 indicates the best possible status and 100 the worst status; original study (Escobar et al, 2012) reports a score of 0 as the worst status and a score of 100 as the best possible status.
** 95 % confidence interval
WOMAC – Western Ontario and McMaster Universities Osteoarthritis Index; THA – Total Hip Arthroplasty; TKA – Total Knee Arthroplasty: PASS – Patient Acceptable Symptom Scale; MCID – Minimal Clinically Important Difference; ROC – Receiver Operating Characteristics; OMERACT-OARSI – Outcome Measures in Rheumatology-Osteoarthritis Research Society International responder criteria; CID – Clinically Important Difference

APPENDIX C - Pubmed/Medline search strategy for the systematic review on determinants of THA outcomes

((((((tha OR hip arthroplast* OR thr OR hip replacement*[Title/Abstract])) OR ((((hip[MeSH Terms]) OR hip joint[MeSH Terms])) AND ((hip prosthesis[MeSH Terms]) OR arthroplasty, replacement[MeSH Terms]))) OR arthroplasty, replacement, hip[MeSH Terms])) AND (((predict* OR determinant* OR determinat* OR determining* OR contribut* OR impact* OR influenc* OR predispos* OR characteristic* OR associat* OR affect*)) AND (((((recovery of function[MeSH Terms]) OR ((postoperative period[MeSH Terms]) AND pain[Title/Abstract])) OR ((pain, postoperative[MeSH Terms]) OR pain measurement)) OR "clinical outcome") OR (((pain OR function*)) AND (postoperati* OR post-operati* OR postsurgical* OR post-surgical* OR outcome* OR limitation* OR status OR disabilit* OR recovery))))

APPENDIX D - Pubmed/Medline search strategy for the systematic review on determinants of TKA outcomes

((((((tka OR knee arthroplast* OR tkr OR knee replacement*[Title/Abstract])) OR ((((knee[MeSH Terms]) OR knee joint[MeSH Terms])) AND ((knee prosthesis[MeSH Terms]) OR arthroplasty, replacement[MeSH Terms]))) OR arthroplasty, replacement, knee[MeSH Terms])) AND (((predict* OR determinant* OR determinat* OR determining* OR contribut* OR impact* OR influenc* OR predispos* OR characteristic* OR associat* OR affect*)) AND (((((recovery of function[MeSH Terms]) OR ((postoperative period[MeSH Terms]) AND pain[Title/Abstract])) OR ((pain, postoperative[MeSH Terms]) OR pain measurement)) OR "clinical outcome") OR (((pain OR function*)) AND (postoperati* OR post-operati* OR postsurgical* OR post-surgical* OR outcome* OR limitation* OR status OR disabilit* OR recovery))))

APPENDIX E - Methodology Checklist for Prognostic Studies

The criteria used in this checklist are adapted from: Hayden JA, Cote P, Bombardier C (2006) Evaluation of the quality of prognosis studies in systematic reviews. Annals of Internal Medicine 144: 427–37.

Study identification *Include author, title, reference, year of publication*				
Guideline topic:		**Review question no:**		
Checklist completed by:				
		Circle one option for each question		
1.1	The study sample represents the population of interest with regard to key characteristics, sufficient to limit potential bias to the results	Yes	No	Unclear
1.2	Loss to follow-up is unrelated to key characteristics (that is, the study data adequately represent the sample), sufficient to limit potential bias	Yes	No	Unclear
1.3	The prognostic factor of interest is adequately measured in study participants, sufficient to limit potential bias	Yes	No	Unclear
1.4	The outcome of interest is adequately measured in study participants, sufficient to limit bias	Yes	No	Unclear
1.5	Important potential confounders are appropriately accounted for, limiting potential bias with respect to the prognostic factor of interest	Yes	No	Unclear
1.6	The statistical analysis is appropriate for the design of the study, limiting potential for the presentation of invalid results	Yes	No	Unclear

Notes on use of Methodology checklist: prognostic studies

The studies covered by this checklist are designed to answer questions about prognosis. Such questions address the likelihood of an outcome for patients from a population at risk for that outcome, based on the presence of a proposed prognostic factor. Prognostic factors may be disease-specific (for example, presence or absence of particular disease feature), demographic (for example, age, sex), or relate to the likely response to treatment or the presence of comorbidities.

This checklist is based on a checklist for the quality appraisal of studies about prognosis developed by Hayden and co-workers (2006).

Checklist items are worded so that a 'yes' response always indicates that the study has been designed and conducted in such a way as to minimise the risk of bias for that item. An 'unclear' response to a question may arise when the answer to an item is not reported or is not reported clearly.

1.1 The study sample represents the population of interest with regard to key characteristics, sufficient to limit potential bias to the results

Measures of prognosis can vary substantially when obtained from populations with different clinical or demographic features. Estimates of prognosis are not useful without information about the population from which they were obtained.

To minimise bias, the study population should be clearly defined and described and should represent the source population of interest. Points to consider include the following:

- Are the source population or the population of interest adequately described with respect to key characteristics?
- Are the sampling frame and recruitment adequately described, possibly including methods to identify the sample (number and type used; for example, referral patterns in healthcare), period of recruitment and place of recruitment (setting and geographical location)?
- Are inclusion and exclusion criteria adequately described (for example, including explicit diagnostic criteria or a description of participants at the start of the follow-up period)?
- Is participation in the study by eligible individuals adequate?
- Is the baseline study sample (that is, individuals entering the study) adequately described with respect to key characteristics?

1.2 Loss to follow-up is unrelated to key characteristics (that is, the study data adequately represent the sample), sufficient to limit potential bias

Attrition refers to the loss of participants during the course of a study. Consideration should be given to why participants dropped out, as well as

how many dropped out. Attrition bias occurs when there are systematic differences between participants lost to the study and those who remain.

To minimise bias, completeness of follow-up should be described and adequate. Points to consider include the following:

- Is the response rate (that is, proportion of study sample completing the study and providing outcome data) adequate?
- Are attempts to collect information on participants who dropped out of the study described?
- Are reasons for loss to follow-up provided?
- Are the key characteristics of participants lost to follow-up adequately described?
- Are there any important differences in key characteristics and outcomes between participants who completed the study and those who did not?

If your review addresses more than one outcome, you should score this item for each outcome individually.

1.3 The prognostic factor of interest is adequately measured in study participants, sufficient to limit potential bias

The prognostic factor under study should be well defined. It should be clear how the investigators determined whether participants were exposed or not to the factor. The same definition and measurement should be used for all participants in the study. Often there may be more than one way of determining the presence or absence of the factor (for example, physical or laboratory tests, questionnaire, reporting of symptoms). The method of measurement should be valid (that is, it measures what it is claimed to measure) and reliable (that is, it measures something consistently).

To minimise bias, prognostic factors should have been defined and measured appropriately. Points to consider include the following:

- Is a clear definition or description of the prognostic factor(s) measured provided (including dose, level, duration of exposure, and clear specification of the method of measurement)?
- Are continuous variables reported, or appropriate cut-off points (that is, not data-dependent) used?
- Are the prognostic factor measured and the method of measurement valid and reliable enough to limit misclassification bias? (This may include relevant outside sources of information on measurement properties, as well as characteristics such as blind measurement and limited reliance on recall.)
- Are complete data for prognostic factors available for an adequate proportion of the study sample?
- Are the method and setting of measurement the same for all study participants?
- Are appropriate methods employed if imputation is used for missing data on prognostic factors?

1.4 The outcome of interest is adequately measured in study participants, sufficient to limit potential bias

The outcome under study should be well defined. It should be clear how the investigators determined whether participants experienced, or did not experience, the outcome. The same methods for defining and measuring outcome should be used for all participants in the study. Often there may be more than one way of measuring an outcome (for example, physical or laboratory tests, questionnaire, reporting of symptoms). The method of measurement used should be valid and reliable.

To minimise bias, the outcome(s) of interest should be defined and measured appropriately. Points to consider include the following:

- Is a clear definition of the outcome of interest provided, including duration of follow-up?
- Are the outcome that was measured and the method of measurement valid and reliable enough to limit misclassification bias? (This may include relevant outside sources of information on measurement properties, as well as characteristics such as 'blind' measurement and limited reliance on recall.)
- Are the method and setting of measurement the same for all study participants?

If your review addresses more than one outcome, you should score this item for each outcome individually.

1.5 Important potential confounders are appropriately accounted for, limiting potential bias with respect to the prognostic factor of interest

Confounding can occur when there are differences between participants, apart from the presence or absence of the prognostic factor, that are related to both the outcome and the prognostic factor. An example of this is if the participants are recruited at different stages of disease progression. The design and analysis of prognostic studies are usually based on some conceptual model about how factors interact to lead to the outcome.

This question is not relevant where the study is being reviewed for the purposes of identifying the absolute risk of the outcome in the group with the prognostic factor.

To minimise bias, important confounders should be defined and measured, and confounding should be accounted for in the design or analysis. Points to consider include the following:

- Are all important confounders, including treatments (key variables in the conceptual model), measured? Are clear definitions of the important confounders measured (including dose, level and duration of exposures) provided?
- Is measurement of all important confounders valid and reliable? (This may include relevant outside sources of information on measurement

properties, as well as characteristics such as 'blind' measurement and limited reliance on recall.)

- Are the method and setting of measurement of confounders the same for all study participants?
- Are appropriate methods employed if imputation is used for missing data on confounders?
- Are important potential confounders accounted for in the study design (for example, matching for key variables, stratification or initial assembly of comparable groups)?
- Are important potential confounders accounted for in the analysis (that is, appropriate adjustment)?

If your review addresses more than one outcome, you should score this item for each outcome individually.

1.6. The statistical analysis is appropriate for the design of the study, limiting potential for the presentation of invalid results

Analysis undertaken within the study that is incorrect or inappropriate for the study design may result in false conclusions being drawn from the data.

To minimise bias, the statistical analysis undertaken should be clearly described and appropriate for the design of the study. Points to consider include the following:

- Is the presentation of data sufficient to assess the adequacy of the analysis?
- Where several prognostic factors are investigated, is the strategy for model building (that is, the inclusion of variables) appropriate and based on a conceptual framework or model?
- Is the selected model adequate for the design of the study?
- Is there any selective reporting of results?
- Are only pre-specified hypotheses investigated in the analyses?

In some circumstances it may be possible to reanalyse the data using the information supplied in the study report, in order to remove the bias.

APPENDIX F – WOMAC Questionnaire

Submit by Email | Print Form

Hôpital Maisonneuve-Rosemont
Centre affilié à l'Université de Montréal

WOMAC

HANCHE: ☐ GENOU: ☐

choisir POST OPÉRATOIRE SUIVI: choisir COTE : choisir

AUTRE :

Nom: Date:

Le questionnaire complet se rapporte à votre état de santé au cours des 4 dernières semaines.

DOULEUR

Les questions suivantes évaluent l'importance de la douleur que vous oyez au niveau de votre hanche ou de votre genou. Chaque situation se rapporte aux 4 dernières semaines.

(choisir un seul chiffre par ligne)

Quelle est 'Importance de la douleur :	Aucune	Légère	Moyenne	Sévère	Extrême
1) Lorsque vous marchez sur une surface plane?	☐	☐	☐	☐	☐
2) Lorsque vous montez ou descendez les marches?	☐	☐	☐	☐	☐
3) Lorsque vous vous levez d'une chaise ou que vous vous assoyez?	☐	☐	☐	☐	☐
4) Vous étés-vous senti(e) calme et serein(e)?	☐	☐	☐	☐	☐
5) Lorsque vous vous tenez debout?	☐	☐	☐	☐	☐

Page 1 sur 2

RAIDEUR ARTICULAIRE

Les questions suivantes évaluent la raideur articulaire (pas la douleur) que vous avez au niveau de votre hanche ou de votre genou. Chaque situation se rapporte **aux 4 dernières semaines.**
La raideur articulaire est une sensation de restriction d'un mouvement ou la diminution de la facilité bouger une articulation.

(choisir un seul chiffre par ligne)

Votre niveau de raideur articulaire	Aucune	Légère	Moyenne	Sévère	Extrême
6) Quelle est l'importance de la raideur de votre hanche ou genou lorsque vous vous levez le matin?	☐	☐	☐	☐	☐
7) Quelle est l'importance de la raideur de votre hanche ou genou après vous titre assis, couché ou repose plus tard dans la journée?	☐	☐	☐	☐	☐

CAPACITÉS PHYSIQUES

Les questions suivantes évaluent vos capacités physiques. Nous voulons évaluer votre habilité à vous déplacer et à vous occuper de vous-même. Indiquez la difficulté que vous avez à effectuer les activités suivantes à cause de votre hanche ou de votre genou. Chaque situation se rapporte **aux 4 dernières semaines.**

(choisir un seul chiffre par ligne)

Quel est la difficulté que vous avez à :	Aucune	Légère	Moyenne	Sévère	Extrême
8) Descendre les marches?	☐	☐	☐	☐	☐
9) Monter les marches?	☐	☐	☐	☐	☐
10) Vous lever de la position assise?	☐	☐	☐	☐	☐
11) Vous tenir debout?	☐	☐	☐	☐	☐
12) Vous pencher vers l'avant?	☐	☐	☐	☐	☐
13) Marcher sur un terrain plat?	☐	☐	☐	☐	☐
14) Entrer et sortir d'une automobile?	☐	☐	☐	☐	☐
15) Faire les emplettes et les commissions?	☐	☐	☐	☐	☐
16) Mettre des bas?	☐	☐	☐	☐	☐
17) Sortir du lit?	☐	☐	☐	☐	☐
18) Enlever vos bas?	☐	☐	☐	☐	☐
19) Vous étendre dans le lit?	☐	☐	☐	☐	☐
20) Entrer ou sortir du bain?	☐	☐	☐	☐	☐
21) Vous asseoir?	☐	☐	☐	☐	☐
22) Vous asseoir ou lever du siège de toilette?	☐	☐	☐	☐	☐
23) Faire le grand ménage?	☐	☐	☐	☐	☐
24) Faire le grand ménage?	☐	☐	☐	☐	☐

ÉVALUATION GLOBALE DU PATIENT

En considérant tous les aspects de votre vie que votre hanche ou genou peut affecter, comment allez-vous?

Très bien	Bien	Normal	Mal	Très mal
☐	☐	☐	☐	☐

APPENDIX G – Hip perception questionnaire

Submit by Email | Print Form

Hôpital Maisonneuve-Rosemont
Centre affilié à l'Université de Montréal

Hanche ☐ droite ☐ gauche

Suivi : choisir

Autr

Nom : _____ Date : _____

Quelle perception avez-vous de votre hanche reconstruite?

☐ 1.- Comme une hanche naturelle

☐ 2.- Hanche artificielle sans limitation fonctionnelle

☐ 3.- Hanche artificielle avec limitations fonctionnelles

☐ 4.- Hanche artificielle avec limitations fonctionnelles

☐ 5.- Articulation non fonctionnelle

Commentaires :

APPENDIX H – Research ethics board approval

Hôpital Maisonneuve-Rosemont
Centre affilié à l'Université de Montréal

Pour vous, pour la vie

Le 09 avril 2004

CHERCHEUR:	Dr. Pascal-André Vendittoli
INFIRMIÈRE :	Daniel Lusignan
PROTOCOLE:	Évaluation de vos capacités fonctionnelles et de votre qualité de vie en pré chirurgie, péri chirurgie et post chirurgie.
OBJET:	Correspondance du 19 mars 2004 au sujet du formulaire de consentement corrigé (Version 2 en date du 18 mars 2004) et du questionnaire sur la qualité de vie

RÉSOLUTION:

Sur recommandation du Sous-comité de la correspondance tenue le 31 mars 2004, les membres du Comité d'éthique de la recherche de l'Hôpital Maisonneuve-Rosemont lors de la réunion plénière tenue le 08 avril 2004, acceptent le formulaire de consentement et le questionnaire.

Jocelyne Tessier, M.D.
Présidente du comité d'éthique de la recherche
JT/dv

Pavillon Maisonneuve
5415, boul. de l'Assomption
Montréal QC H1T 2M4
Téléphone : (514) 252-3400

Pavillon Rosemont
5689, boul. Rosemont
Montréal QC H1T 2H1
Téléphone : (514) 252-3400

Pavillon Rachel-Tourigny
5305, boul. de l'Assomption
Montréal QC H1T 2M4
Téléphone : (514) 252-3400

Pavillon pédiatrique Thérèse-de-Yturralde
6900, 42e Avenue
Montréal QC H1T 2T2
Téléphone : (514) 374-7940

Centre d'accueil Judith-Jasmin
8850, rue Bisaillon
Montréal QC H1K 4N2
Téléphone : (514) 354-5990

Le 12 février 2014

Dr Pascal-André Vendittoli
Orthopédie
Hôpital Maisonneuve-Rosemont

OBJET:	Correspondance du 08 janvier 2014 au sujet de la réapprobation annuelle, du formulaire d'information et de consentement (Version 7 datée du 11 janvier 2011) et de la politique de banque de données cliniques (Version 4 datée du 17 novembre 2011)
PROTOCOLE:	Évaluation de vos capacités fonctionnelles et de votre qualité de vie en pré chirurgie, péri chirurgie et post chirurgie.
RÉFÉRENCE CÉR:	04026
INFIRMIER(IÈRE):	Monsieur Daniel Lusignan

Docteur,

Lors de la réunion du Sous-comité de correspondance du Comité d'éthique de la recherche de l'Hôpital Maisonneuve-Rosemont tenue le 12 février 2014, les membres ayant procédé à l'évaluation de la correspondance mentionnée en titre :

> ***« réapprouvent cette étude jusqu'au 11 mars 2015 et acceptent le formulaire d'information et de consentement ainsi que la politique de banque. »***

Pour toute correspondance avec le CÉR, je vous demanderais de bien vouloir mentionner le numéro attribué par notre institution à votre projet de recherche (Réf. CÉR). Pour plus amples informations, je vous saurais gré de communiquer avec Madame Vaillancourt, assistante administrative au CÉR.

Vous remerciant de votre engagement dans la quête de l'excellence en recherche dans le respect de la dignité des sujets, je vous prie d'agréer, Docteur, l'expression de mes salutations les plus distinguées.

Lynne Senécal, M.D.
Présidente
Comité d'éthique de la recherche

5415, boul. de l'Assomption
Montréal QC H1T 2M4
Téléphone : (514) 252-3400
www.maisonneuve-rosemont.org

APPENDIX I – Information and consent form

Hôpital Maisonneuve-Rosemont
Centre affilié à l'Université de Montréal

Pour vous, pour la vie

Formulaire d'Information et de Consentement

Titre de l'étude :

Évaluation de vos Capacités Fonctionnelles et de Votre Qualité de Vie en Pré Chirurgie, Péri Chirurgie et Post Chirurgie.

Chercheur Principal : Dr Pascal André Vendittoli

Co chercheurs : Dr David Blanchette
Dr Michel Fallaha
Dr Marc Isler
Dr Martin Lavigne
Dr Patrick Lavigne
Dr Alain Roy
Dr Sophie Mottard
Dr Vincent Massé

Organisme subventionnaire :

Fondation de recherche et d'éducation en orthopédie de Montréal.

Préambule

Nous sollicitons votre participation à un projet de recherche. Cependant, avant d'accepter de participer à ce projet et de signer ce formulaire d'information et de consentement, veuillez prendre le temps de lire, de comprendre et de considérer attentivement les renseignements qui suivent.

Ce formulaire peut contenir des mots que vous ne comprenez pas. Nous vous invitons à poser toutes les questions que vous jugerez utiles au chercheur responsable du projet ou aux autres membres du personnel affecté au projet de recherche et à leur demander de vous expliquer tout mot ou renseignement qui n'est pas clair.

But et Objectif de l'étude :

Vous êtes atteint d'une maladie articulaire qui modifie grandement vos habitudes de vie et vos capacités fonctionnelles. Vous serez éventuellement amené à avoir une chirurgie correctrice qui procédera à un autre changement d'habitudes de vie.

Le but de cette étude est d'évaluer ces changements par le biais de questionnaires. Ces questionnaires mesureront les effets de la chirurgie et donneront des indications pour votre réadaptation.

Ces questionnaires sont validés scientifiquement et utilisés à travers le monde dans ce type de recherche. Ces questionnaires sont le « Womac, le SF-36, le Harris score, le Knee score, le Merle d'Aubigné et des questionnaires construits par le département d'arthroplastie de l'HMR.

Nous allons aussi utiliser d'autres questionnaires qui sont validés universellement et qui sont devenus des standards dans l'évaluation de l'acte médical en orthopédie. Ces questionnaires multiples sont disponibles sur internet et cela pour toutes personnes intéressées.

Toutes les données recueillis iront dans notre banque de données qui se nomme comme suit:

Politique d'encadrement de la banque de données pour la recherche orthopédique

L'objectif de la banque vise à la constitution d'une réserve de données cliniques sur des sujets malades et en santés, ayant consentis à participer pour les fins de la recherche à des questionnaires orthopédiques validés et standardisés, ainsi que des questionnaires créés par l'équipe d'orthopédie. Les informations orthopédiques issues du dossier médical et les données radiologiques seront aussi cumulées dans cette banque de données.

Toutes les sujets opérés pour la hanche et le genou peuvent contribuer à cette banque en ayant consenti par un formulaire d'information et de consentement (comme celui-ci) spécifique à chaque projet de recherche. Cette banque est optionnelle, vous n'aurez qu'à initialer pour y participer et cocher oui à la dernière page du consentement.

Notre banque de données contiendra aussi l'information de cette recherche. Donc, le participant a le droit à tout moment de révoquer son consentement de participation à la banque et de demander le retrait de ses données qu'il a confiées à des fins de recherche scientifique. Ce retrait n'entraînera ni désavantage ni pénalité pour la personne concernée.

De plus pour notre banque de données, les informations recueillies seront traités comme suit :

- Tous les renseignements sont conservés sous forme confidentielle et un numéro de code est attribué à chaque patient. Le gestionnaire de la banque détient le registre des participants dans un ordinateur portable sécurisé par des mots de passe. Il attribue un code et il est le seul à pouvoir faire le lien avec le participant.
- Les données de la banque sont la propriété de l'Hôpital Maisonneuve Rosemont, elles sont conservées dans un local sécurisé pendant 25 ans à HMR, au 5415 boul. l'Assomption, pavillon Rachel Tourigny, porte 4152.

Aucune donnée ne sera détruite sans autorisation à moins que le participant n'en fasse la demande. Une confirmation écrite de la destruction des données lui sera envoyée.

Cette banque de données est soumise à une politique de banque de données qui est accepté par le Comité d'éthique et de la recherche de HMR.

Votre participation à cette recherche se limite à accepter de compléter ces questionnaires qui prendront une quinzaine de minutes de votre temps.

Cette étude sera proposée aux patients de l'HMR, de l'hôpital Santa Cabrini et de l'hôpital Jean-Talon, qui sont sous les soins des chirurgiens orthopédistes reliés à cette étude et qui opèrent dans ces hôpitaux. Cela veut dire que vous les autoriser à consulter votre dossier médical.

Risques et Inconvénients :

Mis à part le temps pris pour compléter ces questionnaires. Il n'y a aucun risque associé à la participation à cette recherche. La participation à cette étude n'a aucune influence sur le traitement que vous recevrez. Le seul risque est le potentiel de bris de confidentialité.

Avantages :

Le seul avantage relié avec cette étude est de contribuer à l'avancement des connaissances dans le domaine des soins orthopédiques post-opératoires.

Compensation Financière :

Aucune compensation financière n'est prévue pour la participation à cette étude.

Indemnisation :

En acceptant de participer à cette étude et en signant le présent formulaire, vous ne renoncez à aucun de vos droits garantis par la loi ni ne libérez les chercheurs, le personnel ou les institutions impliquées dans cette recherche de leurs responsabilités légales et professionnelles.

Participation volontaire et possibilité de retrait

La participation à cette étude est volontaire, c'est-à-dire que vous pouvez décider librement de participer à l'étude ou non. En tout temps, avant ou après votre opération, vous pouvez décider de vous retirer de l'étude. Votre décision de ne pas participer à ce projet de recherche ou de vous en retirer n'aura aucune conséquence sur la qualité des soins et des services auxquels vous avez droit ou sur votre relation avec le chercheur responsable du projet et les autres intervenants.

Communications des résultats :

En tout temps vous pourrez être informé des résultats vous concernant, vous devrez en faire la demande à votre chirurgien ou à la personne responsable de l'étude.

Votre nom ne sera transmis comme donnée dans la banque de données, il fait partie d'un répertoire à part qui est gardé par le responsable de la banque. L'utilisation de la banque est autorisée que par le responsable de la banque. Aucune donnée permettant de vous identifier ne sera transmise.

Confidentialité :

Durant votre participation à ce projet, le chercheur responsable ainsi que son personnel recueilleront et consigneront dans un dossier de recherche les renseignements vous concernant. Seuls les renseignements nécessaires pour répondre aux objectifs scientifiques de ce projet seront recueillis.

Ces renseignements peuvent comprendre les informations contenues dans vos dossiers médicaux concernant votre état de santé passé et présent, vos habitudes de vie ainsi que les

résultats de tous les tests, examens et procédures que vous aurez à subir durant ce projet ainsi que les réponses aux questionnaires. Votre dossier de recherche peut aussi comprendre d'autres renseignements tels que votre nom, votre sexe, votre date de naissance et votre origine ethnique.

Tous les renseignements recueillis demeureront strictement confidentiels dans les limites prévues par la loi. Afin de préserver votre identité et la confidentialité des renseignements, vous ne serez identifié que par un numéro de code. La clé du code reliant votre nom à votre dossier de recherche sera conservé par le chercheur responsable.

Le chercheur responsable et son personnel feront parvenir au commanditaire ou à ses représentants, les données vous concernant. Ces données n'incluent pas votre nom ni votre adresse.

Le commanditaire utilisera les données à des fins de recherche dans le but de répondre aux objectifs scientifiques du projet décrits dans le formulaire d'information et de consentement.

Les données pourront être publiées dans des revues spécialisées ou faire l'objet de discussions scientifiques, mais il ne sera pas possible de vous identifier.

À des fins de surveillance et de contrôle, votre dossier de recherche ainsi que vos dossiers médicaux, s'il y a lieu, pourront être consultés par une personne mandatée par le comité d'éthique de la recherche de l'HMR, par les établissements concernés ou par l'organisme subventionnaire. Toutes ces personnes et organismes adhèrent à une politique de confidentialité.
À des fins de protection, notamment afin de pouvoir communiquer avec vous rapidement, vos noms et prénoms, vos coordonnées et la date de début et de fin de votre participation au projet seront conservés dans un répertoire à part maintenu par le chercheur responsable.

Vous avez le droit de consulter votre dossier de recherche pour vérifier les renseignements recueillis, et les faire rectifier au besoin, et ce, aussi longtemps que le chercheur responsable du projet ou l'établissement détiennent ces informations. Cependant, afin de préserver l'intégrité scientifique du projet, vous pourriez n'avoir accès à certaines de ces informations qu'une fois votre participation terminée.

Financement du projet de recherche

Le chercheur responsable du projet a reçu un financement de l'organisme subventionnaire pour mener à bien ce projet de recherche.

Personnes-ressources

Si vous avez des questions concernant le projet de recherche ou la participation à cette banque, vous pouvez communiquer avec le chercheur responsable du projet de recherche, Dr PA Vendittoli, au 514-252-3400 poste 4558 ou Daniel Lusignan, infirmier et coordonnateur recherche au 514-252-3400, poste 5701.

Pour toute question concernant vos droits en tant que sujet participant à ce projet de recherche ou si vous avez des plaintes ou des commentaires à formuler vous pouvez

communiquer avec le commissaire local aux plaintes et à la qualité des services de l'HMR u 514-252-3400 poste 3510.

Informations administratives

Ce projet de recherche et ce formulaire d'information et de consentement ont été approuvés par le CÉR de l'HMR, il en assure le suivi. Si vous avez des questions sur vos droits à titre de sujet participant à une étude de recherche, vous pouvez joindre la coordonnatrice du 514-252-3400 poste 5708.

Si vous avez une plainte à formuler, vous pouvez contacter le Commissaire local aux plaintes et à la qualité des services de l'Hôpital Maisonneuve Rosemont au numéro (514) 252-3400, poste 3510.

Consentement

Évaluation de vos capacités fonctionnelles et de votre qualité de vie
en pré chirurgie, péri chirurgie et post chirurgie.

Consentement du sujet

J'ai pris connaissance du formulaire d'information et de consentement. Je reconnais qu'on m'a expliqué le projet, qu'on a répondu à mes questions et qu'on m'a laissé le temps voulu pour prendre une décision.

Je consens à participer à ce projet de recherche aux conditions qui y sont énoncées. Une copie signée et datée du présent formulaire d'information et de consentement m'a été remise. Une autre copie sera versée à mon dossier médical. En conséquence, je comprends que cette information sera disponible à toute personne ou compagnie à laquelle je donnerai accès à mon dossier médical.

_____

Nom et signature du sujet de recherche Date

J'autorise la collecte des données mentionnées ci haut pour une participation à la banque de données suivante : ☐ oui ____ (initiales Patient)

Politique d'encadrement de la banque de données
pour la recherche orthopédique

Personne qui a obtenu le consentement

J'ai expliqué au sujet de recherche les termes du présent formulaire d'information et de consentement et j'ai répondu aux questions qu'il m'a posées.

_____

Nom et signature de la personne qui obtient le consentement Date

Engagement du chercheur

Je certifie qu'on a expliqué au sujet de recherche les termes du présent formulaire d'information et de consentement, que l'on a répondu aux questions que le sujet de recherche avait à cet égard et qu'on lui a clairement indiqué qu'il demeure libre de mettre un terme à sa participation, et ce, sans préjudice. Je m'engage, avec l'équipe de recherche, à respecter ce qui a été convenu au formulaire d'information et de consentement et à en remettre une copie signée au sujet de recherche.

_____

Nom et signature du chercheur responsable du projet de recherche Date

APPENDIX J – Proof of manuscript acceptance

From: **Revue de Chirurgie Orthopedique** rco@elsevier.com
Subject: OTSR-RCO - Votre article No. OTSR-RCO-D-15-00542R1 est accepte pour publication
Date: December 3, 2015 at 11:35
To:

RD

ELSEVIER EDITORIAL SYSTEM
Orthopaedics & Traumatology: Surgery & Research
Ref.: Ms. No. OTSR-RCO-D-15-00542R1
Titre : A systematic review of preoperative determinants of patient-reported pain and physical function up to two years following primary total hip arthroplasty

Dear Colleague,

We are pleased to inform you that your manuscript has been accepted by the reviewers of OTSR and will be published shortly.

You will soon receive the proofs.

Thank you very much for your interest in our journal.

With our best regards.

The Editorial Board
Orthopaedics & Traumatology: Surgery & Research

APPENDIX K – STROBE guidelines for the prediction algorithm

	Item No	Recommendation	Pages
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	63
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	63
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	67
Objectives	3	State specific objectives, including any prespecified hypotheses	68
Methods			
Study design	4	Present key elements of study design early in the paper	68
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	68
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	68
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	69-70
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	69-70
Bias	9	Describe any efforts to address potential sources of bias	71
Study size	10	Explain how the study size was arrived at	69-70
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	70-71
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	70-71
		(*b*) Describe any methods used to examine subgroups and interactions	N/A
		(*c*) Explain how missing data were addressed	N/A
		(*d*) If applicable, explain how loss to follow-up was addressed	N/A
		(*e*) Describe any sensitivity analyses	73 – Internal validation
Results			

Participants	13	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	71-72
		(b) Give reasons for non-participation at each stage	71
		(c) Consider use of a flow diagram	N/A
Descriptive data	14	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	78
		(b) Indicate number of participants with missing data for each variable of interest	78
		(c) Summarise follow-up time (eg, average and total amount)	72
Outcome data	15	Report numbers of outcome events or summary measures over time	72-73
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	73
		(*b*) Report category boundaries when continuous variables were categorized	73
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	N/A
Discussion			
Key results	18	Summarise key results with reference to study objectives	73-74
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	76
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	77
Generalisability	21	Discuss the generalisability (external validity) of the study results	76
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	77

APPENDIX L – Additional prediction algorithms that were considered

Figure 1 Graphical representation of prediction algorithm (PA) 1

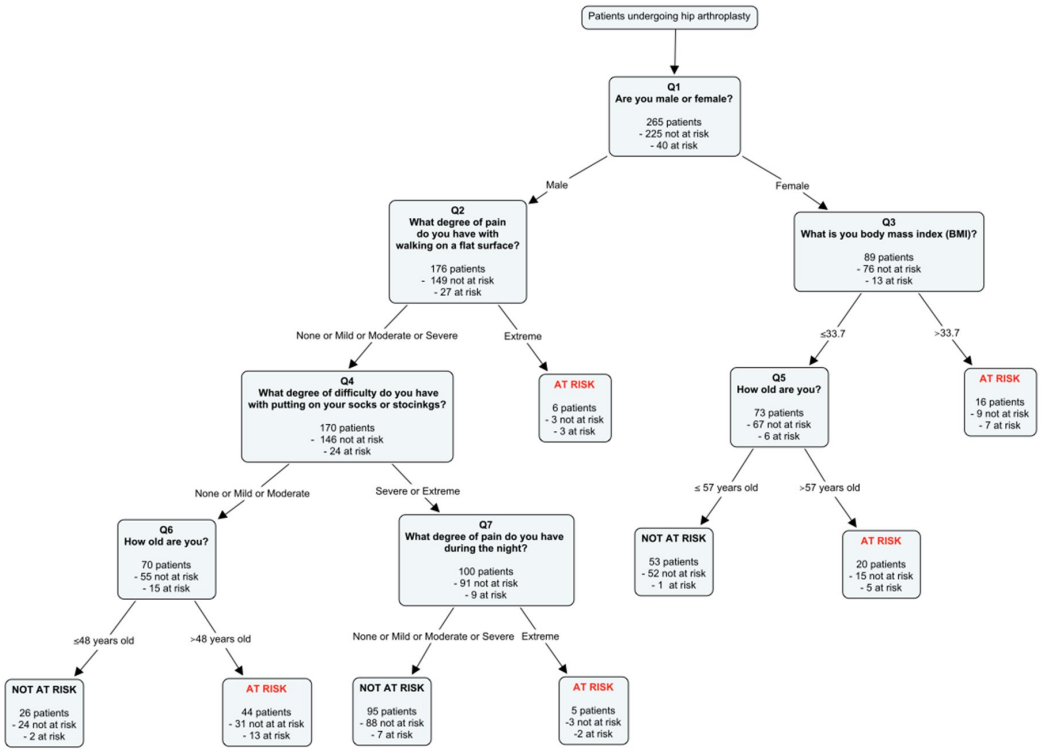


Table 1.1 Two by Two table of predicted versus actual outcomes of the PA1

	Actual Outcome	
	AT RISK	**NOT AT RISK**
Predicted outcome	Worst postoperative WOMAC quartile (>11.5/100) & "Artificial with minimal or major limitations" joint perception	Postoperative WOMAC ≤ 11.5/100 or ''Artificial with no limitations'' or ''Natural joint'' joint perception
AT RISK	30	61
NOT AT RISK	10	164
TOTAL	*40*	*225*

Table 1.2 Validity measures of the PA1

Measure	**Estimates in training sample**
Sensitivity % (95% CI)	75.0 (59.8-85.8)
Specificity % (95% CI)	72.9 (66.7-78.3)
Positive predictive value % (95% CI)	33.0 (24.2-43.1)
Negative predictive value % (95% CI)	94.3 (89.7-96.8)
Positive likelihood ratio (95% CI)	2.77 (2.09-3.66)
Negative likelihood ratio (95% CI)	0.34 (0.20-0.59)

Figure 2 Graphical representation of PA2

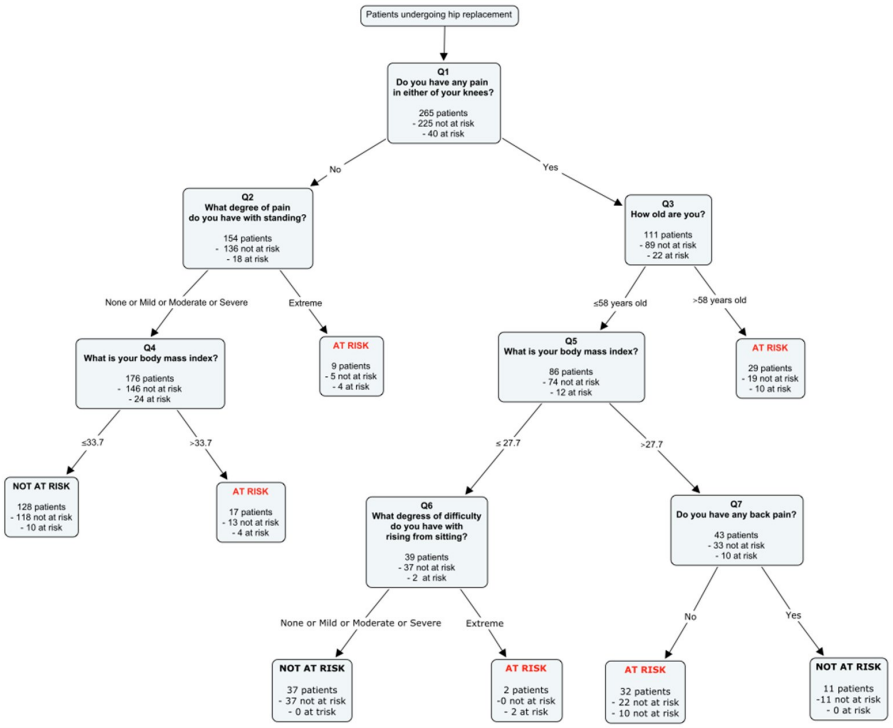


Table 2.1 Two by Two table of predicted versus actual outcomes of the PA2

	Actual Outcome	
	AT RISK	**NOT AT RISK**
Predicted outcome	Worst postoperative WOMAC quartile (>11.5/100) & "Artificial with minimal or major limitations" joint perception	Postoperative WOMAC ≤ 11.5/100 or ''Artificial with no limitations'' or ''Natural joint'' joint perception
AT RISK	30	59
NOT AT RISK	10	166
TOTAL	*40*	*225*

Table 1.2 Validity measures of the PA2

Measure	**Estimates in training sample**
Sensitivity % (95% CI)	75.0 (59.8-85.8)
Specificity % (95% CI)	73.8 (67.7-79.1)
Positive predictive value % (95% CI)	33.7 (24.7-44.0)
Negative predictive value % (95% CI)	94.3 (89.9-96.9)
Positive likelihood ratio (95% CI)	2.86 (2.16-3.80)
Negative likelihood ratio (95% CI)	0.34 (0.20-0.58)

Figure 3 Graphical representation of PA3

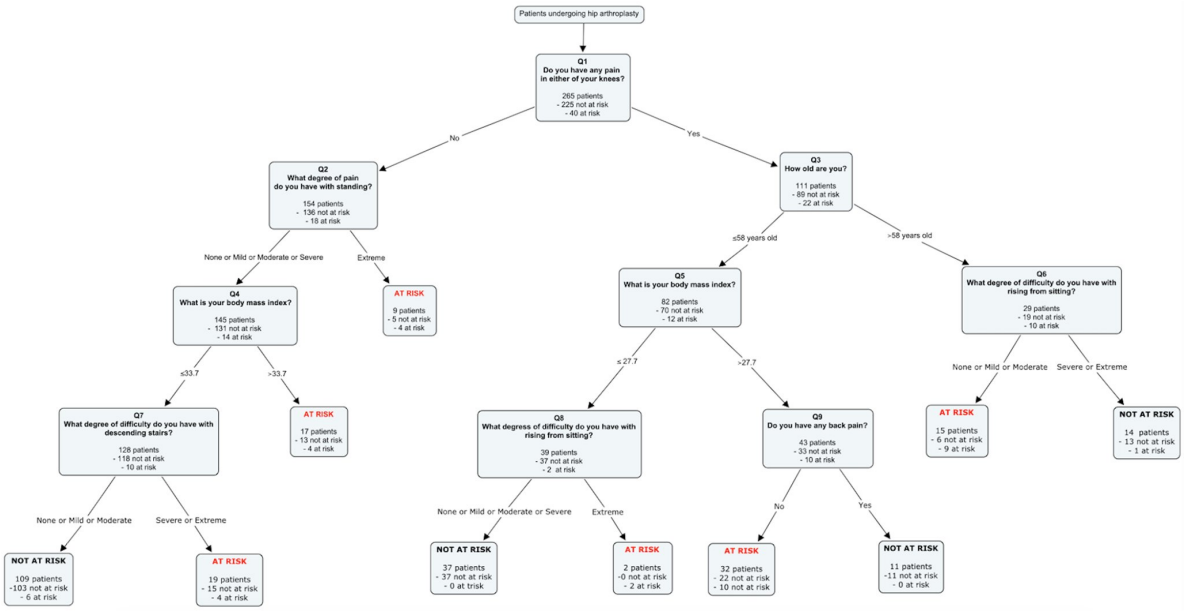


Table 3.1 Two by Two table of predicted versus actual outcomes of the PA3

	Actual Outcome	
	AT RISK	**NOT AT RISK**
Predicted outcome	Worst postoperative WOMAC quartile (>11.5/100) & "Artificial with minimal or major limitations" joint perception	Postoperative WOMAC ≤ 11.5/100 or ''Artificial with no limitations'' or ''Natural joint'' joint perception
AT RISK	33	61
NOT AT RISK	7	164
TOTAL	*40*	*225*

Table 3.2 Validity measures of the PA3

Measure	**Estimates in training sample**
Sensitivity % (95% CI)	82.5 (68.1-91.3)
Specificity % (95% CI)	72.9 (66.7-78.3)
Positive predictive value % (95% CI)	35.1 (26.2-45.2)
Negative predictive value % (95% CI)	95.9 (91.8-98.0)
Positive likelihood ratio (95% CI)	3.04 (2.35-3.94)
Negative likelihood ratio (95% CI)	0.24 (0.12-0.47)

Figure 4 Graphical representation of PA4

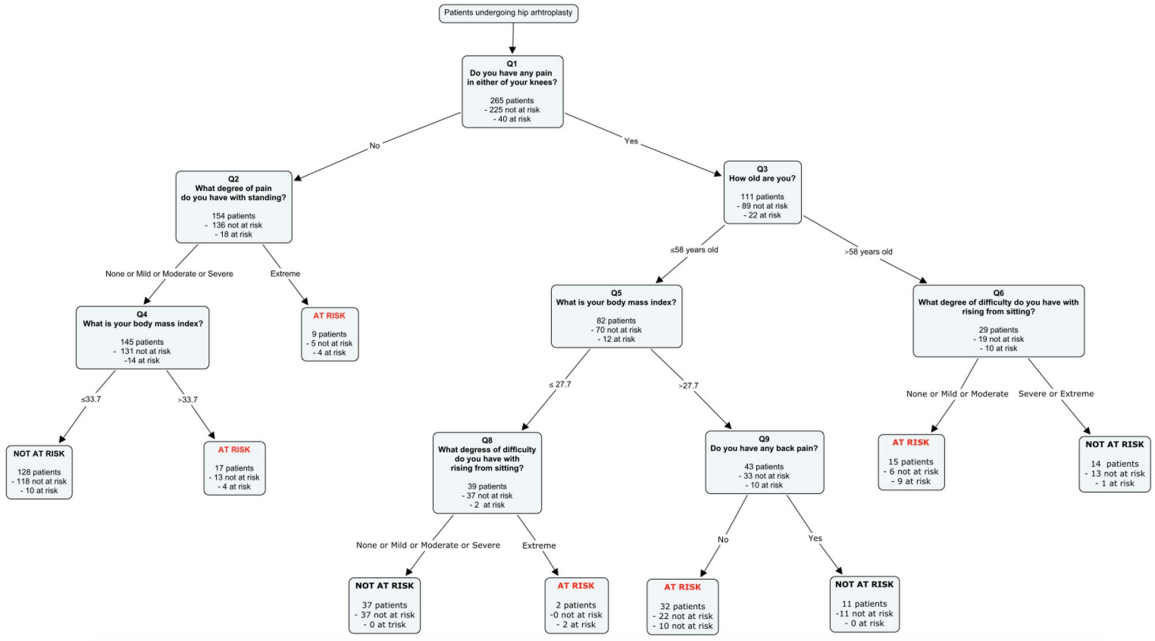


Table 4.1 Two by Two table of predicted versus actual outcomes of the PA4

	Actual Outcome	
	AT RISK	**NOT AT RISK**
Predicted outcome	Worst postoperative WOMAC quartile (>11.5/100) & "Artificial with minimal or major limitations" joint perception	Postoperative WOMAC ≤ 11.5/100 or ''Artificial with no limitations'' or ''Natural joint'' joint perception
AT RISK	29	46
NOT AT RISK	11	179
TOTAL	*40*	*225*

Table 4.2 Validity measures of the PA4

Measure	**Estimates in training sample**
Sensitivity % (95% CI)	72.5 (57.2-83.9)
Specificity % (95% CI)	79.6 (73.8-84.3)
Positive predictive value % (95% CI)	38.7 (28.5-50.0)
Negative predictive value % (95% CI)	94.2 (89.9-96.7)
Positive likelihood ratio (95% CI)	3.55 (2.57-4.89)
Negative likelihood ratio (95% CI)	0.35 (0.21-0.57)

Figure 5 Graphical representation of PA5

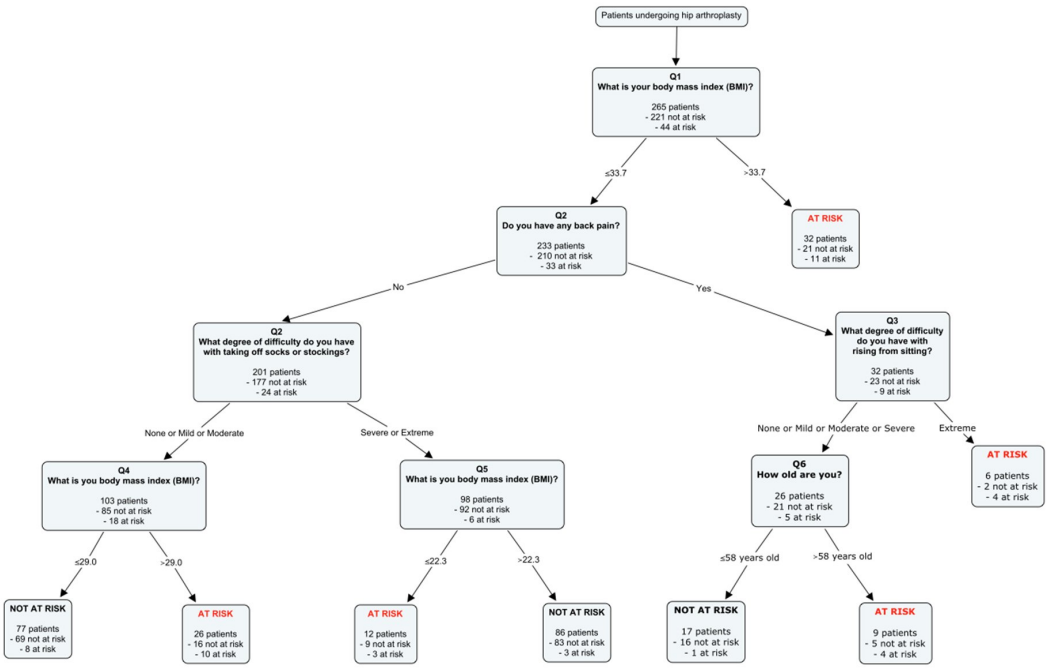


Table 5.1 Two by Two table of predicted versus actual outcomes of the PA5

	Actual Outcome	
	AT RISK	**NOT AT RISK**
Predicted outcome	Worst postoperative WOMAC tertile (>9.4/100) & "Artificial with minimal or major limitations" joint perception	Postoperative WOMAC ≤ 9.4/100 or ''Artificial with no limitations'' or ''Natural joint'' joint perception
AT RISK	32	53
NOT AT RISK	12	168
TOTAL	*44*	*221*

Table 5.2 Validity measures of the PA5

Measure	**Estimates in training sample**
Sensitivity % (95% CI)	72.7 (58.2-83.7)
Specificity % (95% CI)	76.0 (70.0-81.2)
Positive predictive value % (95% CI)	37.6 (28.1-48.3)
Negative predictive value % (95% CI)	93.3 (88.7-96.1)
Positive likelihood ratio (95% CI)	3.03 (2.26-4.08)
Negative likelihood ratio (95% CI)	0.36 (0.22-0.59)

Figure 6 Graphical representation of PA6

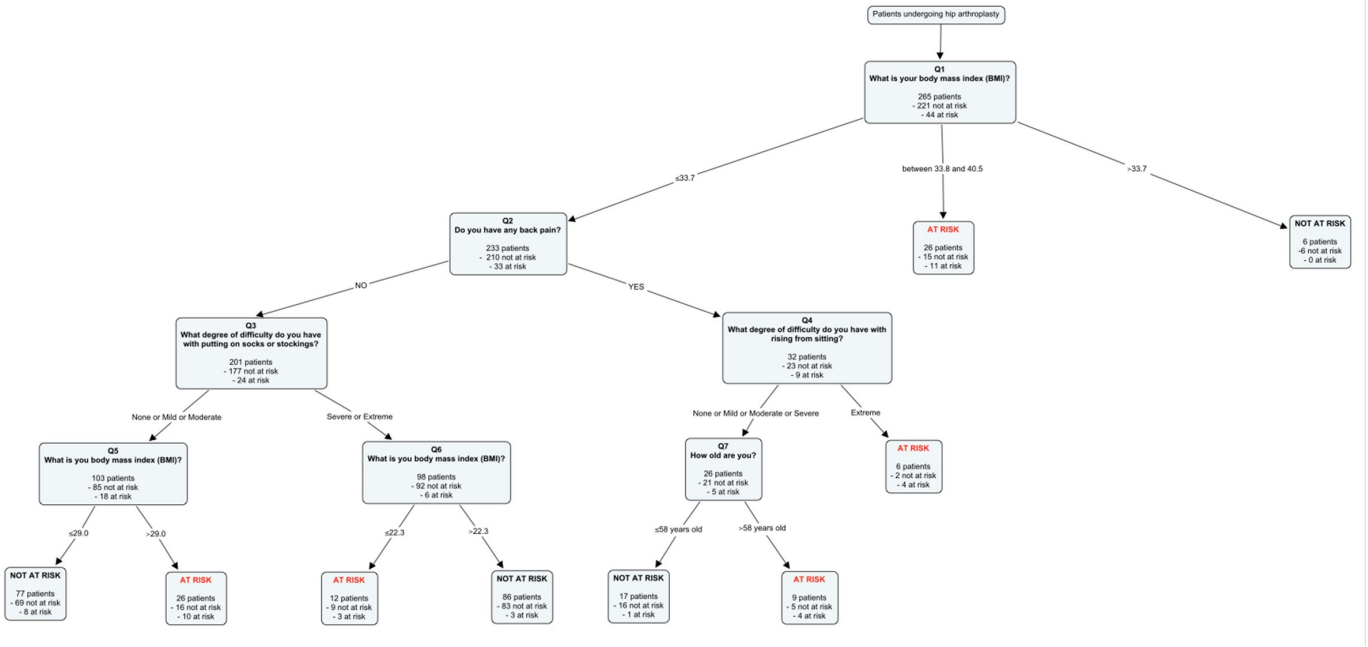


Table 6.1 Two by Two table of predicted versus actual outcomes of the PA6

	Actual Outcome	
	AT RISK	**NOT AT RISK**
Predicted outcome	Worst postoperative WOMAC tertile (>9.4) & "Artificial with minimal or major limitations" joint perception	Postoperative WOMAC ≤ 9.4/100 or ''Artificial with no limitations'' or ''Natural joint'' joint perception
AT RISK	32	47
NOT AT RISK	12	174
TOTAL	*44*	*221*

Table 6.2 Validity measures of the PA6

Measure	**Estimates in training sample**
Sensitivity % (95% CI)	72.7 (58.2-83.7)
Specificity % (95% CI)	78.7 (72.9-83.6)
Positive predictive value % (95% CI)	40.5 (30.4-51.5)
Negative predictive value % (95% CI)	93.5 (89.1-96.3)
Positive likelihood ratio (95% CI)	3.42 (2.50-4.67)
Negative likelihood ratio (95% CI)	0.35 (0.21-0.56)

APPENDIX M – STROBE guidelines for the clinical prediction rule

	Item No	Recommendation	Pages
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	125
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	125
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	129
Objectives	3	State specific objectives, including any prespecified hypotheses	130
Methods			
Study design	4	Present key elements of study design early in the paper	130
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	130
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	131
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	131-132
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	132
Bias	9	Describe any efforts to address potential sources of bias	133
Study size	10	Explain how the study size was arrived at	Reference [121]
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	133
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	133
		(*b*) Describe any methods used to examine subgroups and interactions	N/A
		(*c*) Explain how missing data were addressed	133
		(*d*) If applicable, explain how loss to follow-up was addressed	140
		(*e*) Describe any sensitivity analyses	136 – Internal validation
Results			

Participants	13	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	134-135
		(b) Give reasons for non-participation at each stage	134-135
		(c) Consider use of a flow diagram	143
Descriptive data	14	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	146
		(b) Indicate number of participants with missing data for each variable of interest	146
		(c) Summarise follow-up time (eg, average and total amount)	135
Outcome data	15	Report numbers of outcome events or summary measures over time	136
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	147
		(*b*) Report category boundaries when continuous variables were categorized	147
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	N/A
Discussion			
Key results	18	Summarise key results with reference to study objectives	136-137
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	139-140
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	140
Generalisability	21	Discuss the generalisability (external validity) of the study results	140
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	142

APPENDIX N – Additional clinical prediction rules that were considered

Figure 1 Graphical representation of Prediction Rule (PR) 1

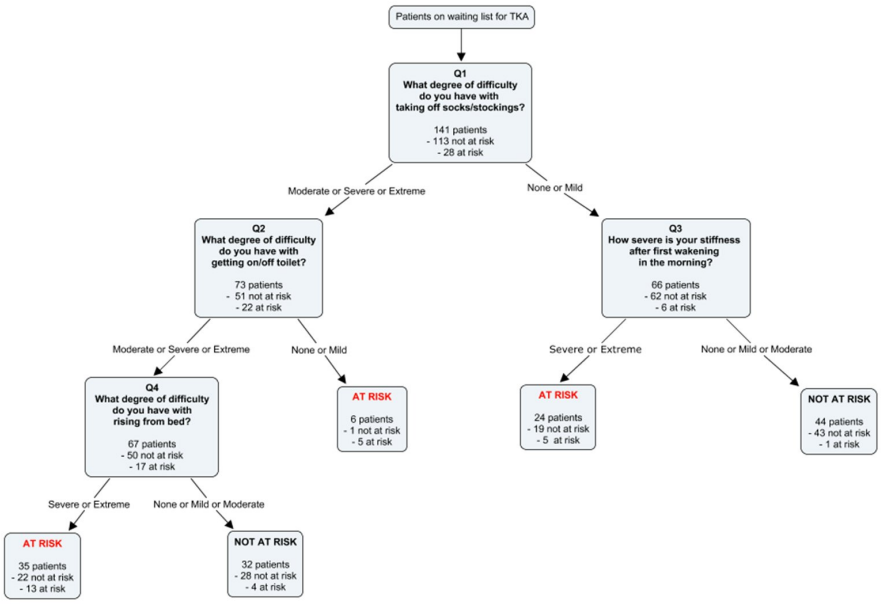


Table 1.1 Two by Two table of predicted versus actual outcomes of the PR1

	Actual Outcome	
Predicted outcome	**AT RISK (Postoperative WOMAC > 40.4)**	**NOT AT RISK (Postoperative WOMAC ≤ 40.4)**
AT RISK	23	42
NOT AT RISK	5	71
TOTAL	*28*	*113*

Table 1.2 Validity measures of the PR1

Measure	**Estimates in training sample**
Sensitivity % (95% CI)	82.1 (64.4-92.1)
Specificity % (95% CI)	62.8 (53.6-71.2)
Positive predictive value % (95% CI)	35.4 (24.9-47.5)
Negative predictive value % (95% CI)	93.4 (85.5-97.2)
Positive likelihood ratio (95% CI)	2.21 (1.65-2.97)
Negative likelihood ratio (95% CI)	0.28 (0.13-0.63)

Figure 2 Graphical representation of PR2

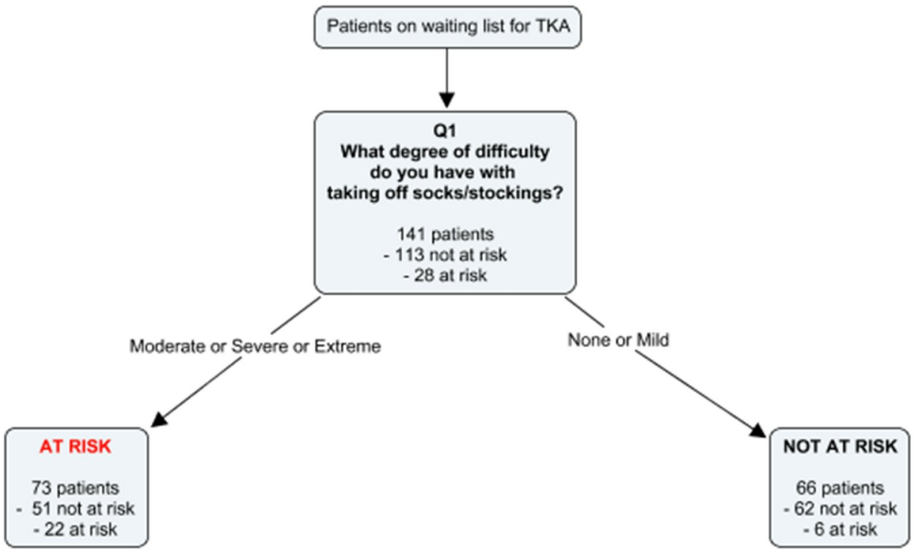


Table 2.1 Two by Two table of predicted versus actual outcomes of the PR2

	Actual Outcome	
Predicted outcome	**AT RISK (Postoperative WOMAC > 40.4)**	**NOT AT RISK (Postoperative WOMAC ≤ 40.4)**
AT RISK	22	51
NOT AT RISK	6	62
TOTAL	*28*	*113*

Table 2.2 Validity measures of the PR2

Measure	**Estimates in training sample**
Sensitivity % (95% CI)	78.6 (60.5-89.8)
Specificity % (95% CI)	54.9 (45.7-63.7)
Positive predictive value % (95% CI)	30.1 (20.8-41.4)
Negative predictive value % (95% CI)	91.2 (82.1-95.9)
Positive likelihood ratio (95% CI)	1.74 (1.32-2.31)
Negative likelihood ratio (95% CI)	0.39 (0.19-0.81)

Figure 3 Graphical representation of PR3

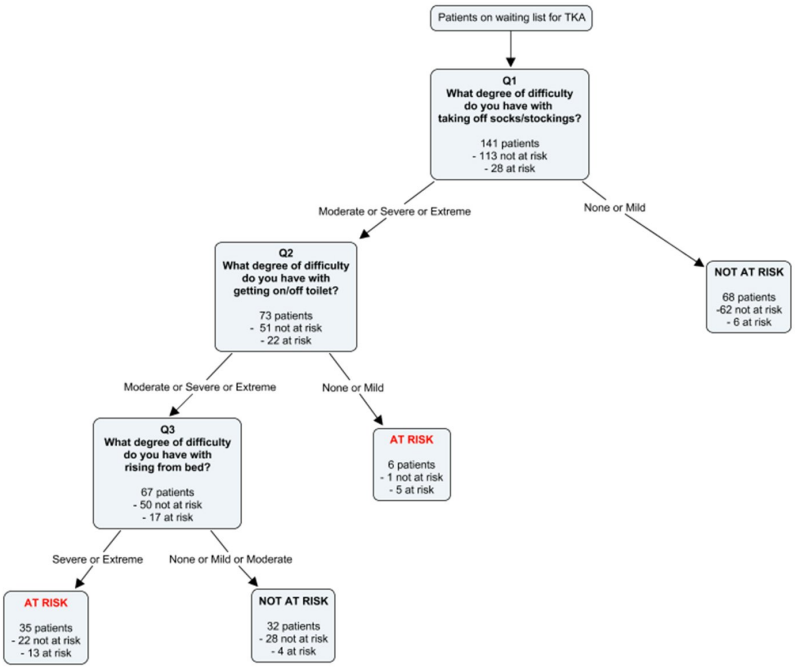


Table 3.1 Two by Two table of predicted versus actual outcomes of the PR3

	Actual Outcome	
Predicted outcome	**AT RISK (Postoperative WOMAC > 40.4)**	**NOT AT RISK (Postoperative WOMAC ≤ 40.4)**
AT RISK	18	23
NOT AT RISK	10	90
TOTAL	*28*	*113*

Table 3.2 Validity measures of the PR3

Measure	**Estimates in training sample**
Sensitivity % (95% CI)	64.3 (45.8-79.3)
Specificity % (95% CI)	79.6 (71.3-86.0)
Positive predictive value % (95% CI)	43.9 (29.9-59.0)
Negative predictive value % (95% CI)	90.0 (82.6-94.5)
Positive likelihood ratio (95% CI)	3.16 (2.00-4.99)
Negative likelihood ratio (95% CI)	0.39 (0.19-0.81)

Figure 4 Graphical representation of PR4

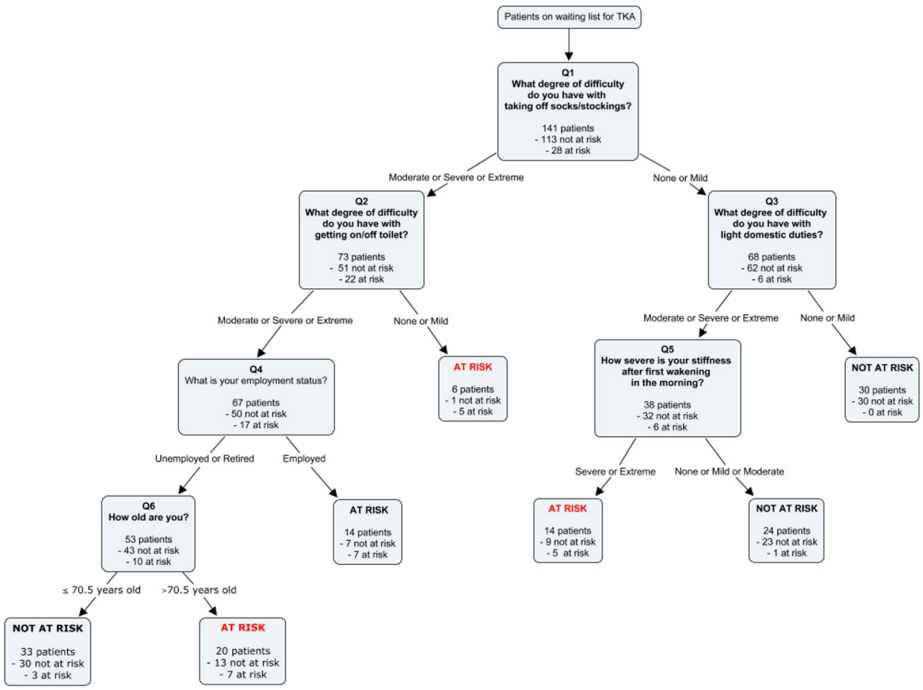


Table 4.1 Two by Two table of predicted versus actual outcomes of the PR4

	Actual Outcome	
Predicted outcome	**AT RISK (Postoperative WOMAC > 40.4)**	**NOT AT RISK (Postoperative WOMAC ≤ 40.4)**
AT RISK	24	30
NOT AT RISK	4	83
TOTAL	*28*	*113*

Table 4.2 Validity measures of the PR4

Measure	**Estimates in training sample**
Sensitivity % (95% CI)	85.7 (68.5-94.3)
Specificity % (95% CI)	73.5 (64.6-80.7)
Positive predictive value % (95% CI)	44.4 (32.0-57.6)
Negative predictive value % (95% CI)	95.4 (88.8-98.2)
Positive likelihood ratio (95% CI)	3.23 (2.29-4.55)
Negative likelihood ratio (95% CI)	0.19 (0.08-0.49)

Figure 5 Graphical representation of PR5

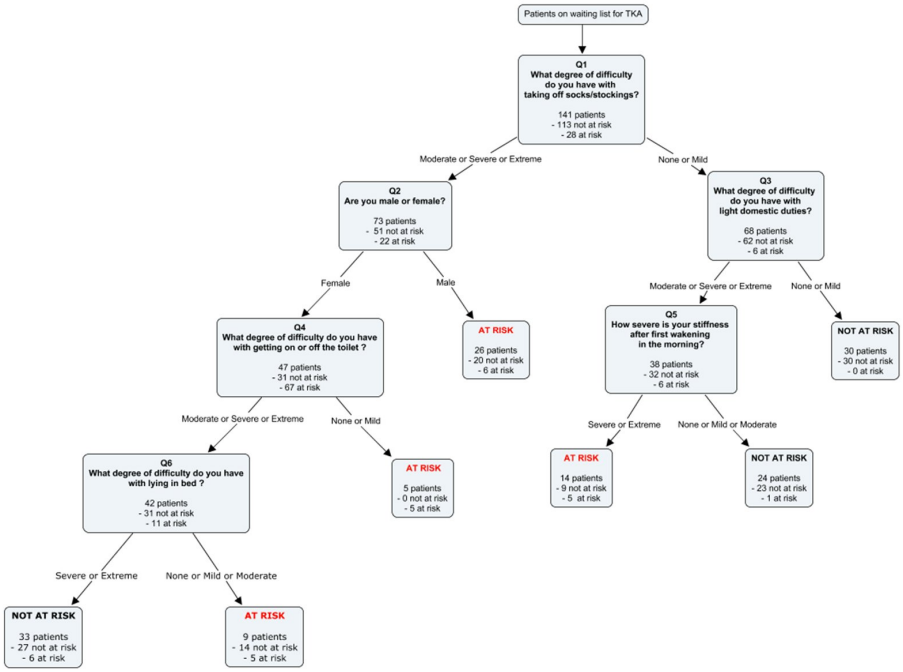


Table 5.1 Two by Two table of predicted versus actual outcomes of the PR5

	Actual Outcome	
Predicted outcome	**AT RISK (Postoperative WOMAC > 40.4)**	**NOT AT RISK (Postoperative WOMAC ≤ 40.4)**
AT RISK	21	33
NOT AT RISK	7	80
TOTAL	*28*	*113*

Table 5.2 Validity measures of the PR5

Measure	**Estimates in training sample**
Sensitivity % (95% CI)	75.0 (56.6-87.3)
Specificity % (95% CI)	70.8 (61.8-78.4)
Positive predictive value % (95% CI)	38.9 (27.0-52.2)
Negative predictive value % (95% CI)	92.0 (84.3-96.0)
Positive likelihood ratio (95% CI)	2.57 (1.80-3.67)
Negative likelihood ratio (95% CI)	0.35 (0.18-0.68)

Figure 6 Graphical representation of PR6

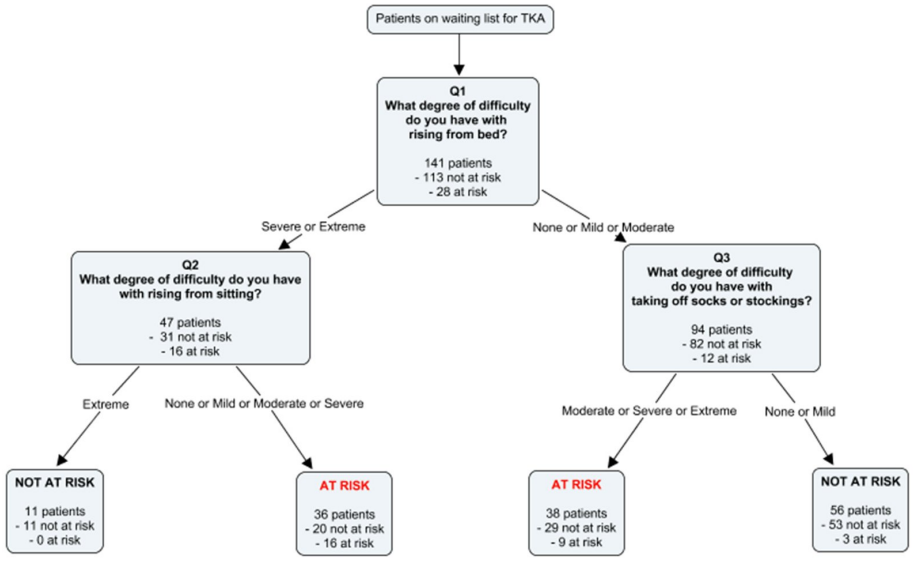


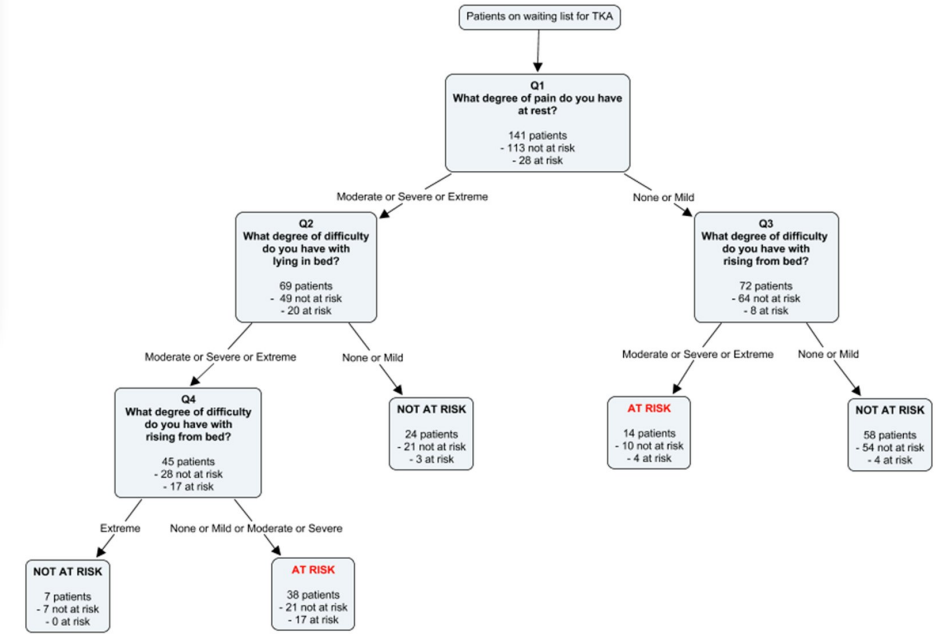
Table 6.1 Two by Two table of predicted versus actual outcomes of the PR6

	Actual Outcome	
Predicted outcome	**AT RISK (Postoperative WOMAC > 40.4)**	**NOT AT RISK (Postoperative WOMAC ≤ 40.4)**
AT RISK	25	40
NOT AT RISK	3	73
TOTAL	*28*	*113*

Table 6.2 Validity measures of the PR6

Measure	**Estimates in training sample**
Sensitivity % (95% CI)	89.3 (72.8-96.3)
Specificity % (95% CI)	64.6 (55.4-72.8)
Positive predictive value % (95% CI)	38.5 (27.6-50.6)
Negative predictive value % (95% CI)	96.1 (89.0-98.6)
Positive likelihood ratio (95% CI)	2.52 (1.91-3.34)
Negative likelihood ratio (95% CI)	0.17 (0.06-0.49)

Figure 7 Graphical representation of PR7



	Actual Outcome	
Predicted outcome	**AT RISK (Postoperative WOMAC > 40.4)**	**NOT AT RISK (Postoperative WOMAC ≤ 40.4)**
AT RISK	21	31
NOT AT RISK	7	82
TOTAL	*28*	*113*

Table 7.1 Two by Two table of predicted versus actual outcomes of the PR7

Table 7.2 Validity measures of the PR7

Measure	**Estimates in training sample**
Sensitivity % (95% CI)	89.3 (72.8-96.3)
Specificity % (95% CI)	64.6 (55.4-72.8)
Positive predictive value % (95% CI)	38.5 (27.6-50.6)
Negative predictive value % (95% CI)	96.1 (89.0-98.6)
Positive likelihood ratio (95% CI)	2.52 (1.91-3.34)
Negative likelihood ratio (95% CI)	0.17 (0.06-0.49)

Figure 8 Graphical representation of PR8

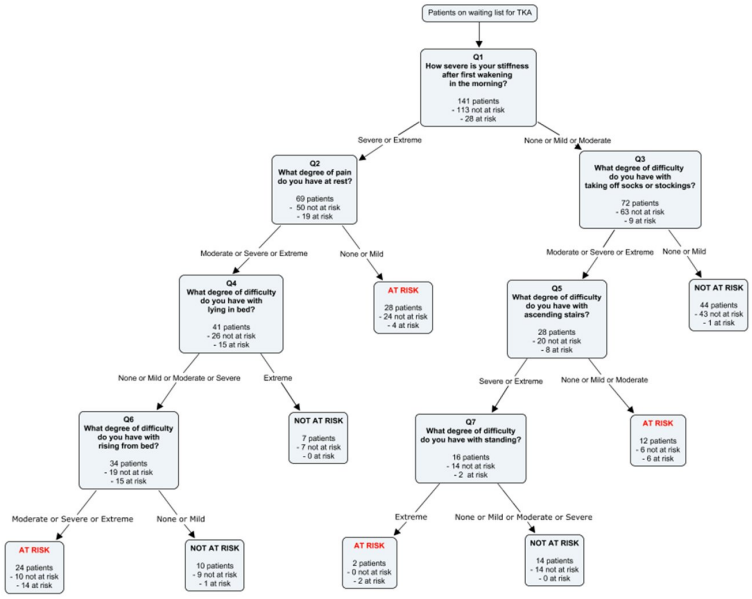


Table 8.1 Two by Two table of predicted versus actual outcomes of the PR8

	Actual Outcome	
Predicted outcome	**AT RISK (Postoperative WOMAC > 40.4)**	**NOT AT RISK (Postoperative WOMAC ≤ 40.4)**
AT RISK	26	40
NOT AT RISK	2	73
TOTAL	*28*	*113*

Table 8.2 Validity measures of the PR8

Measure	**Estimates in training sample**
Sensitivity % (95% CI)	92.9 (77.4-98.0)
Specificity % (95% CI)	64.6 (55.4-72.8)
Positive predictive value % (95% CI)	39.4 (28.5-51.5)
Negative predictive value % (95% CI)	97.3 (90.8-99.3)
Positive likelihood ratio (95% CI)	2.62 (2.00-3.43)
Negative likelihood ratio (95% CI)	0.11 (0.03-0.42)