Systematic review

Predictive Models for Clinical Outcomes in Total Knee Arthroplasty: A Systematic Analysis

Cécile Batailler, MD [a, b, *], Timothy Lording, FRACS [c], Daniele De Massari, PhD [d],
Sietske Witvoet-Braam [d], Stefano Bini, MD, PhD [e], Sébastien Lustig, MD, PhD [a, b]

[a] Orthopedic Surgery Department, Croix-Rousse Hospital, Lyon, France
[b] IFSTTAR, LBMC UMR_T9406, Université Claude Bernard Lyon 1, Villeurbanne, France
[c] Orthopedic surgery department, Melbourne Orthopaedic Group, Windsor, Australia
[d] Stryker Department, Amsterdam, The Netherlands
[e] Orthopedic surgery department, University of California, San Francisco, USA

ARTICLE INFO

ABSTRACT

Background: Predictive modeling promises to improve our understanding of what variables influence patient satisfaction after total knee arthroplasty (TKA). The purpose of this article was to systematically review the relevant literature using predictive models of clinical outcomes after TKA. The aim was to identify the predictor strategies used for systematic data collection with the highest likelihood of success in predicting clinical outcomes.
Methods: A Preferred Reporting Items for Systematic Reviews and Meta-Analyses protocol systematic review was conducted using 3 databases (MEDLINE, EMBASE, and PubMed) to identify all clinical studies that had used predictive models or that assessed predictive features for outcomes after TKA between 1996 and 2020. The ROBINS-I tool was used to evaluate the quality of the studies and the risk of bias.
Results: A total of 75 studies were identified of which 48 met our inclusion criteria. Preoperative predictive factors strongly associated with postoperative clinical outcomes were knee pain, knee-specific Patient-Reported Outcome Measure (PROM) scores, and mental health scores. Demographic characteristics, pre-existing comorbidities, and knee alignment had an inconsistent association with outcomes. The outcome measures that correlated best with the predictive models were improvement of PROM scores, pain scores, and patient satisfaction.
Conclusions: Several algorithms, based on PROM improvement, patient satisfaction, or pain after TKA, have been developed to improve decision-making regarding both indications for surgery and surgical strategy. Functional features such as preoperative pain and PROM scores were highly predictive for clinical outcomes after TKA. Some variables such as demographics data or knee alignment were less strongly correlated with TKA outcomes.
Level of evidence: Systematic review − Level III.

Introduction

Total knee arthroplasty (TKA) is an efficient surgical treatment for knee osteoarthritis. However, patient dissatisfaction and sub-optimal patient reported outcomes are reported to be as high as to 20% [1,2]. With the rise of robotic surgery, a time may come when the procedure itself will no longer be considered a feature that significantly determines outcomes. In such a scenario, understanding how other groups of variables such as patient-specific attributes, functional measures, socio-economic indicators, or perioperative recovery location influence clinical outcomes will become increasingly important. Conceivably, using relevant data points incorporated into an algorithm, the insights derived for any given patient could impact surgical indications, procedure type, venue of surgery, and even recovery site.

Predictive models are usually deployed in contexts where the measurement of the output is difficult, time-demanding, and expensive [3]. The increasing availability of large digital health-care

data sets has facilitated the application of predictive models. Several studies have published predictive models for TKA outcomes that have taken into account several features such as functional scores, preoperative pain [4], comorbidities [5], demographic characteristics, and psychological features [6–8]. The goal was to use these probabilistic models to estimate and predict the likelihood of improvements in function and satisfaction after TKA with the goal of supporting surgeon decision-making [9]. Ever more complex algorithmic approaches have been developed; however, none of these have so far been able to replicate standard surgeon intuition [10,11].

Currently, to our knowledge, no study summarizes which features have been identified as the most predictive of clinical outcomes or which algorithms have been most successfully used in predictive analytics after TKA. The purpose of this article is therefore to systematically review the relevant literature on predictive factors and predictive models for outcome after TKA. This article will describe the preoperative predictive features which have been identified as having the strongest correlation with outcomes and patient satisfaction after TKA. Second, it will review the machine learning models of TKA results.

Material and methods

Article identification and selection process

A query in December 2020 was performed to identify all available literature that described or used predictive models for outcomes after TKA. The search was performed through PubMed, EMBASE, and MEDLINE data bases from 1996 to 2020 inclusive using the 2009 Preferred Reporting Items for Systematic Reviews and Meta-Analyses protocol (PRISMA).

Inclusion criteria for the search strategy included all English language studies reporting information regarding the use of predictive models or the identification of preoperative predictive factors for outcomes after TKA. The following terms were used: "total knee arthroplasty" or "total knee replacement"; "predictive factor" or "predictive model" or "predictive modeling" or "predictive feature" or "predict"; and "outcomes", "satisfaction", "pain" or "PROMs" or "dissatisfaction". Exclusion criteria consisted of (1) editorial articles, (2) systematic reviews or meta-analyses, (3) articles on revision TKA, and (4) articles evaluating joints other than the knee. The abstracts from all identified articles were independently reviewed by 2 investigators.

Quality assessment

The Risk Of Bias In Non-Randomised Studies of Interventions (ROBINS-I) tool [12] was used to evaluate the quality of the included studies and their relative risk of bias. The categories for risk of bias judgements are "Low risk", "Moderate risk", "Serious risk" and "Critical risk". To increase the reliability of this classification, the same observer evaluated all articles with the ROBINS-I tool 2 times separated by an interval of 4 weeks. If the assessment of the study quality was not the same during these 2 evaluations, a second observer evaluated the concerned article with the ROBINS-I tool.

Results

Included articles and study characteristics

The PRISMA flow diagram for study selection is shown in Figure 1. Of the 75 potential articles, 19 were excluded as not relevant, and 8 were excluded because of their scoring a "critical risk of bias" score leaving 48 studies for inclusion. The risk of bias for these studies is reported in Table 1.

Predictive factors

Several parameters were consistently identified in different studies as impacting outcomes after TKA. These parameters have been classified into 3 groups according to the strength of their association with outcome (Supplementary Table 1): (1) strong and consistent association; (2) strong but inconsistent association; (3) weak and inconsistent association (Table 2). The predictive factors classified in group I (strong association) were significantly correlated with outcomes after TKA ($P < .05$) in all studies with low or moderate risk, which assessed these factors. The predictive factors classified in group II (strong but inconsistent association) were significantly correlated with TKA outcomes ($P < .05$) in low- or moderate-risk studies, but not in all. For this group, relevant studies found a significant correlation, but other relevant studies did not find the same strong association. The predictive factors classified in group III (weak association) were not significantly correlated with TKA outcomes in the low- or moderate-risk studies.

Strong and clear association

Preoperative pain. Patients with higher levels of pain before TKA surgery had lower postoperative functional scores. However, improvements in pain scores, functional knee scores, and patient satisfaction were greater in this group [13–16]. Huijbregts et al. found that a one-point increase in preoperative NRS-pain (Numerical Rating Scale) resulted in a 1.73-point decrease in 1-year Oxford Knee Score (OKS) [17].

Preoperative PROM score. Preoperative PROM scores, particularly knee scores, were strongly correlated with TKA outcome [5,14,15,18–21]. Maratt et al. has defined the minimally clinically important difference (MCIDs) for the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) scores at 2 years after TKA [22]. The preoperative WOMAC scores were the strongest predictive factors for improvement in postoperative WOMAC scores in a cohort of 2350 TKAs.

Mental health. Anxiety and depression before surgery are frequently identified as risk factors for lower patient-reported outcomes after TKA [16], in particular with regard to patient dissatisfaction [6,14,15,20,23,24], knee pain [7], and walking limitations [5,25–27].

Inconsistent association

Demographic characteristics. Huijbregts et al. reported an inverse correlation between age and satisfaction with knee surgery [17,28,29].

With respect to sex, several studies have suggested that residual pain and stiffness [30], and consequently dissatisfaction, were more prevalent in female patients [15,25].

Body mass index (BMI) is a statistically significant predictor of satisfaction (Knee Society Score [KSS] satisfaction subscale) [14,31], postoperative PROM scores [32–34], and postoperative range of motion [30]. Dowsey et al. described poorer functional outcomes (Knee and function KSS) in morbidly obese patients (BMI > 40 kg/m^2) [35]. Nevertheless, this correlation is not consistently demonstrated by all authors [29,36–38].

Some studies had grouped several demographic parameters (age, female gender, and BMI) to create a single demographic
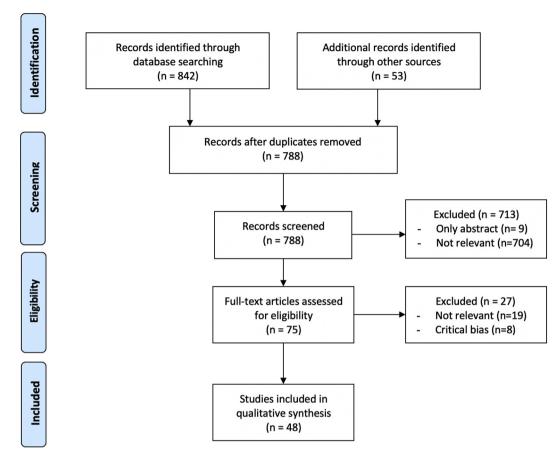
Figure 1. Flow chart from initial literature search through to data extraction from final list of included studies.

criterion [13,34]. A predictive model for postoperative PROM scores found a significant correlation between age and gender and OKS [33]. For example, younger women (age < 60 years) had better OKS outcomes than men, but in the oldest age group (age 80 years or older), women had worse outcomes than men [33].

Clinical comorbidities. Several studies identified clinical comorbidities as significant predictive factors of poor outcomes after TKA [16]. Diabetes [20,39] or allergies [40,41] were singled out.

Osteoarthritis severity. On 996 TKAs, Schnurr et al. reported that the severity of preoperative osteoarthritis was the only feature which inversely correlated with patient satisfaction [36]. In comparison to severe arthritis grade IV, the risk for dissatisfaction was 2.6-fold higher for arthritis grade III ($P < .001$) and 3.0-fold higher for grade II ($P = .001$).

Surgical history. In some studies, the number of previous knee surgeries was correlated to postoperative outcomes [39] or to pain during the hospitalization ($P < .002$) [13].

Preoperative knee alignment. Sueyoshi et al. described a significant association between preoperative varus greater than 5° and postoperative pain ($P = .0096$) [37]. However, Twiggs et al. found no correlation between preoperative knee alignment and pain scores at 12 months [42].

Preoperative range of motion. Similarly, preoperative range of motion (ROM) has an uncertain impact on postoperative clinical

outcomes after TKA despite having a direct correlation with postoperative ROM [25,30,43].

Low association
Some features were not identified as significant predictors of clinical outcomes, such as education, socioeconomic status, smoking, and alcohol habits, and patient expectations were not found to be predictive of outcomes or pain after TKA [13,14,42,44,45]. Preoperative pain medication use was not a significant predictive factor of postoperative satisfaction after TKA [36]. Surgical time and tourniquet time were not clearly identified as independent predictors of postoperative pain or dissatisfaction after TKA [13,29]. These features were not found to be predictive of postoperative outcomes in the literature reviewed and are unlikely to contribute to a predictive model.

Patient-reported outcome measures

Twenty-three different outcome measurement parameters were found in the included studies. These parameters could be grouped as follows: (1) postoperative validated measures, (2) patient satisfaction measures, (3) pain measures, and (4) improvement in PROMs. Several predictive models have been developed to estimate various measures of postoperative outcome after TKA (Table 3).

Postoperative PROMs (specific of the knee or general)
Some predictive models have focused on postoperative knee PROMs. Sanchez-Santos et al. described a predictive model for the postoperative OKS questionnaire at 12 months after TKA using data

Table 1
Summary of quality assessment of the studies included in our analysis, according to ROBINS-I tool (Risk Of Bias In Non-Randomised Studies of Interventions) [12].

Study	Confounding	Selection of patients	Classification of interventions	Deviations from intended interventions	Missing data	Measurement of outcomes	Selection of reported results	Study risk
Brander et al. (2003) [7]	Moderate	Low	Low	Low	Low	Moderate	Moderate	Moderate
Lingard et al. (2004) [56]	Moderate	Low	Moderate	Low	Moderate	Moderate	Moderate	Moderate
Bourne et al. (2007) [28]	Moderate	Moderate	Low	Low	Low	Moderate	Low	Moderate
Escobar et al. (2007) [10]	Moderate	Low	Low	Low	Low	Moderate	Low	Moderate
Davis et al. (2008) [44]	Moderate	Low	Low	Low	Moderate	Moderate	Moderate	Moderate
Franklin et al. (2008) [32]	Low	Low	Low	Low	Low	Low	Low	Low
Nilsdotter et al. (2009) [45]	Serious	Moderate	Moderate	Low	Moderate	Serious	Moderate	Serious
Rajgopal et al. (2008) [38]	Moderate	Low	Low	Low	Moderate	Low	Moderate	Moderate
Dowsey et al. (2010) [35]	Moderate	Low	Low	Low	Low	Moderate	Low	Moderate
Bourne et al. (2010) [1]	Moderate	Moderate	Moderate	Low	Low	Serious	Moderate	Serious
Blackburn et al. (2012) [27]	Moderate	Moderate	Moderate	Low	Moderate	Moderate	Low	Moderate
Judge et al. (2012) [6]	Moderate	Low	Moderate	Low	Serious	Serious	Moderate	Serious
Baker et al. (2012) [5]	Moderate	Low	Low	Low	Moderate	Moderate	Low	Moderate
Schnurr et al. (2013) [36]	Moderate	Moderate	Low	Low	Moderate	Moderate	Moderate	Moderate
Barlow et al. (2014) [57]	Serious	Moderate	Moderate	Low	Serious	Serious	Moderate	Serious
Lungu et al. (2014) [46]	Moderate	Moderate	Low	Low	Serious	Serious	Moderate	Serious
Sueyoshi et al. (2015) [37]	Moderate	Low	Moderate	Low	Moderate	Moderate	Low	Moderate
Huijbregts et al. (2016) [17]	Moderate	Low	Low	Low	Moderate	Low	Moderate	Moderate
Maratt et al. (2015) [22]	Moderate	Moderate	Moderate	Low	Moderate	Moderate	Moderate	Moderate
Feldmann et al. (2015) [58]	Serious	Moderate	Low	Low	Moderate	Serious	Low	Serious
Maempel et al. (2016) [30]	Low	Low	Low	Low	Low	Low	Low	Low
Van Onsem et al. (2016) [15]	Moderate	Moderate	Low	Low	Moderate	Moderate	Low	Moderate
Giurea et al. (2016) [24]	Moderate	Low	Moderate	Low	Moderate	Moderate	Moderate	Moderate
Hinarejos et al. (2016) [40]	Moderate	Low	Moderate	Low	Low	Moderate	Moderate	Moderate
Kremers et al. (2017) [59]	Low	Moderate	Low	Moderate	Serious	Moderate	Serious	Serious
Jain et al. (2017) [60]	Serious	Serious	Moderate	Low	Moderate	Serious	Moderate	Serious
Sanchez Santos et al. (2018) [33]	Low	Low	Low	Low	Low	Low	Low	Low
Clement et al. (2018) [20]	Moderate	Low	Low	Low	Moderate	Moderate	Low	Moderate
Van Onsem et al. (2018) [43]	Serious	Moderate	Moderate	Low	Serious	Serious	Moderate	Serious
Clement et al. (2018) [23]	Moderate	Low	Moderate	Low	Moderate	Moderate	Moderate	Moderate
Abrecht et al. (2019) [13]	Moderate	Moderate	Moderate	Low	Low	Moderate	Low	Moderate
Calkins et al. (2019) [31]	Low	Serious	Moderate	Low	Low	Serious	Low	Serious
Twiggs et al. (2019) [42]	Low	Low	Low	Low	Low	Low	Low	Low
Tolk et al. (2020) [18]	Moderate	Low	Moderate	Low	Moderate	Moderate	Moderate	Moderate
Zabawa et al. (2019) [14]	Moderate	Moderate	Moderate	Low	Low	Moderate	Low	Moderate
Kunze et al. (2019) [39]	Low	Low	Moderate	Low	Low	Moderate	Low	Moderate
Clement et al. (2019) [19]	Moderate	Low	Low	Low	Moderate	Moderate	Moderate	Moderate
Ramkumar et al. (2019) [61]	Low	Low	Low	Low	Low	Low	Low	Low
Pua et al. (2019) [25]	Moderate	Low	Moderate	Low	Low	Moderate	Moderate	Moderate
Xu et al. (2020) [62]	Moderate	Moderate	Low	Low	Moderate	Moderate	Low	Moderate
Vissers et al. (2020) [63]	Moderate	Low	Moderate	Low	Low	Moderate	Moderate	Moderate
Kunze et al. (2020) [41]	Moderate	Moderate	Low	Low	Moderate	Low	Low	Moderate
Belford et al. (2020) [64]	Moderate	Moderate	Moderate	Low	Moderate	Moderate	Moderate	Moderate
Pua et al. (2020) [26]	Moderate	Low	Low	Low	Low	Moderate	Moderate	Moderate
Farooq et al. (2020) [29]	Moderate	Low	Low	Low	Moderate	Moderate	Low	Moderate
Harris et al. (2021) [21]	Moderate	Low	Low	Low	Moderate	Low	Moderate	Moderate
Itou et al. (2020) [65]	Serious	Serious	Moderate	Low	Moderate	Serious	Moderate	Serious
Anis et al. (2020) [34]	Low	Low	Low	Low	Low	Low	Low	Low

The categories for risk of bias judgements are "Low risk", "Moderate risk", "Serious risk", and "Critical risk". The worst judgment bias assigned within any one domain gives the judgment score of the complete study.

Table 2

Table reporting the different predictive factors and the strength of their correlation with TKA outcomes, for each included study.

Parameters	Brander (2003) [7]	Lingard (2004) [56]	Bourne (2007) [28]	Escobar (2007) [10]	Davis (2008) [44]	Franklin (2008) [32]	Nilsdotter (2008) [45]	Rajgopal (2008) [38]	Dowsey (2010) [35]	Bourne (2010) [1]	Blackbum (2012) [27]	Judge (2012) [6]	Baker (2012) [5]	Schnurr (2013) [36]	Barlow (2014) [57]	Lungu (2014) [46]	Sueyoshi (2015) [37]	Huijbregts (2015) [17]	Maratt (2015) [22]	Feldmann (2015) [58]	Maempel (2016) [30]	Van Onsem (2016) [15]
Quality assessment	Mod.	Mod.	Mod.	Mod.	Mod.	Low	Serious	Mod.	Mod.	Serious	Mod.	Serious	Mod.	Mod.	Serious	Serious	Mod.	Mod.	Mod.	Serious	Low	Mod.
Preoperative predictive factors																						
Preop VAS Pain	SA																	SA				SA
Preop pain medication																						
Neurological disease/backpain				SA																		
Preop KOOS score												SA										SA
Preop KSS score																						
Preop WOMAC	SA	SA						SA		IA					SA			SA	SA			SA
Preop SF-12 PCS/SF-36		SA		IA		SA								NA				SA				IA
Preop SF-12 MCS		SA				SA		SA														
Preop OKS Score												SA	SA		NA			SA			SA	
Preop EQ5D VAS													SA				SA					
Preop ROM																IA					SA	
Joint comorbidity/previous knee surgery												SA			NA		SA					
Severity osteoarthritis (Kellgren)														SA								
Preop knee alignment																IA	SA					IA
Quadriceps strength					SA																	
Depression/Anxiety	SA										SA	SA		SA	NA							SA
Ability to cope																						SA
Allergy (>1 self-reported)																						
Medical comorbidities/ASA score		IA												SA				NA	NA			
BMI	NA	IA	IA			SA	IA	IA	IA	NA	NA				NA		NA	NA	NA		SA	
Gender	NA	IA	NA	IA		NA	IA	NA		NA	IA						NA	NA	NA		SA	IA
Age	NA	SA	SA	IA		SA	IA	NA		SA	IA		SA		NA	IA	SA	NA	IA		SA	IA
Geography (UK vs US/AUS)		SA																				
Income					IA																	
Decreased social support																	NA					
Education/Socioeconomic status (SES)				IA	NA															SA		
Smoking/Drinking																						
Employment status										NA												

Table 2 (continued)

Parameters	Brander (2003) [7]	Lingard (2004) [56]	Bourne (2007) [28]	Escobar (2007) [10]	Davis (2008) [44]	Franklin (2008) [32]	Nilsdotter (2008) [45]	Rajgopal (2008) [38]	Dowsey (2010) [35]	Bourne (2010) [1]	Blackbum (2012) [27]	Judge (2012) [6]	Baker (2012) [5]	Schnurr (2013) [36]	Barlow (2014) [57]	Lungu (2014) [46]	Sueyoshi (2015) [37]	Huijbregts (2015) [17]	Maratt (2015) [22]	Feldmann (2015) [58]	Maempel (2016) [30]	Van Onsem (2016) [15]
Quality assessment	Mod.	Mod.	Mod.	Mod.	Mod.	Low	Serious	Mod.	Mod.	Serious	Mod.	Serious	Mod.	Mod.	Serious	Serious	Mod.	Mod.	Mod.	Serious	Low	Mod.
Expectation							NA			SA									NA			
Ethnicity																			NA			
Patient-reported outcome measures																						
Pain Catastrophizing Scale (PCS)																				X		X
VAS pain	X																					
KSCR improvement		X	X																			
WOMAC Score	X			X	X			X		X					X	X			X	X		
WOMAC improvement			X					X		X						X			X			
SF12 PCS score						X			X													
SF12 MCS score									X													
SF12 PCS improvement		X	X			X			X													
SF-36 score				X					X						X							
KSS score						X											X					X
KOOS score							X												X			X
KOOS Improvement																						
IKS score									X										X			
IKS improvement									X													
OKS score											X	X	X		X			X				X
OKS improvement													X									
EQ-5D score													X									X
EQ-5D improvement													X									
Satisfaction																X		X				X
Post op ROM																					X	
Revision risk																						

Parameters	Giurea (2016) [24]	Hinarejos (2016) [40]	Kremers (2017) [59]	Jain (2017) [60]	Sanchez Santos (2018) [33]	Clement (2018) [20]	Van Onsem (2018) [43]	Clement (2018) [23]	Abrecht (2019) [13]	Calkins (2019) [31]	Twiggs (2019) [42]	Tolk (2019) [18]	Zabawa (2019) [14]	Kunze (2019) [39]	Clement (2019) [19]	Ramkumar (2019) [61]	Pua (2019) [26]	Xu (2019) [62]	Vissers (2020) [63]	Kunze (2020) [41]	Belford (2020) [64]
Quality assessment	Mod.	Mod.	Serious	Serious	Low	Mod.	Serious	Mod.	Mod.	Serious	Low	Mod.	Mod.	Mod.	Mod.	Low	Mod.	Mod.	Mod.	Mod.	Mod.
Preoperative predictive factors																					
Preop VAS Pain		NA							SA				SA								SA
Preop pain medication										NA											
Neurological disease/backpain						SA		IA			SA				SA						
Preop KOOS score				IA			SA		NA	SA	SA										
Preop KSS score		NA					SA						SA	SA						SA	
Preop WOMAC score						SA									SA						SA
						SA			NA						SA						

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20
Preop SF-12 PCS/SF-36																				
Preop SF-12 MCS					SA				NA					SA			SA			
Preop OKS score				SA		SA					SA							SA		SA
Preop EQ5D VAS											SA									
Preop ROM				SA		SA							SA			SA				
Joint comorbidity/previous knee surgery				SA				IA		NA			SA							
Severity osteoarthritis (Kellgren)													NA					SA		
Preop knee alignment										NA			NA							
Quadriceps strength						SA														
Depression/anxiety	SA			SA			IA	IA				SA		IA			SA			SA
Ability to cope	SA			SA																
Allergy (>1 self-reported)		SA											SA						SA	
Medical comorbidities/ASA score				SA	SA		NA					SA		NA	SA	IA	NA	SA		SA
BMI	NA		NA	SA	NA	NA		SA	SA	SA	IA	SA	SA	NA		IA		NA		
Gender	NA		NA	SA	SA	NA		SA	NA	NA	SA	SA		NA		IA		NA		
Age	NA		NA	IA	IA	NA	NA	SA	NA	NA	IA			NA		IA		NA		SA
Geography (UK vs US/AUS)																				
Income				SA								NA								
Decreased social support	NA									NA										
Education/socioeconomic status (SES)													NA			NA				
Smoking/drinking										NA				SA						
Employment status										NA										
Expectation			IA																	
Ethnicity																				
Patient-reported outcome measures																				
Pain Catastrophizing Scale (PCS)																				
VAS Pain								X								X				
KSCR improvement																				
WOMAC Score		X		X									X			X			X	
WOMAC improvement				X																
SF12 PCS score			X	X																
SF12 MCS score			X	X																
SF12 PCS improvement				X																
SF-36 score																				
KSS score		X		X					X		X	X							X	
KOOS score			X	X						X								X		

Table 2 (continued)

Parameters	Giurea (2016) [24]	Hinarejos (2016) [40]	Kremers (2017) [59]	Jain (2017) [60]	Sanchez Santos (2018) [33]	Clement (2018) [20]	Van Onsem (2018) [43]	Clement (2018) [23]	Abrecht (2019) [13]	Calkins (2019) [31]	Twiggs (2019) [42]	Zabawa (2019) [14]	Tolk (2019) [18]	Kunze (2019) [39]	Clement (2019) [19]	Ramkumar (2019) [61]	Pua (2019) [26]	Xu (2019) [62]	Vissers (2020) [63]	Kunze (2020) [41]	Belford (2020) [64]
Quality assessment	Mod.	Mod.	Serious	Serious	Low	Mod.	Serious	Mod.	Mod.	Serious	Low	Mod.	Mod.	Mod.	Mod.	Low	Mod.	Mod.	Mod.	Mod.	Mod.
KOOS improvement																					
IKS score																					
IKS improvement																					
OKS score					X		X						X						X		X
OKS improvement					X																
EQ-5D score													X						X		
EQ-5D improvement																					
Satisfaction	X					X		X						X	X						
Post op ROM														X	X		X				
Revision risk			X																		

BMI, body mass index; EQ-5D, Euro QOL score; KOOS, Knee injury and Osteoarthritis Outcome Score; KSS, Knee Society Score; IA, inconsistent association; Mod., moderate; NA, no association; OKS, Oxford Knee Score; PROM, Patient-Reported Outcome Measures; ROM, range of motion; SA, strong association; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index; VAS, Visual Analog Scale.

from a cohort of 1649 patients [33]. Tolk et al. estimated the probability of residual symptoms after TKA based on individual PROM questions and PROM total scores [18,33]. However, the use of these models in clinical practice is neither intuitive nor practical. Furthermore, these models are built on relatively small data sets and not have wide applicability.

Patient satisfaction

In a cohort of 484 patients, Kunze et al. have developed a preoperative knee survey score to predict patient outcome and satisfaction at 1 year after TKA [39]. They concluded that a knee survey score of 96.5 would confer a 97.5% sensitivity and 95.7% negative predictive value for satisfaction, and a knee survey score of less than or equal to 96.5 increased the probability of experiencing postoperative dissatisfaction.

Patient pain

From a cohort of 4026 TKAs, Pua et al. developed a predictive model designed to determine the expected knee ROM, knee pain, and walking limitations of a patient 6 months postoperatively [25]. They created a web application to facilitate its use in clinical practice (https://sgh-physio.shinyapps.io/predicTKR/). However, no study has assessed the predictive value of this model.

Twiggs et al. created an algorithm designed to predict a patient's knee pain score 12 months after TKA [42]. The algorithm is based on a preoperative self-administered questionnaire and predicts the likelihood that a patient's change in the pain score will be equal to or greater than the MCID in the PROM score. The use of the MCID allows the patient and the surgeon to know during the preoperative consultation if the patient will likely experience a clinically significant improvement in pain after TKA.

PROM improvement

The most commonly used measurement to assess outcome after TKA is the improvement seen between results collected before surgery and postoperative data collected using validated instruments or objective measures such as ROM [19,21,34].

Discussion

The aim of this article was to identify and group the preoperative predictive factors and outcome measurement parameters which have been found to be predictive of clinical outcomes after TKA in the current literature. This compendium identifies those variables that are the most likely to be useful features in the context of predictive algorithms for clinical outcomes after TKA. These features are summarized in Table 4.

The challenge of developing useful predictive algorithms is twofold: (1) choosing the right predictive factors and (2) selecting those outcomes that are both "predictable" and useful measures of clinical satisfaction [18,46]. Indeed, in several studies, predictive factors and the outcomes are correlated independently of patient variables because the same questions, the same scores, and parameter are assessed before and after TKA. For this reason, multiple variable analysis is essential to assess the independent contribution of each feature to the prediction of postoperative outcomes. For example, a patient with a preoperative fixed flexion deformity has a higher risk of having a postoperative fixed flexion deformity [25]. In this context, the improvement in PROM scores can be very useful to assess TKA outcomes. The clinical relevance of the improvements can be quantified by the MCID. Unfortunately, MCIDs are not universally valid across populations and cultures and vary by instrument.

Another challenge of using predictive models built and validated with data from one population is extrapolating the results of

Table 3
Studies reporting a predictive model for TKA outcomes.

Study	Year	Sample size	Location	Predictive factors	Outcome measurement parameters	Delay	Predictive model
Judge et al. [6]	2012	1991	UK	Age, gender, BMI, primary diagnosis, ASA score, Index of Multiple Deprivation, OKS, EQ5D	Satisfaction, OKS	6 mo	N/A
Lungu et al. [46]	2014	141	Canada	5 Preoperative WOMAC questions: 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	WOMAC	6 mo	Predictive rule, based on 5 preop WOMAC questions
Van Onsem et al. [15]	2016	113	Belgium	Question selections based on KOOS, OKS, PCS, EQ-5D, KSS, age, and gender	KSS satisfaction subscore	3 mo	Algorithm: Satisfaction at M3 = 26.10 + 2.3*gender + 0.13*age + 1.58*Q3 − 1.40*Q4 − 1.08*Q5 − 0.75*Q6 − 1*Q7 − 1.12*Q8 − 0.88*Q9 − 1.10*Q10
Sanchez et al. [33]	2018	1649 (External validation on 595)	UK	Age, gender, marital status, Index of Multiple Deprivation, BMI, anxiety/depression, OKS, ASA score, etiology, previous knee arthroscopy, flexion contracture, ACL status.	OKS	12 mo	N/A
Van Onsem et al. [43]	2018	57	Belgium	Preop ROM, quadriceps and hamstring force, sit-to-stand test, 6-min walk test	KOOS, KSS, OKS	6 mo	N/A
Twiggs et al. [42]	2019	330 (2 external validations)	US/Australia	Age, gender, KOOS items, back pain, occurrence of hip pain, occurrence of falls in past year	Knee pain MCID = 10 points of KOOS pain score	12 mo	Predictive model with a web application
Tolk et al. [18]	2020	7071	NL	Age, gender, ASA score, BMI, smoking, previous knee surgery, Charnley score, KOOS-PS, OKS, EuroQoL 5D-3L, NRS	Residual symptoms (pain at rest and activity, sit-to-stand movement, stair negotiation, walking, performance of activities of daily living, kneeling and squatting).	6 and 12 mo	Predictive model for residual symptoms
Kunze et al. [39]	2019	484	US	BMI, drug allergies, osteophytes, soft tissue thickness, flexion contracture, diabetes, opioid use, comorbidity, previous knee surgery, surgical indication, smoking	Patient-reported health state, KSS, ROM, satisfaction => Knee survey score	12 mo	Knee survey score on 110 pts. 4 risks of experiencing postoperative dissatisfaction: Score 96.5-110 = low risk Score 75-96.4 = mild risk Score 60-74.9 = medium risk Score <60 = high risk
Ramkumar et al. [61]	2019	171,025	US	Age, gender, ethnicity, emergency department, risk of mortality, severity of illness, comorbidity weekend admission, hospital type, income	LOS charges/cost, discharge disposition		Model code (https://github.com/JaretK/ NeuralNetArthroplasty)
Pua et al. [25]	2019	4026	Singapore	Age, gender, race, educational level, diabetes, preop gait aids, contralateral knee pain, psychological distress	Knee extension, knee flexion, knee pain, walking limitation	6 mo	Prediction model with a web application (https://sgh-physio.shinyapps.io/predicTKR/)
Anis et al. [34]	2020	5958 and 2391	US	Age, gender, BMI, race, educational level, smoking, comorbidity, KOOS items, 12PCS, 12MCS	LOS, 90-days readmission, PROMS	12 mo	N/A

EQ-5D, Euro QOL score; KOOS, Knee injury and Osteoarthritis Outcome Score; KSS, Knee Society Score; LOS, length of stay; OKS, Oxford Knee Score; PCS, Pain Catastrophizing Scale; ROM, range of motion; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.

Table 4
Summary of the main preoperative predictive factors and outcome measurement parameters.

Strength of association	Predictive factors	Outcome measurement parameters	Delay

Strong correlation	• Pain
	  - VAS pain
	  - Back pain
	• Knee-specific PROMs
	  - KOOS
	  - WOMAC
	• Knee characteristics
	  - ROM
	• General PROMs
	  - EQ-5D
	• Mental health
	  - Anxiety/Depression
	  - SF-12
		• Improvement of knee-specific PROMs
		  - OKS
		  - KOOS
		  - WOMAC
		• Satisfaction
		  - Self-assessment of improvement
		  - KSS satisfaction subscale
		• Pain
		  - VAS pain
		  - WOMAC pain
		  - Persistent pain
			6 and 12 mo

Inconsistent correlation	• Comorbidities/ASA score
	• BMI
	• Gender
	• Age
	• Previous knee surgery
	• Severity of osteoarthritis
	• Preop knee alignment
		• Knee-specific PROMs
		  - OKS
		  - KOOS
		  - WOMAC
		  - SF-36
		• General PROMs	

BMI, body mass index; EQ-5D, Euro QOL score; KOOS, Knee injury and Osteoarthritis Outcome Score; KSS, Knee Society Score; OKS, Oxford Knee Score; PROM, Patient-Reported Outcome Measures; ROM, range of motion; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index; VAS, Visual Analog Scale.

the algorithm using data from other populations. For example, Zabawa et al. [14] and Calkins et al. [31] did not support the validity of the Van Onsem prediction tool [15]. These differences can have several explanations. First, the populations under study can be very different from one country to another. For example, the mean BMI may vary considerably between 2 populations, and if this parameter is identified as a predictive factor in the algorithm, the model may simply be inaccurate in the second population. Second, the indications for TKA can change between countries or between centers in the same country, particularly relative to age, BMI, and osteoarthritis stage. Thus, the result from the model for a given patient may meet threshold criteria in one country/center but not in another. For this reason, predictive models using data from very large populations numbering in the hundreds of thousands, including several centers or countries, are more relevant and reliable [18].
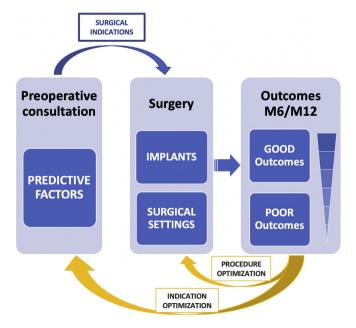

Figure 2. Diagram explaining the correlation between predictive factors and outcomes in a predictive model.

It is also worth noting that while some correlations were identified between preoperative variables and postoperative clinical outcomes, the strength of even the best correlations was underwhelming. While it is possible that larger and more accurate data sets may increase the validity and predictive value of the algorithms, it is also possible that entirely different end points will be required. The existing "gold standard" PROMs are now several decades old and measured outcomes that were tied to problems faced by implant technology and surgical techniques that are now antiquated. While the primary concern in the latter 20th century was with implant survivorship of TKA, in the first 2 decades of the 21st century, attention has turned to functional outcomes after that procedure. However, the functional aspect of most scoring systems sets a reasonably low bar for success such as standing up from a chair rather than playing a round of golf. This is one reason so many PROMs have well-defined ceiling effects [47–50]. Some more demanding functional scores are sometimes used, such as the forgotten joint score, the WOMAC score, and the UCLA score. But they remain rarely used to assess the TKA outcomes. Maybe other assessment methods would be necessary, such as the gait analysis or digital care management platforms with tools to have connected patients. Nevertheless, these devices are currently lightly used and not described or assessed in the studies on the predictive models.

In terms of modeling technique, a substantial number of publications relied on traditional regression models which are robust and provide a quantitative assessment of the predictor's relationship with the output through the investigation of the model's coefficients [10,18,20,36,37]. However, machine learning techniques have been shown to outperform linear regression model in specific tasks, such as prediction of post-TKA EQ-5D-3L visual analog scale [51], estimating risk of total joint replacement [52], and more recently length-of-stay prediction after TKA [34,53]. More and more studies try to use the machine learning to predict the TKA outcomes and to adapt the practices according to the established predictive features [21,26,29,41]. The development of clinical decision-making tools generated from machine learning, which can be used in consultation to help discuss risk stratification with patients, could provide a means of better understanding which patients are at a greater risk to experience dissatisfaction after primary TKA [34,54,55]. Therefore, we expect the level of adoption of machine learning models to increase in light of the promising

results reported in some of the publications reviewed herein [44]. We hope that future research in this field will adopt the best practice of benchmarking different algorithms to a given prediction problem as we clearly have seen that there is no "one-size-fits-all" best solution in predictive modeling for TKA clinical outcomes.

Our findings should be considered in the light of the key limitations of the data set. First, the inclusion criteria, such as English language or the requirement of full text access, may have excluded relevant research. Second, the methodology score has known limitations with regard to the type of studies included (predictive cohort-based studies) and the difficulties in assessing the validity of the analyses conducted without having access to the raw data. Third, there was an important variability between the studies with respect to the type of predictive features or outcome measurements, the follow-up period, the patient population, and the analyses performed. This heterogeneity limits the possibility of performing a true meta-analysis. Finally, the accuracy of predictive algorithms is derived from 2 critical aspects of the data set on which they are constructed: their size and the accuracy and completeness of the data sets within them. The larger the number of variables that can influence an outcome and the more complex the interaction between these variables, the larger the data set needs to be to discern and predict these complex interactions.

The indications for TKA are complex and multifactorial. Selecting patients for surgery based strictly on the prediction of clinical outcomes alone is probably not reasonable. However, predictive models can communicate information and insight to both patients and surgeons that can be included in a shared decision-making process. The output of these algorithms might 1 day be expanded from simply predicting outcomes to providing a stratification in the variation of possible outcomes based on the pursuit of different surgical strategies. In this scenario, the surgeon and patient would essentially customize the procedure to optimize the likelihood of meeting the patient's needs. Examples include the decision between partial and total knee replacement or the choice of a cruciate retaining or cruciate sacrificing TKA. By feeding results back to the data set, the models evolve and improve over time and become increasingly accurate (Fig. 2). We expect that such predictive models, when trained with appropriate and accurate data sets, could become an important adjunct to daily clinical practice in the near future.

Conclusion

The existing literature on predictive modeling of clinical outcomes after TKA has identified preoperative variables that have at least some correlation with clinical results. Functional features such as pain, PROMs scores, or mental health were highly predictive for clinical outcomes after TKA. Some variables such as demographics data, surgical history, or knee alignment were less strongly correlated with TKA outcomes. The challenge of developing useful predictive algorithms is further complicated by the need to select the most appropriate measurement parameters of TKA outcomes such as improvement in PROMs, patient satisfaction, or postoperative pain. Creating accurate and reproducible predictive algorithms may 1 day provide advanced tools for shared decision-making relative to surgical indications and expected outcomes. However, the data gathered also suggested that work is still required to define outcomes measures that more accurately correlate with preoperative variables and better reflect patient satisfaction.

Conflicts of interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing

References

[1] Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KD. Patient satisfaction after total knee arthroplasty: who is satisfied and who is not? Clin Orthop Relat Res 2010;468(1):57.

[2] 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):e000435.

[3] Bini SA, Shah RF, Bendich I, Patterson JT, Hwang KM, Zaid MB. Machine learning algorithms can use wearable sensor data to accurately predict six-week patient-reported outcome scores following joint replacement in a prospective trial. J Arthroplasty 2019;34(10):2242.

[4] Baker PN, van der Meulen JH, Lewsey J, Gregg PJ. National Joint Registry for E, Wales. The role of pain and function in determining patient satisfaction after total knee replacement. Data from the National Joint Registry for England and Wales. J Bone Joint Surg Br 2007;89(7):893.

[5] Baker PN, Deehan DJ, Lees D, 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.

[6] Judge A, Arden NK, Cooper C, et al. Predictors of outcomes of total knee replacement surgery. Rheumatology (Oxford) 2012;51(10):1804.

[7] Brander VA, Stulberg SD, Adams AD, et al. Predicting total knee replacement pain: a prospective, observational study. Clin Orthop Relat Res 2003;(416):27.

[8] Wylde V, Rooker J, Halliday L, Blom A. Acute postoperative pain at rest after hip and knee arthroplasty: severity, sensory qualities and impact on sleep. Orthop Traumatol Surg Res 2011;97(2):139.

[9] Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health Inf Sci Syst 2014;23.

[10] Escobar A, Quintana JM, Bilbao A, et al. Development of explicit criteria for prioritization of hip and knee replacement. J Eval Clin Pract 2007;13(3):429.

[11] Riddle DL, Perera RA, Jiranek WA, Dumenci L. Using surgical appropriateness criteria to examine outcomes of total knee arthroplasty in a United States sample. Arthritis Care Res (Hoboken) 2015;67(3):349.

[12] Sterne JA, Hernan MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016:355i4919.

[13] Abrecht CR, Cornelius M, Wu A, et al. Prediction of pain and opioid utilization in the perioperative period in patients undergoing primary knee arthroplasty: psychophysical and psychosocial factors. Pain Med 2019;20(1):161.

[14] Zabawa L, Li K, Chmell S. Patient dissatisfaction following total knee arthroplasty: external validation of a new prediction model. Eur J Orthop Surg Traumatol 2019;29(4):861.

[15] Van Onsem S, Van Der Straeten C, Arnout N, Deprez P, Van Damme G, Victor J. A new prediction model for patient satisfaction after total knee arthroplasty. J Arthroplasty 2016;31(12):2660.

[16] Lewis GN, Rice DA, McNair PJ, Kluger M. Predictors of persistent pain after total knee arthroplasty: a systematic review and meta-analysis. Br J Anaesth 2015;114(4):551.

[17] Huijbregts HJ, Khan RJ, Fick DP, Jarrett OM, Haebich S. Prosthetic alignment after total knee replacement is not associated with dissatisfaction or change in Oxford Knee Score: a multivariable regression analysis. Knee 2016;23(3):535.

[18] Tolk JJ, Waarsing JEH, Janssen RPA, van Steenbergen LN, Bierma-Zeinstra SMA, Reijman M. Development of preoperative prediction models for pain and functional outcome after total knee arthroplasty using the Dutch arthroplasty register data. J Arthroplasty 2020;35(3):690.

[19] Clement ND, Merrie KL, Weir DJ, Holland JP, Deehan DJ. Asynchronous bilateral total knee arthroplasty: predictors of the functional outcome and patient satisfaction for the second knee replacement. J Arthroplasty 2019;34(12):2950.

[20] Clement ND, Walker LC, Bardgett M, et al. Patient age of less than 55 years is not an independent predictor of functional improvement or satisfaction after total knee arthroplasty. Arch Orthop Trauma Surg 2018;138(12):1755.

[21] Harris AHS, Kuo AC, Bowe TR, Manfredi L, Lalani NF, Giori NJ. Can machine learning methods produce accurate and easy-to-use preoperative prediction models of one-year improvements in pain and functioning after knee arthroplasty? J Arthroplasty 2021;36(1):112.

[22] Maratt JD, Lee YY, Lyman S, Westrich GH. Predictors of satisfaction following total knee arthroplasty. J Arthroplasty 2015;30(7):1142.
[23] Clement ND, Bardgett M, Weir D, Holland J, Gerrard C, Deehan DJ. Three groups of dissatisfied patients exist after total knee arthroplasty: early, persistent, and late. Bone Joint J 2018;100-B(2):161.
[24] Giurea A, Fraberger G, Kolbitsch P, et al. The impact of personality traits on the outcome of total knee arthroplasty. Biomed Res Int 2016. 20165282160.
[25] Pua YH, Poon CL, Seah FJ, et al. Predicting individual knee range of motion, knee pain, and walking limitation outcomes following total knee arthroplasty. Acta Orthop 2019;90(2):179.
[26] Pua YH, Kang H, Thumboo J, et al. Machine learning methods are comparable to logistic regression techniques in predicting severe walking limitation following total knee arthroplasty. Knee Surg Sports Traumatol Arthrosc 2020;28(10):3207.
[27] Blackburn J, Qureshi A, Amirfeyz R, Bannister G. Does preoperative anxiety and depression predict satisfaction after total knee replacement? Knee 2012;19(5):522.
[28] Bourne RB, McCalden RW, MacDonald SJ, Mokete L, Guerin J. Influence of patient factors on TKA outcomes at 5 to 11 years followup. Clin Orthop Relat Res 2007:46427.
[29] Farooq H, Deckard ER, Ziemba-Davis M, Madsen A, Meneghini RM. Predictors of patient satisfaction following primary total knee arthroplasty: results from a traditional statistical model and a machine learning algorithm. J Arthroplasty 2020;35(11):3123.
[30] Maempel JF, Clement ND, Brenkel IJ, Walmsley PJ. Range of movement correlates with the Oxford knee score after total knee replacement: a prediction model and validation. Knee 2016;23(3):511.
[31] Calkins TE, Culvern C, Nahhas CR, et al. External validity of a new prediction model for patient satisfaction after total knee arthroplasty. J Arthroplasty 2019;34(8):1677.
[32] Franklin PD, Li W, Ayers DC. The Chitranjan Ranawat Award: functional outcome after total knee replacement varies with patient attributes. Clin Orthop Relat Res 2008;466(11):2597.
[33] Sanchez-Santos MT, Garriga C, Judge A, et al. Development and validation of a clinical prediction model for patient-reported pain and function after primary total knee replacement surgery. Sci Rep 2018;8(1):3381.
[34] Anis HK, Strnad GJ, Klika AK, et al. Developing a personalized outcome prediction tool for knee arthroplasty. Bone Joint J 2020;102-B(9):1183.
[35] Dowsey MM, Liew D, Stoney JD, Choong PF. The impact of pre-operative obesity on weight change and outcome in total knee replacement: a prospective study of 529 consecutive patients. J Bone Joint Surg Br 2010;92(4):513.
[36] Schnurr C, Jarrous M, Gudden I, Eysel P, Konig DP. Pre-operative arthritis severity as a predictor for total knee arthroplasty patients' satisfaction. Int Orthop 2013;37(7):1257.
[37] Sueyoshi T, Lackey WG, Malinzak RA, et al. Predicting pain in total and partial knee arthroplasty. Open J Orthop 2015:5151.
[38] 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.
[39] Kunze KN, Akram F, Fuller BC, Zabawa L, Sporer SM, Levine BR. Internal validation of a predictive model for satisfaction after primary total knee arthroplasty. J Arthroplasty 2019;34(4):663.
[40] Hinarejos P, Ferrer T, Leal J, Torres-Claramunt R, Sanchez-Soler J, Monllau JC. Patient-reported allergies cause inferior outcomes after total knee arthroplasty. Knee Surg Sports Traumatol Arthrosc 2016;24(10):3242.
[41] Kunze KN, Polce EM, Sadauskas AJ, Levine BR. Development of machine learning algorithms to predict patient dissatisfaction after primary total knee arthroplasty. J Arthroplasty 2020.
[42] Twiggs JG, Wakelin EA, Fritsch BA, et al. Clinical and statistical validation of a probabilistic prediction tool of total knee arthroplasty outcome. J Arthroplasty 2019;34(11):2624.
[43] Van Onsem S, Verstraete M, Dhont S, Zwaenepoel B, Van Der Straeten C, Victor J. Improved walking distance and range of motion predict patient satisfaction after TKA. Knee Surg Sports Traumatol Arthrosc 2018;26(11):3272.
[44] Davis ET, Lingard EA, Schemitsch EH, Waddell JP. Effects of socioeconomic status on patients' outcome after total knee arthroplasty. Int J Qual Health Care 2008;20(1):40.
[45] Nilsdotter AK, Toksvig-Larsen S, Roos EM. Knee arthroplasty: are patients' expectations fulfilled? A prospective study of pain and function in 102 patients with 5-year follow-up. Acta Orthop 2009;80(1):55.
[46] Lungu E, Desmeules F, Dionne CE, Belzile EL, Vendittoli PA. Prediction of poor outcomes six months following total knee arthroplasty in patients awaiting surgery. BMC Musculoskelet Disord 2014:15299.
[47] Lim CR, Harris K, Dawson J, Beard DJ, Fitzpatrick R, Price AJ. Floor and ceiling effects in the OHS: an analysis of the NHS PROMs data set. BMJ Open 2015;5(7):e007765.
[48] Hamilton DF, Giesinger JM, MacDonald DJ, Simpson AH, Howie CR, Giesinger K. Responsiveness and ceiling effects of the Forgotten Joint Score-12 following total hip arthroplasty. Bone Joint Res 2016;5(3):87.
[49] Lyman S, Lee YY, Franklin PD, Li W, Mayman DJ, Padgett DE. Validation of the HOOS, JR: a short-form hip replacement survey. Clin Orthop Relat Res 2016;474(6):1472.
[50] Steinhoff AK, Bugbee WD. Knee Injury and Osteoarthritis Outcome Score has higher responsiveness and lower ceiling effect than Knee Society Function Score after total knee arthroplasty. Knee Surg Sports Traumatol Arthrosc 2016;24(8):2627.
[51] Huber M, Kurz C, Leidl R. Predicting patient-reported outcomes following hip and knee replacement surgery using supervised machine learning. BMC Med Inform Decis Mak 2019;19(1):3.
[52] Qiu R, Jia Y, Wang F, et al. Predictive modeling of the total joint replacement surgery risk: a deep learning based approach with claims data. AMIA Jt Summits Transl Sci Proc 2019:2019562.
[53] Li H, Jiao J, Zhang S, Tang H, Qu X, Yue B. Construction and comparison of predictive models for length of stay after total knee arthroplasty: regression model and machine learning analysis based on 1,826 cases in a single Singapore center. J Knee Surg 2020.
[54] Cochrane JA, Flynn T, Wills A, Walker FR, Nilsson M, Johnson SJ. Clinical decision support tools for predicting outcomes in patients undergoing total knee arthroplasty: a systematic review. J Arthroplasty 2020.
[55] Mahajan SM, Nguyen C, Bui J, Kunde E, Abbott BT, Mahajan AS. Risk factors for readmission after knee arthroplasty based on predictive models: a systematic review. Arthroplast Today 2020;6(3):390.
[56] 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(10):2179.
[57] Barlow T, Dunbar M, Sprowson A, Parsons N, Griffin D. Development of an outcome prediction tool for patients considering a total knee replacement–the Knee Outcome Prediction Study (KOPS). BMC Musculoskeletal Disord 2014:15451.
[58] Feldman CH, Dong Y, Katz JN, Donnell-Fink LA, Losina E. Association between socioeconomic status and pain, function and pain catastrophizing at presentation for total knee arthroplasty. BMC Musculoskelet Disord 2015:1618.
[59] Maradit Kremers H, Kremers WK, Berry DJ, Lewallen DG. Patient-reported outcomes can Be used to identify patients at risk for total knee arthroplasty revision and potentially individualize postsurgery follow-up. J Arthroplasty 2017;32(11):3304.
[60] Jain D, Nguyen LL, Bendich I, et al. Higher patient expectations predict higher patient-reported outcomes, but not satisfaction, in total knee arthroplasty patients: a prospective multicenter study. J Arthroplasty 2017;32(9S):S166.
[61] Ramkumar PN, Karnuta JM, Navarro SM, et al. Deep learning preoperatively predicts value metrics for primary total knee arthroplasty: development and validation of an artificial neural network model. J Arthroplasty 2019;34(10):2220.
[62] Xu J, Twiggs J, Parker D, Negus J. The association between anxiety, depression, and locus of control with patient outcomes following total knee arthroplasty. J Arthroplasty 2020;35(3):720.
[63] Vissers LCM, van Hove RP, van der Zwaard BC. Predicting self-reported functional improvement one year after primary total knee arthroplasty using pre- and postoperative patient-reported outcome measures. Knee 2020;27(3):683.
[64] Belford K, Gallagher N, Dempster M, et al. Psychosocial predictors of outcomes up to one year following total knee arthroplasty. Knee 2020;27(3):1028.
[65] Itou J, Itoh M, Kuwashima U, Okazaki K. Assessing the validity of a new prediction model for patient satisfaction after total knee arthroplasty: a retrospective cross-sectional study. Orthop Res Rev 2020:12133.

Supplementary Table 1

Table reporting the 3 different types of predictive factors according to the strength of their association with TKA outcomes.

			Joint specific PROMs								Function	General PROMs	Satisfaction		Pain			
			OKS (Q-score)	Improvement in OKS	WOMAC	WOMAC Func	WOMAC Stiffness	KOOS	Change in KOOS	SF-36	post-op ROM	EQ5 D VAS	Self assessment of outcomes improvement	KSS satisfaction subscale	VAS Pain	WOMAC Pain	No pain relief	Opioid consumption
Clear association with (poor) outcomes	Pain	Pre-op VAS Pain		Huijbregts (2016)										Van Onsem (2016) Zabawa (2019)	Abrecht (2019)			Abrecht (2019)
		Neurological disease / Back pain				Escobar (2007)	Escobar (2007), Clement (2019)		Twiggs (2019)				Clement (2018)			Escobar (2007)		
	Joint specific PRE-op PROMs	Pre-op knee function scores (KOOS pain/ function)	Sanchez-Santos (2018)		Lungu (2014)	Lingard (2018) Lungu (2014)	Lungu (2014)		Twiggs (2019)					Van Onsem (2016)		Lungu (2014)		
		Pre-op WOMAC Function			Allyson Jones (2003) Lingard (2004) Rajgopal (2008) Nunez (2009)	Escobar (2007) Lingard (2018) Nunez (2007)	Clement (2019)			Allyson Jones (2003) Lingard (2004)						Lopez-Olivio (2011) Clement (2019)		
		Pre-op WOMAC Pain											Clement (2018)	Van Onsem (2016)		Nunez (2007) Clement (2019)		
		Worse Pre-op WOMAC Stiffness			Lungu (2014)	Lungu (2014)	Clement (2019) Nunez (2007)							Van Onsem (2016)		Lungu (2014) Clement (2019)		
		Pre-op SF-12 PCS/SF-36		Huijbregts (2016)	Lingard (2004)	Escobar (2007) Clement (2019)	Escobar (2007)			Lingard (2004)			Clement (2019)			Escobar (2007) Clement (2019)		
	Knee	Absent or damaged ACL pre-op		Sanchez-Santos (2018)														
		Pre-op Range of Motion (ROM)		Sanchez-Santos (2018)										Van Onsem (2018)				
	gen. PROM	Pre-op EQ5D VAS	Huber (2019)		Maratt (2015)							Huber (2019)						
	Mental health	Depression/ Anxiety			Xu (2019)	Lopez-Olivio (2011)					Judge (2012)		Clement (2018) Giurea (2016)	Van Onsem (2016) Zabawa (2019)	Abrecht (2019)	Clement (2019)		Abrecht (2019)
		Ability to cope	Sanchez-Santos (2018)			Lopez-Olivio (2011)							Giurea (2016)	Van Onsem (2016)				
		Hospital Anxiety and Depression Scale (HAD)		Blackburn (2012)	Xu (2019)													
		Pre-op SF-12 MCS			Rajgopal (2008) Xu (2019)	Escobar (2007) Clement (2019)	Escobar (2007) Clement (2019)			Lingard (2004) Franklin (2008)			Clement (2018)			Escobar (2007) Clement (2019)		
	other	Geography (UK vs US/ AUS)			Lingard (2018)													

Supplementary Table 1 (continued)

		Variable	Joint specific PROMs	Function	General PROMs	Satisfaction	Pain
Inconsistent association	Demographics	Joint co-morbidity	Rajgopal (2008); Twiggs (2019)				
		Occurance of falls in preceding year				Kunze (2019)	Hinarejos (2016)
		Allergy (>1 self-reported)	Nunez (2007)	Nunez (2007)			Dave (2017) Nunez (2007)
		Wide-spread body pain	Hinarejos (2016) Nunez (2007)			Schnurr (2013) Kunze (2019)	
		Severity osteoarthritis (Kellgren-Lawrence)	Vissers (2020) Judge (2012)				
		Medical co-morbidities / ASA score	Sanchez-Santos (2018); Lingard (2018) Allyson Jones (2003) Lingard (2004) Nunez (2009) Nunez (2011); Escobar (2007)	Lingard (2004) Hilton (2016)			Allyson Jones (2003) Escobar (2007) Lopez-Olivio (2011) Hilton (2016)
		BMI	Sanchez-Santos (2018) Judge (2012); Allyson Jones (2003) Lingard (2004) Bourne (2007) Rajgopal (2008); Nunez (2007) Nunez (2009) Lopez-Olivio (2011); Twiggs (2019)	Maempel (2016); Allyson Jones (2003) Lingard (2004) Cushnaghan (2008) Franklin (2008) Dowsey (2010)		Clement (2018) Kunze (2019); Zabawa (2019) Calkins (2019)	Abrecht (2019); Sueyoshi (2015) Abrecht (2019)
		Gender	Judge (2012); Allyson Jones (2003) Nunez (2007) Bourne (2007) Rajgopal (2008); Lingard (2004) Escobar (2007) Nunez (2009); Twiggs (2019)	Maempel (2016); Kiebzak (2002) Allyson Jones (2003) Cushnaghan (2008) Franklin (2008)		Van Onsem (2016) Zabawa (2019) Van Onsem (2018) Calkins (2019)	Abrecht (2019); Nunez (2007-2009) Lopez-Olivio (2011) Clement (2019); Sueyoshi (2015) Abrecht (2019)
		Age	Clement (2012); Nunez (2007) Bourne (2007) Rajgopal (2008); Allyson Jones (2003) Escobar (2007) Cushnaghan (2008); Escobar (2007) Twiggs (2019)	Maempel (2016); Allyson Jones (2003) Lingard(2004) Cushnaghan (2008) Franklin (2008) Clement (2012)	Abrecht (2019)	Van Onsem (2016) Calkins (2019); Huijbregts (2016) Schnurr (2013)	Abrecht (2019); Lingard (2004) Escobar (2007) Clement(2019); Clement (2019) Escobar (2007); Sueyoshi (2015) Abrecht (2019)
	Knee	(No) previous knee surgery	Huber (2019) Sanchez-Santos (2018); Rajgopal (2008); Twiggs (2019)			Kunze (2019)	Abrecht (2019)
	Knee	Pre-op knee alignment	Twiggs (2019)				Sueyoshi (2015)
	Social	Income	Sanchez-Santos (2018); Davis (2008)				Davis (2008)
	Social	Decreased social support	Sanchez-Santos (2018) Judge (2012); Escobar (2007) Lopez-Olivio (2011); Escobar (2007); Twiggs (2019)				Escobar (2007)
Low or no significant correlation	Education / Socioeconomic status (SES)		Davis (2008) Feldman (2015); Escobar (2007) Lopez-Olivio (2011)	Pua (2019) (2015)		Pua (2019)	Lopez-Olivio (2011)

Smoking /
Drinking
Employment
status
Expectation

Ethnicity

Quadriceps st
rength
Pre-op pain
medication

Nilsdotter
(2009)

Twiggs
(2019)

Cushnaghan
(2008)

Twiggs
(2019)

Lopez-Olivio
(2011)

Twiggs
(2019)

Van Onsem
(2018)