Clinical Biomechanics 95 (2022) 105625

Contents lists available at ScienceDirect

Clinical Biomechanics

journal homepage: www.elsevier.com/locate/clinbiomech

Assessing functional recovery following total knee replacement surgery using objective classification of level gait data and patient-reported outcome measures

G.M. Whatling [a,b,*], P.R. Biggs [a,b], C. Wilson [b,c], C.A. Holt [a,b]

[a] Cardiff School of Engineering, College of Physical Sciences and Engineering, Cardiff University, Cardiff, UK
[b] Biomechanics and Bioengineering Research Centre Versus Arthritis, Cardiff University, Cardiff, UK
[c] University Hospital of Wales, Cardiff, UK

A R T I C L E I N F O

Keywords:
Total knee replacement
Gait
Motion analysis
Biomechanics

A B S T R A C T

Background: Patient recovery can be quantified objectively, via gait analysis, or subjectively, using patient reported outcome measures. Association between these measures would explain the level of disability reported in patient reported outcome measures and could assist with therapeutic decisions.
Methods: Total knee replacement outcome was assessed using objective classification and patient-reported outcome measures (Knee Outcome Survey and Oxford Knee Scores). A classifier was trained to distinguish between healthy and osteoarthritic characteristics using knee kinematics, ground reaction force and temporal gait data, combined with anthropometric data from 32 healthy and 32 osteoarthritis knees. For the osteoarthritic cohort, classification of 20 subjects quantified changes at up to 3 timepoints post-surgery.
Findings: Osteoarthritic classification was reduced for 17 subjects when comparing pre- to post-operative assessments, however only 6 participants achieved non-pathological classification and only 4 of these were classified as non-pathological at 12 months. In 15 cases, the level of osteoarthritic classification did not decrease between every post-operative assessment. For an individual's recovery, classification outputs correlated ($r > 0.5$) with knee outcome survey for 75% of patients and oxford knee score for 78% of patients (based on 20 and 9 subjects respectively). Classifier outputs from all visits of the combined total knee replacement sample correlated moderately with knee outcome survey ($r > 0.4$) and strongly with oxford knee score ($r > 0.6$).
Interpretation: Biomechanical deficits existed in most subjects despite improvements in Patient Reported Outcome Measures, with larger changes reported subjectively as compared to measured objectively. Objective Classification provides additional insight alongside Patient Reported Outcomes when reporting recovered outcomes.

1. Introduction

The main goals of total knee replacement (TKR) are to reduce pain associated with osteoarthritis (OA) and restore joint function (Andriacchi, 1993; Myles et al., 2002; Shenoy et al., 2013). Useful feedback on TKR biomechanical performance can be obtained from a range of assessments including physician based rating scales, patient questionnaires (Bachmeier et al., 2001; Gao et al., 2017; Murray and Frost, 1998; Yap et al., 2021), fluoroscopic assessment with image registration (Williams et al., 2020), and motion analysis (Andriacchi et al., 1982; Catani et al., 2003; McClelland et al., 2007; Milner, 2009; Rahman et al., 2015). Patient-reported outcome measures (PROMs) such as the Knee Outcome Survey (KOS) (Irrgang et al., 1998) and Oxford Knee Score (OKS) (Dawson et al., 1998) provide a subjective measure of a patient's wellbeing in terms of pain and daily living activities at a particular instance in time. Scores completed at different times during patient recovery provides subjective measures of changes in function. Motion analysis is a more detailed and objective approach to quantifying biomechanics and level of joint function pre and post-surgery. It provides detailed information about the performance of a joint and how it differs to healthy joints, thus providing insight into limitations to movement associated with OA and changes as a result of a TKR (McClelland et al., 2007; Milner, 2009; Rahman et al., 2015).

The general goal of motion analysis is to communicate reliable, objective information on which to base clinical decisions. This can be used to quantify surgical outcomes or the effects of therapeutic or

* Corresponding author at: Cardiff School of Engineering, College of Physical Sciences and Engineering, Cardiff University, Cardiff, UK.
E-mail address: Whatlinggm@cardiff.ac.uk (G.M. Whatling).

https://doi.org/10.1016/j.clinbiomech.2022.105625
Received 23 July 2020; Accepted 11 March 2022
Available online 16 March 2022

orthotic interventions. No individual biomechanical variable can provide a complete description of a subject's gait (Jacobs et al., 1972), and the relevance of each variable in discriminating levels of joint function is generally unknown. Thus, several variables should be considered in combination. TKR performance is commonly defined using discrete variables which are statistically different from healthy or OA controls (Milner, 2009; Rahman et al., 2015). However, using this information to provide a decision on outcome relies on expert opinion and may be considered subjective. Further, it does not provide a definitive classification of knee function to allow direct comparison between subjects or assessment sessions. For these reasons, there is increasing interest in summary gait measures.

Objective classification can automate a diagnosis from motion analysis data (Jones and Holt, 2008). It provides quantitative information and visual outputs of the level of pre-operative knee function and recovery post-TKR. The method is highly accurate in objectively classifying osteoarthritic knees. The classifier is trained using data (anthropometric, knee kinematics, ground reaction force and temporal) from subjects with knee OA and those with no–pathology (NP). A subject's knee is then classified in terms of its level of NP and OA characteristics. Datasets from different assessments can be compared directly to identify the level of benefit achieved by surgery. Gait following TKR does not return to normal (Andriacchi, 1993; Benedetti et al., 2003; Fuchs et al., 2002; Milner, 2009; Naili et al., 2017) and the classifier provides a method, both quantitatively and visually, to explore the level of recovery, using diagnosis specific variables, and explore what discrepancies still exist when compared to a healthy population. The classification of gait data is useful for diagnosis and monitoring. The importance of being able to identify how gait deviates from NP is demonstrated by the numerous approaches used today (Cimolin and Galli, 2014), including the movement deviation profile (Barton et al., 2012), normalcy index (Schutte et al., 2000), gait deviation index (Schwartz and Rozumalski, 2008) and gait profile score (Baker et al., 2009). PROMs are widely used in clinics as a tool to determine TKR outcome and this study determines whether there is agreement between KOS and OKS with the objectively generated classifier outputs.

This study aimed to determine whether the pattern of recovery for an individual is described similarly when using patient perceived PROMS and objective functional measures. The novelty lies in the investigation of patient recovery trajectories measured at multiple time points from pre- post TKR using both PROMS and a summative, objective measure that describes the changing levels of osteoarthritis characteristics exhibited by patients in response to their surgery.

2. Materials and methods

This study (i) quantifies the levels of NP and OA knee function exhibited by subjects prior to TKR surgery using objective functional classification of gait and anthropometric measures, (ii) monitors how their arthritic characteristics change during a 9- to 15-month recovery period following surgery, (iii) demonstrates the interpretation of classification outputs and (iv) determines whether these changes are mirrored by recovery determined using subjective KOS and OKS.

2.1. Sample

Data collected from two cohorts were used to train the classifier to identify the differences between OA and NP. 32 subjects without lower limb pathology represented NP, and 29 patients (32 knees) listed for TKR surgery represented knee OA. Data for one knee from the NP cohort and all the affected knees from the OA cohort were reported.

A post-TKR sample was collected for 20 knees (from 19 subjects) from the OA cohort, where a further two to three gait assessments were performed at approximately three, six- and twelve-months post-operation. Ethical approval was granted by Bro Taf Authority Local Research Ethics Committee (Reference number 98/2610). The inclusion criteria for patients were the ability to provide informed consent, no previous injury to the joint under investigation or other pathologies that may affect the way they walk. In addition to these, the inclusion criteria for the NP participants were no history of pathology or instability of the joint under investigation. Patients with bi-lateral OA were not excluded.

2.2. Data collection protocol

Anthropometric measurements recorded were height, mass, distal thigh girth, medio-lateral (ML) knee width and anterior-posterior (AP) knee depth. Anatomical calibration using manual palpation and a digitising pointer with one second recordings, identified the upper border of the trochanter, medial and lateral epicondylar gaps and medial malleolus (Beynon et al., 2006). These were used to establish femoral and tibial anatomical coordinate systems. Rigid marker clusters were attached laterally to the thigh and shank, positioned to minimise skin movement (Cappello et al., 1997). The positional relationship between marker clusters and anatomical axes were determined allowing segments to be tracked using the marker clusters. Subjects performed six trials of barefoot level gait at a self-selected pace. Three-dimensional motion analysis was performed using eight 120 Hz infra-red motion capture units capturing at 60 Hz (Qualisys, Gothenburg, Sweden) and two force platforms (Bertec Corporation, Ohio, USA) embedded in the floor, capturing at 1080 Hz. Knee kinematics, temporal parameters and ground reaction forces (GRFs) were quantified (Beynon et al., 2006).

Each subject completed the KOS (Irrgang et al., 1998) and nine subjects completed the OKS (Dawson et al., 1998) at each assessment. This is a retrospective analysis and the limited number of subjects with completed OKS is due to the more recent introduction of this score into the data collection protocols. Both scores allow participants to subjectively evaluate their pain, symptoms and functional limitations to daily activities imposed by their knee pathology or surgery. Scores were calculated as a percentage, where 100% indicates no symptoms.

2.3. Data processing

Motion and GRF data were processed using custom Matlab software and previously published methods of Holt et al. (2000). Knee kinematics waveforms were re-sampled over a 100% gait cycle and an ensemble average of up to six gait trials computed. GRF waveforms were re-sampled at 60 Hz, over 100% stance, normalised to body weight, and trials were averaged.

Principal Component Analysis (PCA) of the knee rotation and ground reaction force waveforms was performed to determine PC scores that describe the waveforms by retaining temporal information (Deluzio et al., 1997). PC reconstruction was performed to interpret the biomechanical feature represented by each component (Brandon et al., 2013). The variables used to represent knee function for classifications, with their interpretations, are listed in Table 1. In addition to PCs of the ground reaction force and knee rotation waveforms, BMI, cadence, percentage stance phase, and measurements of knee width, depth and distal thigh girth were included as these were demonstrated to be discriminatory in classifying knee OA (Jones et al., 2008).

2.4. Classification procedure

The classification method employed has been demonstrated previously in knee OA analyses (Beynon et al., 2006; Biggs et al., 2019a, 2019b; Jones et al., 2006; Jones et al., 2008; Jones and Holt, 2008) and is based on the Dempster-Shafer theory of evidence, providing a non-Bayesian way of using mathematical probability to quantify subjective judgements. It allows for a degree of uncertainty in the decision-making process as to whether a gait variable indicates OA, to deal with the conflicting and corroborating nature of motion analysis data. Levels of support are assigned to each measurement variable and these are combined to classify knee function as NP or OA.

Table 1
Summary of the variables used to classify the level of recovery towards NP gait following knee replacement surgery.

Variable	Variable description	Variable interpretation
v_1	Body mass index (BMI) (kg/ m^2)	Contributor to knee loading
v_2	Cadence (min^{-1})	Indicator of ability to walk with a normal gait
v_3	Stance phase (per cent gait cycle)	Indicator of ability to walk with a normal gait
v_4	PCs representing the anterior-posterior GRF waveform	Magnitude of anterior and posterior peaks with more emphasis on the posterior peak; and the timing of the transition from a posterior to an anterior force.
v_5		Magnitude of anterior and posterior peaks with more emphasis on the anterior peak; and the timing of the transition from a posterior to an anterior force.
v_6		Timing of the anterior GRF peak and rate of load transfer in late stance.
v_7	PCs representing the vertical GRF waveform	A high PC value represents a loss of bi-phasic double peak and reduced rate of load transfer at early and late stance.
v_8		Magnitude of the vertical GRF peaks and the rate of load acceptance
v_9	PCs representing the flexion-extension waveform	Magnitude offset of knee flexion during the gait cycle
v_{10}		Range of motion in stance and magnitude of knee flexion during swing phase
v_{11}		Timing of peak knee flexion in swing
v_{12}	PCs representing the abduction-adduction waveform	Magnitude of abduction (high PC) or adduction (low PC) during gait
v_{13}		Magnitude of abduction (low PC) or adduction (high PC) from early to mid-swing
v_{14}		Magnitude of abduction (low PC) or adduction (high PC) at loading response and terminal swing
v_{15}	PC representing the internal –external rotation waveform	Magnitude offset of knee internal-external rotation during gait
v_{16}	Medio-lateral knee width (cm)	Indicator of knee swelling
v_{17}	Anterior–posterior knee depth (cm)	Indicator of knee swelling
v_{18}	Distal thigh girth (cm)	Indicator of muscle mass

Each input variable is transformed into a set of three belief values (i) level of OA knee function ($m(\{OA\})$), (ii) level of NP knee function ($m(\{NP\})$) and (iii) level of uncertainty ($m(\{\Theta\})$) where knee function may be OA or NP. Collectively, these form a characteristic body of evidence (BOE) for each input variable and these are combined using Dempster's rule of combination to produce a final combined BOE (BOE_c) which is represented as a point on a simplex plot, (Fig. 1a). Several sets of motion analysis data can be interpreted simultaneously enabling monitoring of periodic changes of TKR function during recovery.

The simplex plot is divided into four classification regions (Fig. 1b). At the central decision boundary $m(\{NP\}) = m(\{OA\})$. To the left of this, the belief that the subject has NP knee function is greater than the belief they have OA function (i.e. $m(\{NP\}) > m(\{OA\})$). To the right, the belief the subject has OA knee function is greater than the belief they have NP function (i.e. $m(\{OA\}) > m(\{NP\})$). Region 1 indicates dominant NP function where ($m(\{NP\}) > 0.5$). Region 2 indicates dominant OA function where ($m(\{OA\}) > 0.5$). Region 3 indicates non-dominant NP function where ($m(\{OA\}) < m(\{NP\}) < 0.5$) and region 4 shows non-dominant OA function where ($m(\{NP\}) < m(\{OA\}) < 0.5$). A point situated in region 1 indicates knee function more characteristic of NP than a point situated in region 3 and thus the BOE_c for this subject has more association with $m(\{NP\})$. The higher the point is positioned within the simplex plot, the greater the belief in the level of uncertainty ($m(\{\Theta\})$), where the subject exhibits characteristics that may be associated with OA or NP. The closer the point is situated to a vertex, the greater the belief is that the subject has these characteristics.

Variables from Table 1 were used to train the classifier to objectively differentiate between the characteristics of NP and OA. The training sample used to determine classifier control variables consisted of 32 NP and 32 OA knees. The training set was classified with four misclassified OA knees and 93.75% accuracy, determined using Leave-One-Out cross-validation. Once trained, the classifier was used to transform the input variables of the TKR sample into a BOE_c for each assessment. The final classifications of all assessments from a patient are represented on the same simplex plot, displaying how the participant's function deviates from the NP cohort pre-operatively and changes post-TKR.

In order to gain further insight into the differences between OA and NP subjects in terms of age, BMI, % stance, cadence, knee width, knee depth and distal thigh girth, significant p-values were determined using independent t-tests and Mann-Whitney tests.

2.5. Comparing objective and subjective outcome measures

A Pearson correlation coefficient was computed to explore the level of within-subject agreement between the KOS scores, OKS, $m_c(\{NP\})$ and $m_c(\{OA\})$ across multiple visits for the TKR follow-up sample of 20 knees. This was achieved by determining whether correlations existed between the PROMs and classifier outputs from visits 1–4 for each subject separately. Correlation coefficients were also computed to determine if there was a linear relationship between these outcome

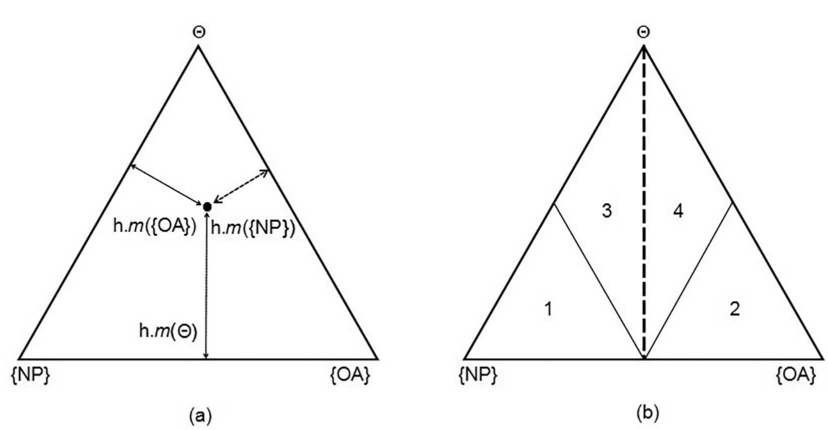

Fig. 1. *(a)* Relationship between the classification belief values and position on the simplex plot illustrating where the subject is characterised based on their objective measures. h is the height of the triangle. *(b)* Regions of dominant (1 and 2) and non-dominant (3 and 4) classification.

measures across all visits of the TKR follow-up sample. Where data failed the test for normality, a Spearman's Rank correlation was calculated.

3. Results

The NP subjects (37.5 ± 14.32 years; BMI 25.5 ± 6.4 kg/m^2) were significantly younger with lower BMI ($p < 0.001$) than the OA subjects (66.5 ± 8.7 years; BMI 30.4 ± 6.0 kg/m^2). The OA cohort used to train the classifier had a mean KOS score of 44.1 ± 14.72%, where 100% is maximum function. The variables in Table 1 were used in combination to classify each participant. As a summary of the NP and OA cohorts in respect to these measures, the OA subjects spent longer in stance (63.6 ± 3.0 vs 61.8 ± 3.2% of the gait cycle; $P = 0.027$) and walked with a slower cadence (47.0 ± 6.1 vs 56.3 ± 3.6 steps per minute; $p < 0.001$). They had a larger ML knee width (11.3 ± 1.2 vs 9.5 ± 1.1 cm; $p < 0.001$), AP knee depth (12.8 ± 1.2 vs 11.5 ± 1.0 cm; $p < 0.001$) and distal thigh girth (44.7 ± 5.2 vs 41.3 ± 4.5 cm; $P = 0.007$).

Group mean kinematics and kinetic waveforms for OA, NP and TKR (at approximately 12 months post-surgery) are illustrated within Fig. 2. In a qualitative comparison of the average kinematic and kinetic waveforms from the two cohorts, it is evident that the PCs used for classification represent key regions where large differences exist. The OA subjects have fixed flexion during stance, lower flexion peaks during stance and swing, exhibit a slower rate of transition to peak knee flexion in swing and to heel strike and approach heel strike with a larger knee flexion angle. This cohort has lower anterior and posterior peak GRFs, do not have clear bi-phasic vertical GRF peaks and an absence of a trough at midstance. They have larger knee adduction and smaller transverse plane rotations.

For the pre- and post-operative assessment times, KOS score, OKS, $m_c(\{OA\})$, $m_c(\{NP\})$ and $m_c(\Theta)$ for the TKR sample (P1 to P20) are listed in the supplementary material where, for each classification, the most influential BOE$_c$ out of $m_c(\{OA\})$ and $m_c(\{NP\})$ are highlighted. Fig. 3 shows simplex plots for each subject, illustrating knee function

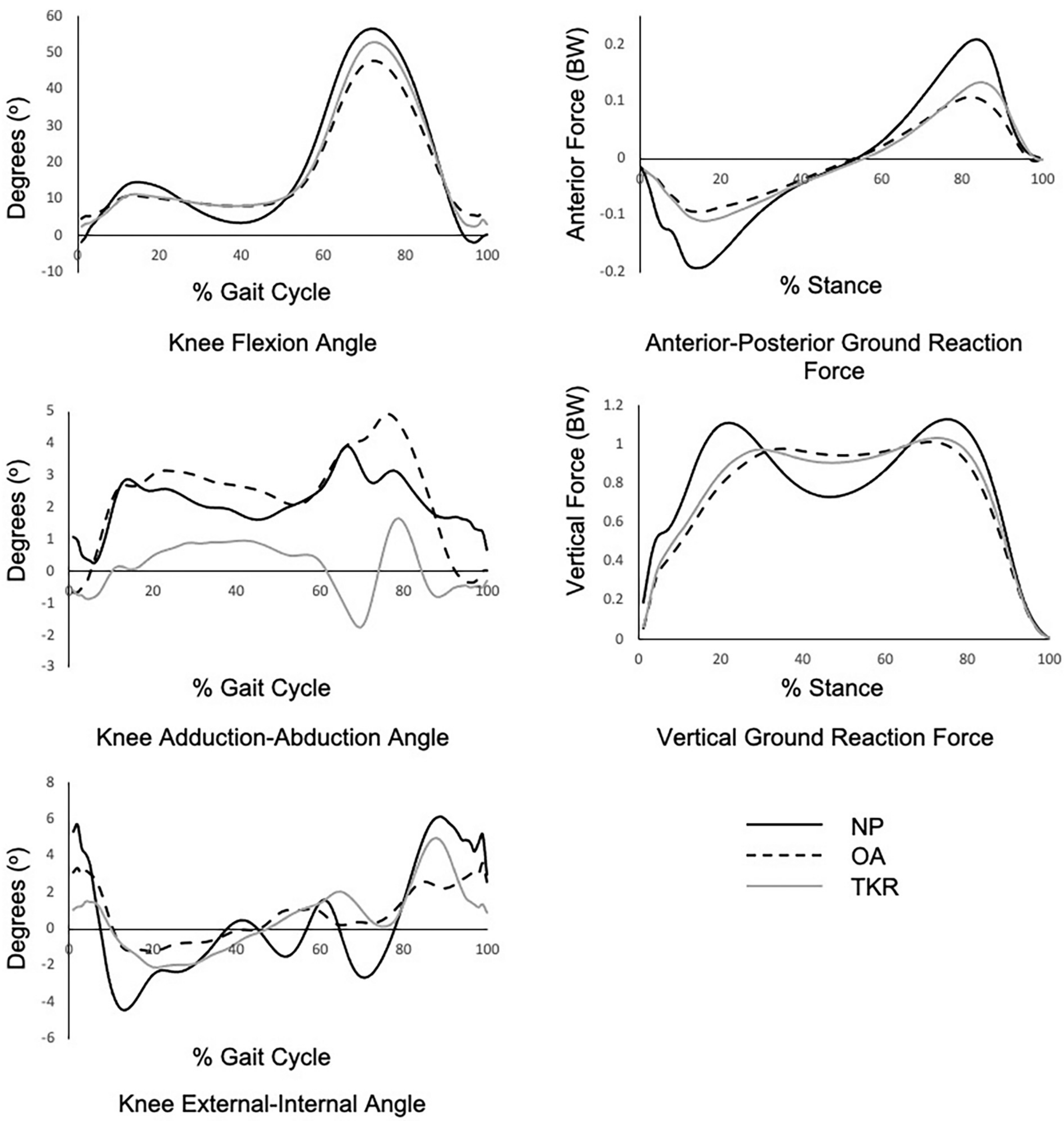

Fig. 2. Mean kinematics and kinetic waveforms for OA, NP and TKR (measured approximately 12 months post-surgery).

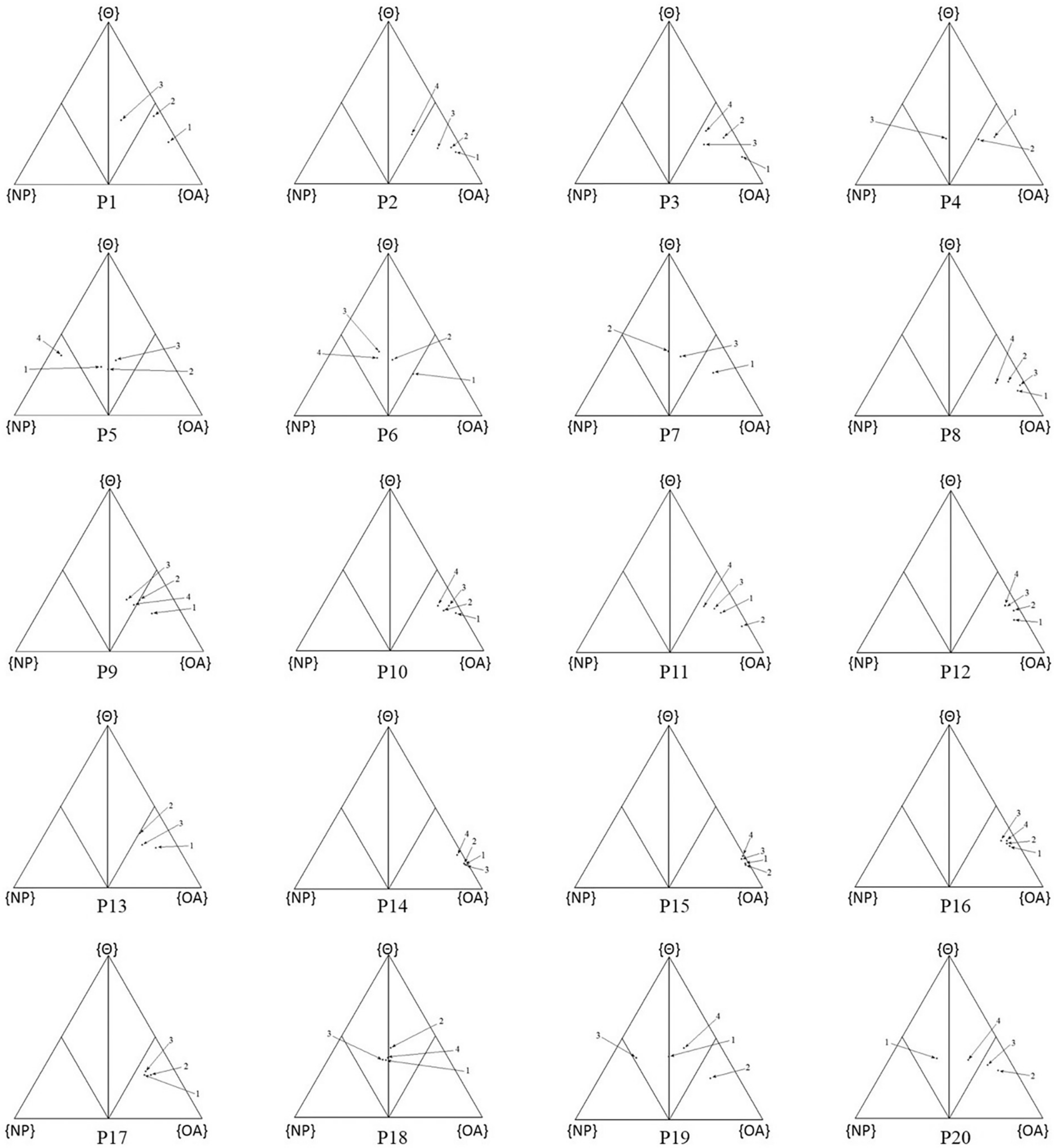

Fig. 3. Simplex plots illustrating the level of recovery towards NP gait following TKR surgery for subjects P1-P20. The numerals indicate visit 1 (pre-operative assessment) and visits 2–4 (post-operative assessments). The plots are displayed in order of greatest to least reduction in m_c({OA}).

classification, plotted as BOE_c indices. To assist with interpretation of the simplex plots, the classification outputs for subjects P6, 14 and 19 representing good, poor and variable recovery respectively, along with their data in Table 2, will be described.

3.1. P6 (good recovery)

Based on the combination of the input variables to the classier, this subject is characterised as being most like patients with knee OA pre-surgery and demonstrates continual improvements in NP characteristics at each post-operative assessment. At visit 1, the OA belief value, m_c({OA}) is the most influential in the classification and since m_c({OA}) = 0.504, P6 is classified on the non-dominant/dominant border. m_c({NP}) increases between visits, suggesting some relief from OA. At visit 3, m_c({NP}) > m_c({OA}) moving P6 into the non-dominant NP region as the subject begins walking with characteristics of the NP training set. Increased uncertainty, $m_c(\Theta)$, moves the classification higher in the simplex plot. Small increases in m_c({NP}) and m_c({OA}) accompanied by reduced uncertainty at visit 4 moves the subject further into and lower in the NP region. There was good agreement (r > ±0.8) between recovery assessed subjectively using PROMs and objectively using classification. The changes in biomechanical signals between visits 1 and 4 mirror the changes in the classification outputs and include a 25% increase in anterior peak GRF, a more prominent double peak vertical GRF with a 6% increased first peak and 12% reduction in midstance, 5% increase in cadence, reduced knee adduction angle and 22% increase in flexion/extension range of motion.

Table 2
Pre- and post-TKR assessment times, KOS score, OKS and classifier outputs ($m_c(\{OA\})$, $m_c(\{NP\})$ and $m_c(\Theta)$) for subjects P6, P15 and P19 representing good, poor and variable recovery.

Subject	Visit	Timing	KOS score	OKS	$m_c(\{OA\})$	$m_c(\{NP\})$	$m_c(\Theta)$
P6 (Good Recovery)	1	2 weeks pre-op	46.25	45.84	0.504	0.238	0.258
	2	3 months post-op	83.75	83.34	0.350	0.305	0.345
	3	6 months post-op	73.75	72.91	0.253	0.351	0.396
	4	12 months post-op	86.25	85.41	0.264	0.380	0.356
P14 (Poor Recovery)	1	10 days pre-op	26.25	16.66	0.834	0.017	0.150
	2	3 months post-op	61.25	62.5	0.823	0.019	0.157
	3	6 months post-op	70	60.41	0.845	0.009	0.145
	4	12 months post-op	81.25	70.84	0.761	0.029	0.210
P19 (Variable Recovery)	1	2 weeks pre-op	68.75	N/A	0.306	0.306	0.388
	2	3 months post-op	66.25	N/A	0.594	0.153	0.253
	3	7 months post-op	80	N/A	0.135	0.484	0.381
	4	11 months post-op	70	N/A	0.359	0.201	0.440

The classification outputs $m_c(\{OA\})$, $m_c(\{NP\})$ and $m_c(\Theta)$ total 1. The closer the value is to 1, the greater the evidence from the input variables that the knee is OA, NP or it is uncertain. The shaded cells indicate the highest belief values demonstrating good, poor and variable recovery for P6, P16 and P19 respectively. N/A, data not available. Classifier outputs are rounded to 3d.p.

3.2. P14 (poor recovery)

Throughout, P14 has dominant OA classification since $m_c(\{OA\}) > 0.5$. Minimal changes between visits indicate poor functional recovery. On inspection of the classification input data, some improvements exist between visit 1 and 4. These include increased cadence, anterior force peak and flexion in swing by 27%, 63% and 5% respectively, decreased ML knee width, AP knee depth and distal knee girth by 6%, 9% and 8% respectively and less knee adduction. The combination of these changes was not sufficient to restore NP functional characteristics. At visit 4, a flat vertical GRF profile was retained with a 4% reduction at first peak and midstance, along with an extended time to reach the second peak, medial peak GRF increased by 14% and there was an absence of extension to a neutral knee position after the first flexion peak in stance. PROMs indicate a large improvement between pre and 3-month post-operative visit, and smaller changes towards 12 months following surgery. Overall, the classifier indicated a different trajectory of improvement, with the only notable small improvement occurring between 6 and 12 months following surgery. The slight decline in the OKS score and increase in $m_c(\{OA\})$ at visit 3 coincided with the subject reporting pain and stiffness.

3.3. P19 (variable recovery)

Initially classified on the decision boundary, the belief in OA rises to >0.5 at visit 2 and the subject moves to the dominant OA region. At visit 3 $m_c(\{NP\}) > m_c(\{OA\})$ moving the subject to the non-dominant NP region, where the participant produces their greatest vertical 1st peak, anterior, posterior and medial peak GRF loading, finishing in the non-dominant OA region at visit 4. These changes strongly correlate with KOS. Review of the input data to the classifier shows variability in gait parameters between visits. It is unsurprising that the subject exhibits overall worse function at visit 4 than 1, with 10% decreased peak knee flexion in swing, reduced flexion in stance, 9% decrease in anterior peak GRF, 34% increased peak medial GRF and no changes in the main features of the vertical GRF.

Six out of the 20 subjects considered (P4, P5, P6, P7, P18 and P19) exhibited NP gait during at least one post-surgery assessment. The remaining subjects have a dominant OA belief value in their classification throughout each post-surgery assessment, meaning these subjects never return to a level of function characteristic of the NP sample. Two subjects (P5 and P20) had NP classification pre-surgery, with another having borderline NP/OA classification (P19). From these three subjects, only P5 had NP classification at their last assessment.

The BOE_c used for each simplex coordinate is listed in the supplementary data, along with the timings of the post-operative assessments.

Tables 3 and 4 display the results of the within-subject agreement between the KOS scores, OKS, $m_c(\{NP\})$ and $m_c(\{OA\})$ across multiple visits. Moderate to strong correlations exist between each questionnaire and classification output measure in most cases, but the strength of the correlation is not consistent across the TKR cohort. Strong correlations exist between the OKS and KOS PROMs, Table 5.

Fig. 4 displays moderate and strong correlations significant at the 0.01 level between the KOS scores, OKS, and classification belief values $m_c(\{NP\})$ and $m_c(\{OA\})$ across all visits of the combined TKR follow-up sample. A negative relationship exists between $m_c(\{OA\})$ and PROMs (OKS and KOS). A positive relationship exists between $m_c(\{NP\})$ and PROMs. The scores from the OKS and KOS were very strongly correlated (Spearman's $r = 0.965$, $p = 0.000$).

4. Discussion

A notable aspect of this study is that correlations existed between objective and subjective measures over multiple timepoints for a high proportion of participants. It is also evident that smaller changes are reported from objective measures than from PROMS, with few patients demonstrating greater NP than OA characteristics during recovery. Including objective measures of OA characteristics within the patient assessment pathway, would provide additional insight into the recovery trajectory, that may have implications when allocating resources to those with poorest recovery.

Recovery was variable in 15 of 20 cases, where the belief of OA function, $m_c(\{OA\})$, did not decrease between each assessment. 6 of these subjects had greater $m_c(\{OA\})$ at their first post-TKR than their pre-TKR assessment. 10 subjects were classified with greater $m_c(\{OA\})$ after an assessment where their $m_c(\{OA\})$ had reduced. Out of the 15 cases with variable recovery, for 7 participants, $m_c(\{OA\})$ had decreased by at least 20% for one or more post-operative visit. Variations in recovery might be due to a range of factors including the recurrence of pain, co-morbidities, prosthesis, termination of physiotherapy or delays due to post-surgery swelling, arthritis or implants in the other limb. Linking these factors to the classifications, although beyond the scope of

Table 3
Correlation (and p values) between KOS and the classification outputs (m_c({OA}) and m_c({NP})) across the pre-post TKR assessments for subjects P1 to P20.

Subject	KOS and m_c({OA})	KOS and m_c({NP})
P1	-0.996 (0.054)	0.500 (0.667) ǂ
P2	-0.708 (0.292)	0.784 (0.216)
P3	-0.911 (0.089)	0.941 (0.059)
P4	-0.885 (0.309)	0.905 (0.280)
P5	-0.039 (0.961)	-0.111 (0.889)
P6	-0.840 (0.160)	0.838 (0.162)
P7	-0.728 (0.481)	0.717 (0.491)
P8	-0.043 (0.957)	-0.166 (0.834)
P9	-0.928 (0.072)	0.936 (0.064)
P10	-0.924 (0.250)	0.742 (0.468)
P11	-0.599 (0.401)	0.685 (0.315)
P12	-0.930 (0.070)	0.800 (0.200) ǂ
P13	-1.000**ǂ	0.500 (0.667) ǂ
P14	-0.548 (0.452)	0.336 (0.664)
P15	-0.073 (0.927)	-0.112 (0.888)
P16	0.047 (0.970)	-0.500 (0.667)
P17	0.640 (0.558)	-0.978 (0.135)
P18	0.263 (0.737)	-0.191 (0.809)
P19	-0.874 (0.126)	0.923 (0.077)
P20	0.567 (0.433)	-0.647 (0.353)

Pearson's Correlation unless indicated by ǂ where Spearman's correlation performed.
**Significant at the 0.01 level (2-tailed).
Shaded cells indicate a correlation where $r > 0.5$ or < 0.5.

this study, are important in future work towards identifying what is limiting the return of NP characteristics for these patients and to be able to predict the expected level of functional recovery for a patient (Sanchez-Santos et al., 2018). Information on likely outcomes could help manage patient expectations, influence satisfaction, and inform clinical decision making (Yap et al., 2021).

In some cases, the minor magnitude change might not exceed the measurement error in determining the classifier belief values. This highlights the importance for research to determine the minimal detectable change, and the level of improvement considered clinically significant for this type of knee OA patient. Exploration in this area by (Bonnefoy-Mazure et al., 2020) identified that patient acceptable symptom state may be more accessible from gait data than minimal clinical important improvement for patients 1-year after TKR.

The level of recovery may also be influenced by the pre-operative condition, suggesting influence of long-standing osteoarthritis, muscle loss and reduced strength. In these cases, the PROMS can show a significant improvement, but this is not mirrored by the biomechanical improvements to the same extent. Biomechanical deficits were apparent in the vast majority of subjects despite improvements in PROMs, in agreement with (Worsley et al., 2016) . The KOS score improved from 41.7 ± 14.9 pre-operatively to 73.6 ± 14.0 at final assessments (9 + months). The OKS score improved from 36.9 ± 13.9 pre-operatively to 74.1 ± 17.8 post-operatively. Judge et al. (2012) found an 11-point magnitude change in OKS predicted patient satisfaction in 95.4% of subjects. The equivalent percentage change would predict satisfaction in 14 of 19 participants using the KOS and 5 of 7 using the OKS. PROMs improved for every subject following surgery. In contrast, three participants did not receive a reduction in m({OA}) at their last visit compared to their first (P18–20). Assuming biomechanical recovery is achieved when m({NP}) > m({OA}), only 6 participants reached this target at some stage in their recovery, with 1 demonstrating high levels of NP

Table 4
Correlation (and p values) between OKS and the classification outputs (m_c({OA}) and m_c({NP})) across the pre-post TKR assessments for 9 subjects.

Subject	OKS and m_c({OA})	OKS and m_c({NP})
P4	-0.925 (0.248)	0.941 (0.219)
P5	0.000 (1.000) ǂ	0.000 (1.000) ǂ
P6	-0.834 (0.166)	0.831 (0.169)
P8	-0.557 (0.443)	0.263 (0.737)
P9	-0.545 (0.633)	0.878 (0.318)
P11	-0.793 (0.207)	0.829 (0.171)
P14	-0.488 (0.512)	0.331 (0.669)
P17	0.822 (0.386)	-0.998 (0.037)*
P20	0.680 (0.320)	-0.750 (0.250)

Pearson's Correlation unless indicated by ǂ where Spearman's correlation performed.
*Significant at the 0.05 level (2-tailed).
Shaded cells indicate a correlation where $r > 0.5$ or < -0.5.

Table 5
Correlation (and p values) between KOS and OKS for 9 subjects.

Subject	KOS and OKS
P4	0.995 (0.061)
P5	0.866 (0.333) ǂ
P6	1.000 (0.000)**
P8	0.672 (0.328)
P9	0.997 (0.052)
P11	0.920 (0.080)
P14	0.974 (0.026)*
P17	0.964 (0.172)
P20	0.985 (0.015)*

Pearson's Correlation unless indicated by ǂ where Spearman's correlation performed.
*Significant at the 0.05 level (2-tailed).
**Significant at the 0.01 level (2-tailed).
Shaded cells indicate a correlation where $r > 0.5$.

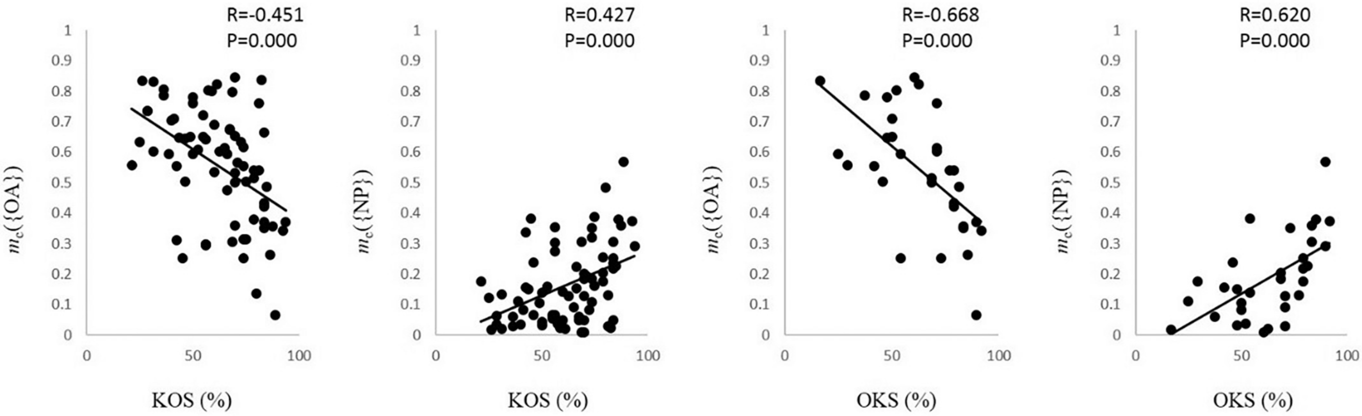

Fig. 4. Relationship between classification belief values m_c({OA}) and m_c({NP}) with KOS (n = 73) and OKS (n = 32) for all subjects and assessment times. Spearman's correlation coefficient displayed where all are significant at the 0.01 level (2-tailed).

function where $m_c(\{NP\}) > 0.5$. Four of these are classified NP at their last assessment. The others, despite improvements, did not exhibit level gait characteristics of an NP subject and this common occurrence agrees with other studies (Andriacchi, 1993; Benedetti et al., 2003; Fuchs et al., 2002; Myles et al., 2002; Whittle and Jefferson, 1989). This study highlights discrepancies between subjective and objective measures of function where larger changes are reported using PROMs in agreement with (Naili et al., 2017).

For each individual throughout their recovery, correlations ($r > 0.5$) were identified between classification output measure m_c({OA}) and KOS for 75% of patients and OKS for 78% of patients (based on 20 and 9 subjects respectively). Where this relationship is not present, the two assessment methods are not monitoring recovery similarly. A limitation of these results is that only 3 or 4 time points were used to determine correlation coefficients.

The classifier outputs m_c({OA}) and m_c({NP}) from all visits of the TKR follow-up sample of 20 subjects correlated moderately with KOS and strongly with OKS. This demonstrates a link between the constructs of knee function assessed by the two techniques. The classifier objectively assesses knee function, whereas the questionnaires subjectively record perceived improvement in knee function during a variety of daily activities, considering clinical parameters such as buckling, instability and have been shown to be related to pain and depression (Maly et al., 2006). There is growing evidence that PROMs do not capture the actual change in biomechanical function following TKR surgery (Mizner et al., 2011). Investigating discrepancies between classification outputs and PROMs would prove useful, especially since around 20% of patients report dissatisfaction with their TKR (Baker et al., 2007).

The transparency of the classifier allows the relationship between input variables and classification output to be ascertained so that changes in OA and NP belief values can be explained in terms of the subject's biomechanical, temporal and anthropometric measures. Objective outputs are determined from the combined evidence of 18 variables. Further work could explore how subjective interpretation of each independent functional measure compares to that of the combined classification. It would also be prudent to explore whether other biomechanical signals such as joint moments and power would enhance the classification (Biggs et al., 2019b; Metcalfe et al., 2017; Worsley et al., 2016). Future work should also explore the level of functional improvement achievable following TKR and develop a threshold that defines a clinically satisfactory biomechanical recovery.

This investigation shows that correlations exist between PROMS and objective measures combining biomechanics, temporal and anthropometric information, when assessing pre to post-TKR recovery across multiple assessments. However, changes in objective biomechanical function are often small and few patients recover function equivalent to a healthy cohort. Similar to the findings of (Mizner et al., 2011; Stratford and Kennedy, 2006; Worsley et al., 2016), deficits in biomechanical recovery were not always captured using PROMs, which provide a snapshot of patient satisfaction whilst also incorporating other factors such as pain and other functional activities. Biomechanical outcomes provide insight into compensation strategies or limitations that may relate to future musculoskeletal complications and thus longer-term outcomes of the TKR.

The classifier is generic (Jones et al., 2006; Jones et al., 2008; Jones and Holt, 2008; Whatling et al., 2008), with the potential to provide insight into what constitutes good or poor recovery and understand the limitations to recovery. This information alongside PROMs, which is more accessible to larger cohorts, could provide powerful information to clinicians and implant manufacturers to inform treatment planning and guide rehabilitation strategies post-surgery to optimize patient mobility.

5. Limitations

The cohorts are not age matched and therefore it may not be realistic to expect patients post-TKR to achieve the function of the NP cohort in this analysis. However, since age is one of the risk factors of OA, this decision was made to avoid subjects from the training body who are at high risk of musculoskeletal conditions affecting biomechanics, given the prevalence of knee OA in asymptomatic adults increases with age (Culvenor et al., 2019). Soft Tissue artefacts are present in studies involving marker-based motion analysis which can introduce errors into the resulting rotations, in particular abduction/adduction and internal/external knee rotations (Stagni et al., 2005).

Declaration of Competing Interest

None.

Acknowledgements

This work was supported by Versus Arthritis grant number 20781 and along with a PhD studentship from DePuy International Ltd. (now DePuy Synthes). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.clinbiomech.2022.105625.

References

Andriacchi, T.P., 1993. Functional analysis of pre and post-knee surgery: total knee arthroplasty and ACL reconstruction. J. Biomech. Eng. 115, 575–581.
Andriacchi, T.P., Galante, J.O., Fermier, R.W., 1982. The influence of total knee-replacement design on walking and stair-climbing. J. Bone Joint Surg. Am. 64, 1328–1335.
Bachmeier, C.J., March, L.M., Cross, M.J., Lapsley, H.M., Tribe, K.L., Courtenay, B.G., Brooks, P.M., Arthritis Cost and Outcome Project Group, 2001. A comparison of outcomes in osteoarthritis patients undergoing total hip and knee replacement surgery. Osteoarthr. Cartil. 9, 137–146.
Baker, P.N., van der Meulen, J.H., Lewsey, J., Gregg, P.J., National Joint Registry for England and Wales, Beard, D., Murray, D., Field, R.E., 2007. The role of pain and function in determining patient satisfaction after total knee replacement: data from the National Joint Registry for England and Wales. J. Bone Joint Surg. 89-B, 893–900. https://doi.org/10.1302/0301-620X.89B7.19091.
Baker, R., McGinley, J.L., Schwartz, M.H., Beynon, S., Rozumalski, A., Graham, H.K., Tirosh, O., 2009. The gait profile score and movement analysis profile. Gait Posture 30, 265–269. https://doi.org/10.1016/J.GAITPOST.2009.05.020.
Barton, G.J., Hawken, M.B., Scott, M.A., Schwartz, M.H., 2012. Movement deviation profile: a measure of distance from normality using a self-organizing neural network. Hum. Mov. Sci. 31, 284–294. https://doi.org/10.1016/j.humov.2010.06.003.
Benedetti, M.G., Catani, F., Bilotta, T.W., Marcacci, M., Mariani, E., Giannini, S., 2003. Muscle activation pattern and gait biomechanics after total knee replacement. Clin. Biomech. (Bristol, Avon) 18, 871–876.
Beynon, M.J., Jones, L., Holt, C.A., 2006. Classification of osteoarthritic and normal knee function using three-dimensional motion analysis and the Dempster-Shafer theory of evidence. IEEE Trans. Syst. Man Cybern. Syst. Hum. 36, 173–186. https://doi.org/10.1109/TSMCA.2006.859098.
Biggs, P.R., Whatling, G.M., Wilson, C., Holt, C.A., 2019a. Correlations between patient-perceived outcome and objectively-measured biomechanical change following Total Knee Replacement. Gait Posture 70, 65–70. https://doi.org/10.1016/J.GAITPOST.2019.02.028.
Biggs, P.R., Whatling, G.M., Wilson, C., Metcalfe, A.J., Holt, C.A., 2019b. Which osteoarthritic gait features recover following total knee replacement surgery? PLoS One 14, e0203417. https://doi.org/10.1371/journal.pone.0203417.
Bonnefoy-Mazure, A., Lubbeke, A., Miozzari, H., Armand, A., Sagawa, Y., Turcot, K., Poncet, A., 2020. Walking speed and maximal knee flexion during gait after Total knee arthroplasty: minimal clinically important improvement is not determinable; patient acceptable symptom state is potentially useful. J. Arthroplast. 35 (10), 2865–2871.
Brandon, S.C.E., Graham, R.B., Almosnino, S., Sadler, E.M., Stevenson, J.M., Deluzio, K. J., 2013. Interpreting principal components in biomechanics: representative extremes and single component reconstruction. J. Electromyogr. Kinesiol. 23, 1304–1310. https://doi.org/10.1016/J.JELEKIN.2013.09.010.
Cappello, A., Cappozzo, A., Della Croce, U., Leardini, A., 1997. Bone position and orientation recronstruction using external markers. In: Three -Dimensional Analysis of Human Locomotion. Wiley, London, pp. 146–171.
Catani, F., Benedetti, M.G., De Felice, R., Buzzi, R., Giannini, S., Aglietti, P., 2003. Mobile and fixed bearing total knee prosthesis functional comparison during stair climbing. Clin. Biomech. (Bristol, Avon) 18, 410–418.
Cimolin, V., Galli, M., 2014. Summary measures for clinical gait analysis: a literature review. Gait Posture 39, 1005–1010.
Culvenor, A.G., Øiestad, B.E., Hart, H.F., Stefanik, J.J., Guermazi, A., Crossley, K.M., 2019. Prevalence of knee osteoarthritis features on magnetic resonance imaging in asymptomatic uninjured adults: a systematic review and meta-analysis. Br. J. Sports Med. 53, 1268–1278. https://doi.org/10.1136/bjsports-2018-099257.
Dawson, J., Fitzpatrick, R., Murray, D., Carr, A., 1998. Questionnaire on the perceptions of patients about total knee replacement. J. Bone Joint Surg. 80, 63–69.
Deluzio, K.J., Wyss, U.P., Zee, B., Costigan, P.A., Serbie, C., 1997. Principal component models of knee kinematics and kinetics: normal vs. pathological gait patterns. Hum. Mov. Sci. 16, 201–217. https://doi.org/10.1016/S0167-9457(96)00051-6.
Fuchs, S., Flören, M., Skwara, A., Tibesku, C.O., 2002. Quantitative gait analysis in unconstrained total knee arthroplasty patients. International journal of rehabilitation research. Internationale Zeitschrift fur Rehabilitationsforschung. Rev. Int. Recherches Readapt. 25, 65–70.
Gao, A., Beard, D.J., Price, A.J., 2017. What exactly is a good outcome for TKR? And what proportion of patients experience one? Osteoarthr. Cartil. 25, S357. https://doi.org/10.1016/J.JOCA.2017.02.608.
Holt, C., Deursen, R.W.M., Hayes, N.J., O'Callaghan, P.T., 2000. Three-dimensional analysis of the tibiofemoral joint using external marker clusters and the JCS approach: a comparison of normal and osteoarhtiric knee function. In: Computer Methods in Biomechanics and Biomedical Engineering. Gordon and Breach Science Publishers SA, New York, pp. 289–294.
Irrgang, J.J., Snyder-Mackler, L., Wainner, R.S., Fu, F.H., Harner, C.D., 1998. Development of a patient-reported measure of function of the knee. J. Bone Joint Surg. Am. 80, 1132–1145.
Jacobs, N.A., Skorecki, J., Charnley, J., 1972. Analysis of the vertical component of force in normal and pathological gait. J. Biomech. 5, 11–34. https://doi.org/10.1016/0021-9290(72)90016-4.
Jones, L., Holt, C.A., 2008. An objective tool for assessing the outcome of total knee replacement surgery. Proc. Inst. Mech. Eng. H J. Eng. Med. 222, 647–655. https://doi.org/10.1243/09544119JEIM316.
Jones, L., Beynon, M.J., Holt, C.A., Roy, S., 2006. An application of the Dempster–Shafer theory of evidence to the classification of knee function and detection of improvement due to total knee replacement surgery. J. Biomech. 39, 2512–2520.
Jones, L., Holt, C.A., Beynon, M.J., 2008. Reduction, classification and ranking of motion analysis data: an application to osteoarthritic and normal knee function data. Comp. Methods Biomech. Biomed. Eng. 11, 31–40. https://doi.org/10.1080/10255840701550956.
Judge, A., Arden, N.K., Kiran, A., Price, A., Javaid, M.K., Beard, D., Murray, D., Field, R. E., 2012. Interpretation of patient-reported outcomes for hip and knee replacement surgery. J. Bone Joint Surg. 94-B, 412–418. https://doi.org/10.1302/0301-620X.94B3.27425.
Maly, M.R., Costigan, P.A., Olney, S.J., 2006. Determinants of self-report outcome measures in people with knee osteoarthritis. Arch. Phys. Med. Rehabil. 87. https://doi.org/10.1016/j.apmr.2005.08.110.
McClelland, J.A., Webster, K.E., Feller, J.A., 2007. Gait analysis of patients following total knee replacement: a systematic review. Knee 14, 253–263. https://doi.org/10.1016/J.KNEE.2007.04.003.
Metcalfe, A.J., Stewart, C.J., Postans, N.J., Biggs, P.R., Whatling, G.M., Holt, C.A., Roberts, A.P., 2017. Abnormal loading and functional deficits are present in both limbs before and after unilateral knee arthroplasty. Gait Posture 55, 109–115. https://doi.org/10.1016/J.GAITPOST.2017.04.008.
Milner, C.E., 2009. Is gait normal after total knee arthroplasty? Systematic review of the literature. J. Orthop. Sci. 14, 114–120. https://doi.org/10.1007/s00776-008-1285-8.
Mizner, R.L., Petterson, S.C., Clements, K.E., Zeni, J.A., Irrgang, J.J., Snyder-Mackler, L., Snyder-Mackler, L., 2011. Measuring functional improvement after total knee arthroplasty requires both performance-based and patient-report assessments: a longitudinal analysis of outcomes. J. Arthroplast. 26, 728–737. https://doi.org/10.1016/j.arth.2010.06.004.
Murray, D.W., Frost, S.J., 1998. Pain in the assessment of total knee replacement. J. Bone Joint Surg. 80, 426–431.
Myles, C.M., Rowe, P.J., Walker, C.R.C., Nutton, R.W., 2002. Knee joint functional range of movement prior to and following total knee arthroplasty measured using flexible electrogoniometry. Gait Posture 16, 46–54.
Naili, J.E., Iversen, M.D., Esbjörnsson, A.-C., Hedström, M., Schwartz, M.H., Häger, C.K., Broström, E.W., 2017. Deficits in functional performance and gait one year after total knee arthroplasty despite improved self-reported function. Knee Surg. Sports Traumatol. Arthrosc. 25, 3378–3386. https://doi.org/10.1007/s00167-016-4234-7.
Rahman, J., Tang, Q., Monda, M., Miles, J., McCarthy, I., 2015. Gait assessment as a functional outcome measure in total knee arthroplasty: a cross-sectional study. BMC Musculoskelet. Disord. 16, 1, 16, 1–9. https://doi.org/10.1186/S12891-015-0525-2.
Sanchez-Santos, M.T., Garriga, C., Judge, A., Batra, R.N., Price, A.J., Liddle, A.D., Javaid, M.K., Cooper, C., Murray, D.W., Arden, N.K., 2018. Development and validation of a clinical prediction model for patient-reported pain and function after primary total knee replacement surgery OPEN. Sci. Rep. 8, 3381. https://doi.org/10.1038/s41598-018-21714-1.
Schutte, L.M., Narayanan, U., Stout, J.L., Selber, P., Gage, J.R., Schwartz, M.H., 2000. An index for quantifying deviations from normal gait. Gait Posture 11, 25–31. https://doi.org/10.1016/S0966-6362(99)00047-8.
Schwartz, M.H., Rozumalski, A., 2008. The gait deviation index: a new comprehensive index of gait pathology. Gait Posture 28, 351–357. https://doi.org/10.1016/J.GAITPOST.2008.05.001.
Shenoy, R., Pastides, P.S., Nathwani, D., 2013. (iii) Biomechanics of the knee and TKR. Orthopaed. Trauma 27, 364–371. https://doi.org/10.1016/J.MPORTH.2013.10.003.
Stagni, R., Fantozzi, S., Cappello, A., Leardini, A., 2005. Quantification of soft tissue artefact in motion analysis by combining 3D fluoroscopy and stereophotogrammetry: a study on two subjects. Clin. Biomech. 20, 320–329. https://doi.org/10.1016/J.CLINBIOMECH.2004.11.012.
Stratford, P.W., Kennedy, D.M., 2006. Performance measures were necessary to obtain a complete picture of osteoarthritic patients. J. Clin. Epidemiol. 59, 160–167. https://doi.org/10.1016/j.jclinepi.2005.07.012.
Whatling, G.M., Dabke, H.V., Holt, C.A., Jones, L., Madete, J., Alderman, P.M., Roberts, P., 2008. Objective functional assessment of total hip arthroplasty following two common surgical approaches: the posterior and direct lateral approaches. Proc. Inst. Mech. Eng. H J. Eng. Med. 222, 897–905. https://doi.org/10.1243/09544119JEIM396.
Whittle, M.W., Jefferson, R.J., 1989. Functional Biomechanical Assessment of the Oxford Meniscal Knee. J Arthroplasty, pp. 231–243. https://doi.org/10.1016/S0883-5403(89)80019-1. https://www.arthroplastyjournal.org/article/S0883-5403(89)80019-1/pdf.
Williams, D., Metcalfe, A., Madete, J., Whatling, G., Kempshall, P., Forster, M., Lyons, K., Holt, C., 2020. The relationship between alignment, function and loading in total knee replacement: in-vivo analysis of a unique patient population. J. Biomech. 112, 110042 https://doi.org/10.1016/J.JBIOMECH.2020.110042.
Worsley, P.R., Whatling, G., Barrett, D., Holt, C., Stokes, M., Taylor, M., 2016. Assessing changes in subjective and objective function from pre- to post-knee arthroplasty using the Cardiff Dempster–Shafer theory classifier. Comp. Methods Biomech. Biomed. Eng. 19, 418–427. https://doi.org/10.1080/10255842.2015.1034115.
Yap, Y.Y.W., Edwards, K.L., Soutakbar, H., Fernandes, G.S., Scammell, B.E., 2021. Oxford knee score 1 year after TKR for osteoarthritis with reference to a normative population: what can patients expect? Osteoarthrit. Cartil. Open 3, 100143. https://doi.org/10.1016/J.OCARTO.2021.100143.