Gait & Posture xxx (2018) xxx-xxx

Contents lists available at ScienceDirect

Gait & Posture

journal homepage: www.elsevier.com

Pre-operative sensor-based gait parameters predict functional outcome after total knee arthroplasty

Felix Kluge[a, *], Julius Hannink[a], Cristian Pasluosta[b], Jochen Klucken[c], Heiko Gaßner[c], Kolja Gelse[d], Bjoern M. Eskofier[a], Sebastian Krinner[d]

[a] Machine Learning and Data Analytics Lab, Department of Computer Science, Friedrich-Alexander University Erlangen-Nürnberg (FAU), Carl-Thiersch-Str. 2b, 91052 Erlangen, Germany
[b] Laboratory for Biomedical Microtechnology, Department of Microsystems Engineering, University of Freiburg, 79110 Freiburg, Germany
[c] Molecular Neurology, University Hospital Erlangen, Schwabachanlage 6, 91054 Erlangen, Germany
[d] Department of Trauma Surgery, University Hospital Erlangen, Krankenhausstrasse 12, 91054 Erlangen, Germany

ARTICLE INFO

Keywords:
Knee replacement surgery
Osteoarthritis
Inertial measurement unit
Accelerometer
Classification
Mobile gait analysis

ABSTRACT

Background: Despite the general success of total knee arthroplasty (TKA) regarding patient-reported outcome measures, studies investigating gait function have shown diverse functional outcomes. Mobile sensor-based systems have recently been employed for accurate clinical gait assessments, as they allow a better integration of gait analysis into clinical routines as compared to laboratory based systems.
Research question: In this study, we sought to examine whether an accurate assessment of gait function of knee osteoarthritis patients with respect to surgery outcome evaluation after TKA using a mobile sensor-based gait analysis system is possible.
Methods: A foot-worn sensor-based system was used to assess spatio-temporal gait parameters of 24 knee osteoarthritis patients one day before and one year after TKA, and in comparison to matched control participants. Patients were clustered into positive and negative responder groups using a heuristic approach regarding improvements in gait function. Machine learning was used to predict surgery outcome based on pre-operative gait parameters.
Results: Gait function differed significantly between controls and patients. Patient-reported outcome measures improved significantly after surgery, but no significant global gait parameter difference was observed between pre- and post-operative status. However, the responder groups could be correctly predicted with an accuracy of up to 89% using pre-operative gait parameters. Patients exhibiting high pre-operative gait function were more likely to experience a functional decrease after surgery. Important gait parameters for the discrimination were stride time and stride length.
Significance: The early identification of post-surgical functional outcomes of patients is of great importance to better inform patients pre-operatively regarding surgery success and to improve post-surgical management.

1. Introduction

Osteoarthritis (OA) is a progressive disease mainly occurring in the latter half of life and is characterized by the wear, softening and thinning of articular cartilage [1]. Symptomatic knee OA exhibits a prevalence of 12% among US adults and symptoms not only include pain, but also severe functional limitations, including gait dysfunction and a reduction of other activities of daily living [2]. With 1.6% of all US adults undergoing knee replacement surgery, total knee arthroplasty (TKA) is the most common invasive intervention for end-stage knee OA patients [3]. Surgery outcome is usually assessed by patient-reported outcome measures (PROMs), which include the assessment of pain, functional measures, and satisfaction based on questionnaires [4]. PROMs, however, only yield a global perspective on TKA outcomes, are rather subjective and do not fully capture functional limitations [5]. It has been shown, that performance-based measures such as gait parameters obtained from instrumented gait analysis contain complemen-

* Corresponding author.
Email addresses: felix.kluge@fau.de (F. Kluge); julius.hannink@fau.de (J. Hannink); cristian.pasluosta@imtek.uni-freiburg.de (C. Pasluosta); jochen.klucken@uk-erlangen.de (J. Klucken); heiko.gassner@uk-erlangen.de (H. Gaßner); kolja.gelse@uk-erlangen.de (K. Gelse); bjoern.eskofier@fau.de (B.M. Eskofier); sebastian.krinner@uk-erlangen.de (S. Krinner)

https://doi.org/10.1016/j.gaitpost.2018.08.026
Received 15 April 2018; Received in revised form 20 August 2018; Accepted 22 August 2018
Available online xxx
0966-6362/

tary information to traditionally used PROMs and that gait function also contributes to satisfaction after surgery [4,6].

Gait biomechanics of OA and changes due to TKA have been investigated in numerous studies and reviews found mixed results on whether gait biomechanics (i.e. kinematic and kinetic parameters) improved after surgery [7–9]. Furthermore, traditionally used laboratory systems such as infrared cinematography for biomechanical assessment are stationary, expensive and require trained personnel, such that they are not routinely used in clinical assessment. Changes in gait function can also be assessed using spatio-temporal gait parameters, with the benefit of allowing the use of various motion capture systems that are able to extract those parameters. Regarding positive TKA outcomes, improved spatio-temporal gait parameters (i.e. higher gait speed, cadence, stride length) have partially been observed after surgery [4,6,10–13].

Sensor-based systems including inertial measurement units have been increasingly used for more mobile, cost efficient and clinically feasible gait assessments in various diseases [14–16]. For example, they allow a fast assessment of gait function changes using spatio-temporal gait parameters, exhibiting higher sensitivities as compared to PROMs in knee OA [17]. The discrimination of pathologic and healthy gait using mobile sensor systems has been demonstrated before [18–22] and also the effect of TKA has been evaluated [6,12], mostly using pelvis-worn sensors.

Despite the general success of TKA regarding PROMs, some patients may experience adverse effects with persistent or worsened functional limitations [23]. By investigating a diverse post-operative patient group, Berliner et al. predicted functional gait improvements based on pre-operative PROMs [23]. Patients with high self-reported pre-operative function were less likely to experience a clinically meaningful functional improvement after TKA. In a different study, biomechanical parameters have been used to predict post-operative abnormal knee joint loading patterns and severity of post-surgery anterior knee pain [24]. Highly accurate predictions of the response to exercise interventions in patients with mild to moderate knee OA has previously been performed using three-dimensional lower limb kinematics or inertial data of gait as well as PROMs [25,26]. Bolink et al. demonstrated that spatio-temporal gait parameters provide complementary information for the discrimination between pre- and post-operative function [6]. But it still remains open whether those parameters can also be used to predict functional surgery outcome on an objective basis.

Therefore, the goals of this study were to (a) demonstrate the feasibility of characterizing pre- and post-operative OA gait in comparison to healthy gait using a mobile foot-worn sensor-based system and to (b) predict post-operative gait function using pre-operative spatio-temporal gait parameters. As mobile sensor-based systems have the potential to assess patients' gait on a large scale without the impediments of laboratory based motion capture setups, they might be more feasible in clinical routine use for intervention assessment and outcome prediction.

2. Methods

2.1. Participants

A total of 24 patients with end-stage unilateral knee OA (mean ± stdev, 8 males, 16 females; age: 64.0 ± 11.0 years; mass: 90.8 ± 19.4 kg; height: 170.6 ± 10.4 cm; BMI: 31.3 ± 6.8 kg m^{-2}) were recruited at the University Hospital Erlangen and underwent TKA with the endoprosthesis "BPK-S Integration" (Brehm, Weisendorf, Germany). Patients with moderate-to-severe knee OA elected for knee replacement surgery were included in the study (mean Kellgren-Lawrence grade of 3.4 ± 0.7). Exclusion criteria were bilateral joint dysfunction and other pathologies potentially interfering with the gait pattern, such as neurological diseases. Gait assessment was performed on average 1.5 ± 0.5 days before and 48.3 ± 11.2 weeks after surgery. Six patients dropped out after the first assessment. Age and sex matched participants were identified for gait parameter comparison from a previously recruited population of healthy participants with no self-reported history of neuro-muscular diseases and joint deterioration of the lower limbs (mean ± stdev, 8 males, 16 females; age: 62.3 ± 9.7 years; mass: 70.7 ± 11.6 kg; height: 168.4 ± 6.2 cm; BMI: 24.9 ± 4.1 kg m^{-2}). The study was approved by the local ethical committee (Ethical approval Re.-No. 181_12 B, Ethics Committee of the Faculty of Medicine, Friedrich-Alexander-University Erlangen-Nürnberg, Germany) and written informed consent was obtained from all study participants before participation.

2.2. Instrumentation

Two Shimmer3 sensors (Shimmer, Dublin, Ireland) were laterally attached to each shoe using rigid sensor mounts to ensure the same sensor positions for each participant. Each sensor contained a three-axis accelerometer (range: ±8 g) and a three-axis gyroscope (range: ±500° s^{-1}) sampling at a rate of 102.4 Hz. The data were transferred via Bluetooth to a mobile device for storage. We detected single strides in the continuous inertial data stream using the multi-dimensional sub sequence dynamic time warping approach (msDTW) which nonlinearly matches time series of different length to a pre-defined stride template [27]. Then, the gait events heel strike (HS) and toe off (TO) were detected for each stride [16]. The inertial measurements from the local frame of measurement of the shoe were transformed into the global coordinate frame and gravity was removed. Finally, the feet's trajectories were calculated using double integration by also accounting for drift effects [28]. The trajectories, orientations, and gait events were used to extract the parameters gait speed, stride length, stance time, swing time, HS angle, TO angle and maximal toe clearance in Matlab R2016b (MathWorks Inc., Natick, MA, USA). The whole system has previously been described in more detail and has been validated with healthy and affected gait patterns [16,27–29]. All participants wore the same shoe model (Adidas Duramo 6, Herzogenaurach, Germany) to reduce potential gait differences arising from wearing differing footwear [30].

2.3. Protocol

PROMs regarding gait and other functional activities as well as quality of life were assessed using the Western Ontario & McMaster Universities Osteoarthritis Index (WOMAC), Oxford Knee Score (OKS), Knee Society Score (KSS, parts 1 and 2), general health status (EQ-5D-3L), and WHO Disability Assessment Schedule 2.0 (WHODAS 2.0, short version) scores. Gait parameters were assessed using a standardized 4 × 10 m overground walking test, which has been suggested as a preferred gait test for knee OA [31]. From initially standing, the patient walked to a mark 10 m away. This distance was covered four times back and forth with 180° turnings (clockwise and counter-clockwise alternately) around marks on the ground.

2.4. Data analysis

Automatic stride segmentation [27] was manually checked to eliminate potentially falsely detected strides. Gait parameters were extracted only for straight walking sequences. A total of 1105 pre-operative, 912 post-operative strides, and 936 strides from healthy participants (average number of strides per participant: 46.0 (pre), 50.7 (post), 39.0 (control)) were used in the analysis. Participant and foot-wise outlier removal was performed based on the median absolute deviation (MAD) around the median [32]. PROMs and gait parameters were assessed for group differences using univariate inferential statistics. As data normality was not generally assured, paired Mann-Whitney u-tests with a priori significance levels α of 0.05 were applied for pairwise comparisons. The coefficient of variation (CV) was considered over all strides for each performed 4×10 m test and each participant separately as a measure of stride-to-stride variability.

2.5. Outcome prediction

We defined a performance-based dichotomization into "positively" and "negatively" responding patients based on the changes of spatio-temporal gait parameters. First, the difference between pre- and post-operative gait parameters was calculated to determine gait improvements. As single strides between post- and pre- session cannot be directly related, the difference of each single post-operative stride x_{post} to the mean parameter of the pre-operative session was calculated:

$$\Delta x_{i,k} = x_{\mathrm{post},i,k} - \frac{1}{n_{\mathrm{pre},k}} \sum_{j}^{n_{\mathrm{pre},k}} x_{\mathrm{pre},j,k} \tag{1}$$

with i being the post-operative stride index, j the pre-operative stride index, k the patient index and $n_{pre,k}$ the number of strides per patient pre-operatively.

Based on the parameter differences, a median split was performed to dichotomize the stride parameters. The definition of "positive" and "negative" was heuristically defined for every gait parameter. Increases in gait speed, stride length, swing time, HS angle and maximum toe clearance were defined as improvements (positive response). Decreases in stance time, stride time and TO angle were equivalently defined as improvements. Accordingly, decreases in gait speed, stride length, swing time, HS angle and maximum toe clearance and increases in stance time, stride time and TO angle were defined as negative responses. By definition in our employed system, higher negative TO angles indicate steeper foot to ground angles.

It was thus possible to label each post-operative single stride as a positive or negative response. Subsequently, the patients were clustered into the associated positive or negative responder class by assigning each patient the label of the most frequent stride type that he exhibited in the respective post-operative 4×10 m test. Each patient could thus be assigned to either positive or negative responder group using the heuristic clustering.

We then predicted this group membership based only on pre-operative gait parameters. Every single pre-operative stride was input into a classification model to predict an outcome (either positive or negative response). Due to variability in the performed strides of the 4×10 m test, both positive and negative outcomes might have been predicted for a patient. To assign a patient a uniform response label, a majority vote was performed on those single stride predictions. This assigned the patient the final class label based on the strides with the highest occurrence frequency. Finally, it was evaluated whether this prediction was correct by comparing the predicted patient label with the label of the heuristic reference clustering. The CVs were not used for prediction, as they represent only summary measures over the whole gait test.

Logistic regression (LR), decision trees (DT), k-nearest neighbors (kNN), Adaboost, and linear support vector machines (SVM) were used to find the best classification scheme. Classification metrics were calculated based on the majority voting results, as the final patient assignment to the respective response cluster was of relevance in this study. Generalizability of the models was evaluated using a leave-one-subject-out cross-validation, in which one patient is left out once as a test observation while the remaining patients are used as training observations. This is repeated until all patients have been used as test subject once. Model parameter selection for each classification algorithm was performed. The performance of the optimal parameter set was evaluated using an inner cross-validation in each of the 18 leave-one-subject-out cross-validation folds. The overall performance was evaluated using accuracy, confidence intervals, sensitivity, specificity, and area under the curve (AUC) of the receiver operating characteristic (ROC) analysis [33].

3. Results

All gait parameters of the healthy population were significantly different from the patient population except for the CVs of some parameters (Table 1). After TKA, all PROMs showed significant improvements (Table 2). Contrarily, gait performance as quantified by stopwatch measures did not improve after surgery. The time to complete the 4×10 m test remained unchanged (pre: 46.0 ± 18.2 s, post: 44.9 ± 12.4 s, $p = 0.70$). Equivalently, no gait parameter changed significantly after surgery (Table 1). Fig. 1 depicts the distribution of the gait parameter differences $\Delta x_{i,k}$ according to the positive and negative response classes. The median of the gait parameter differences was approximately zero for all parameters (Fig. 1 and Table 3), indicating an equal distribution of positive and negative responders according to the heuristic response definition. Eight patients were clustered as positive responders and ten patients were clustered as negative responders by dichotomization. All gait parameters showed significant differences between both responder groups (Table 3). When predicting those re-

Table 1
Spatio-temporal gait parameters for "Control" participants and patients before ("Pre") and after ("Post") surgery. Given are the mean and standard deviation (SD) for all gait parameters. p values are stated for the pairwise comparisons "control participants vs. pre-operative patients" (p_{c-pre}), "pre- vs. post-operative patients" ($p_{pre-post}$), and "control participants vs. post-operative patients" (p_{c-post}).

Parameter	Control		Pre		p_{c-pre}	Post		$p_{pre-post}$	p_{c-post}
	Mean	SD	Mean	SD		Mean	SD		
Gait speed [m/s]	1.38	0.18	1.06	0.24	< 0.001	1.10	0.25	0.98	< 0.001
Stride time [s]	1.02	0.06	1.14	0.09	< 0.001	1.11	0.09	0.36	< 0.001
Swing time [%]	36.60	1.45	33.59	3.10	< 0.001	34.50	2.36	0.12	< 0.001
Stance time [%]	63.40	1.45	66.41	3.10	< 0.001	65.50	2.36	0.12	< 0.001
Stride length [cm]	140.06	15.11	118.98	23.33	< 0.001	121.04	23.18	0.62	< 0.001
Max. toe clear. [cm]	7.96	2.80	5.47	1.86	< 0.001	5.33	2.12	0.42	< 0.001
TO angle [°]	−66.20	5.62	−56.55	12.11	< 0.001	−57.47	14.18	0.43	0.001
HS angle [°]	19.68	5.66	15.59	5.91	< 0.001	16.24	6.14	0.26	0.02
Gait speed CV [%]	7.76	1.76	7.32	2.49	0.34	6.82	2.49	0.41	0.21
Stride time CV [%]	3.11	1.05	3.40	1.81	0.36	3.27	1.90	0.15	0.53
Swing time CV [%]	3.18	1.14	7.17	7.77	< 0.001	5.12	4.51	0.29	0.01
Stance time CV [%]	1.86	0.75	3.41	3.37	< 0.001	2.66	2.32	0.28	0.02
Stride length CV [%]	6.90	1.93	6.40	2.45	0.33	5.87	2.43	0.06	0.26

Table 2
Overview over pre- and post-operative PROMs. Mean values and standard deviation (SD), *p* value, and effect size (ES).

PROM	Pre		Post		*p*	ES
	Mean	SD	Mean	SD		
WOMAC	59.1	14.8	21.5	20.0	< 0.001	1.4
OKS	18.8	6.2	36.1	9.2	< 0.001	1.9
KSS 1	41.3	12.2	84.6	14.0	< 0.001	2.7
KSS 2	58.1	10.6	81.7	19.3	< 0.001	1.8
EQ.5D.3L	54.1	7.1	79.4	17.8	0.002	1.3
WHODAS	12.3	5.2	4.8	3.9	< 0.001	2.0

sponse outcomes using pre-operative gait parameters, all classifiers achieved accuracies higher than chance level, which was 66.7% for this small data set [34]. The highest accuracy 89% was achieved by a decision tree (Table 4), which falsely classified the responses of two positively responding patients which were misclassified as negative responders (sensitivity = 0.75). All negatively responding patients were correctly identified (specificity = 1.00). All other classifiers misclassified three patients, from which also one negative responder was labeled as a positive responder. For the best performing classifier, the final decision tree model was trained on the whole data set. Only stride time and stride length contributed to the model (Fig. 2a). Short stride times and long strides predicted patients to respond negatively. The misclassifications in the decision tree mainly occurred due to stride time (Fig. 2b).

4. Discussion

Currently, the main outcome measures for TKA assessment have been PROMs including pain and self-reported function. The motivation for this study was to explore the potential of a mobile gait analysis system (a) to objectively characterize gait of end-stage knee OA patients undergoing TKA as compared to healthy participants and (b) to predict functional outcome after TKA based on spatio-temporal gait parameters.

The employed system comprising of two foot-worn inertial measurement units was able to discriminate between gait of healthy participants and end-stage knee OA patients, which is a necessary requirement for treatment assessment [35]. The differences were comparable to those in literature using various other systems [4,19,36]. However, TKA had no significant effect on gait performance in terms of gait parameters. This is similar to previous studies, which have shown diverse results on gait changes after TKA, so that no global gait improvement after TKA could a priori be assumed [4,6,10–13]. The significant improvement of all PROMs is in line with literature, which has shown improved post-operative PROMs [6,12], but it also indicates a discrepancy between PROMs and gait parameters. PROMs are important parameters in clinical evaluation. Although they remain subjective, they might relate better to quality of life and subjective well-being than

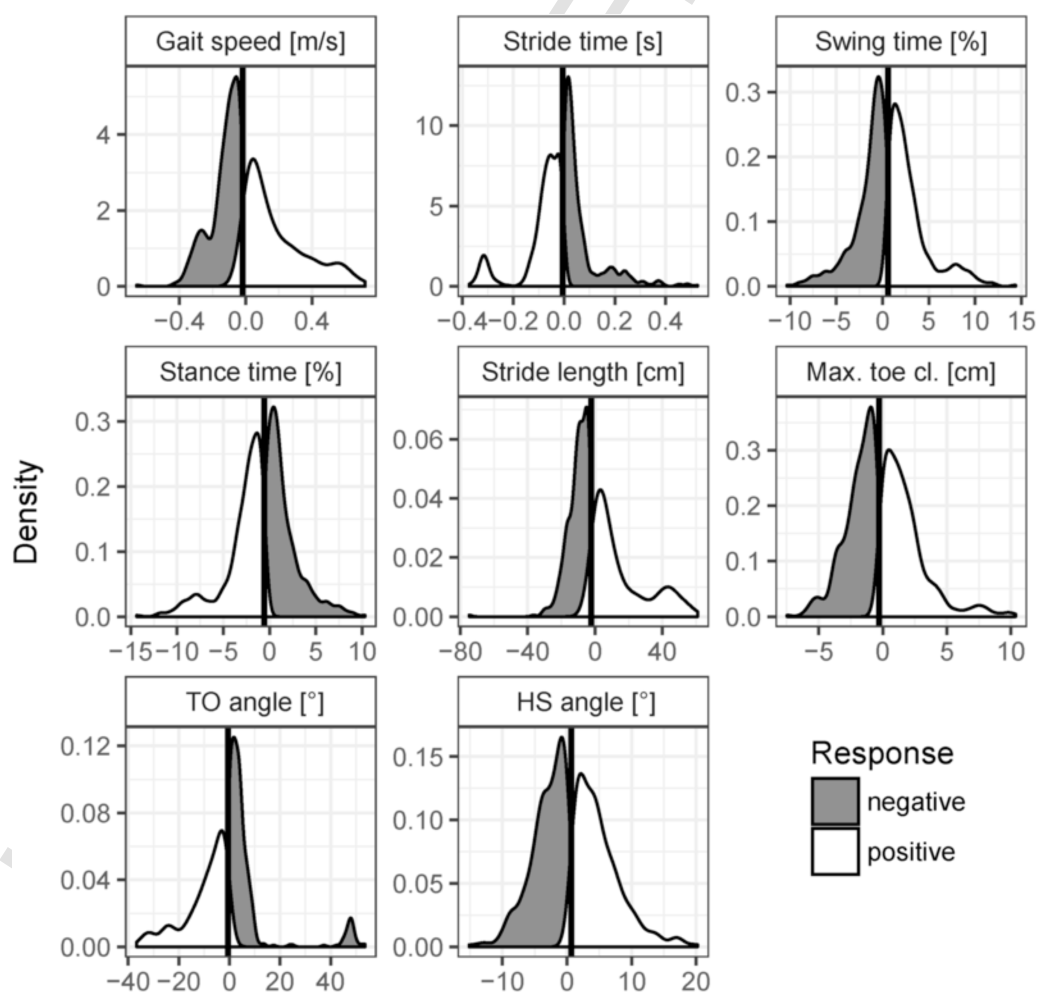

Fig. 1. Dichotomized gait parameters based on the single stride differences. The black lines indicate the medians of the gait parameters that were used to split the patient group into "positive" and "negative" responders.

Table 3
Gait parameter changes due to surgery for "positive" and "negative" responders. Given are the mean and standard deviation (SD), *p* value, effect size (ES), and the median which splits the two groups.

Parameter	Positive		Negative		*p*	ES	Median
	Mean	SD	Mean	SD			
Gait speed [m/s]	0.19	0.21	−0.10	0.12	< 0.001	1.79	−0.02
Stride time [s]	−0.08	0.11	0.04	0.10	< 0.001	1.13	−0.01
Swing time [%]	2.62	3.27	−0.72	2.21	< 0.001	1.22	0.60
Stance time [%]	−2.62	3.27	0.72	2.21	< 0.001	1.22	−0.60
Stride length [cm]	15.39	18.79	−7.86	8.41	< 0.001	1.67	−2.15
Max toe clearance [cm]	0.65	2.09	−0.57	2.65	< 0.001	0.50	−0.27
TO angle [°]	−9.97	9.89	5.15	12.40	< 0.001	1.33	−0.54
HS angle [°]	2.56	5.78	−0.26	4.39	< 0.001	0.56	0.63

Table 4
Classification results of the outcome prediction (classes: "positive" and "negative" responders) using pre-operative gait parameters (CI = confidence interval, AUC = area under the curve is based on the single stride prediction ROC curve). The chance level for this small data set is 66.7% [34]. Sensitivity (true positive rate) is the probability of correctly detecting positive responders. Specificity (true negative rate) is the probability of correctly identifying negative responders.

Classifier	Accuracy (CI)	Sensitivity	Specificity	AUC
Decision tree	0.89 (0.65, 0.99)	0.75	1.00	0.76
kNN	0.83 (0.59, 0.96)	0.75	0.90	0.80
AdaBoost	0.83 (0.59, 0.96)	0.75	0.90	0.77
Logistic regression	0.83 (0.59, 0.96)	0.75	0.90	0.77
SVM (linear)	0.83 (0.59, 0.96)	0.75	0.90	0.77

other parameters. However, the assessment of gait parameters introduces objectivity in the disease and rehabilitation assessment (further studies regarding clinical relevance need however to be performed). Therefore, gait parameters currently have to be regarded as complementary parameters. Nevertheless, the importance of including gait parameters as objective measures also in clinical follow-ups has already been highlighted [6,14,18,19].

The gait-based measures showed a differentiated picture of improved and deteriorated gait performance for individual patients, while the PROMs improved for all patients. Therefore, we were not able to use those self-reported measures as a reference for the clustering of the patients' functional responses. Instead, we introduced a clustering of patients into positive and negative responder classes based on post- vs. pre-operative gait parameter changes. A two class approach was used by dichotomizing the gait parameters based on a median split. Patients showing improved or deteriorated gait parameters were considered positive and negative responders, respectively. The approach of defining positive and negative improvement is only heuristic but a first step towards the quantification of functional improvement using objectively measured gait parameters. It needs to be further validated based on larger sample sizes. However, the direction of positive change of gait corresponding to gait improvement follows biomechanical considerations (e.g. higher gait speed). The median split thresholds lay around zero for all gait parameters, which was equivalent to the heuristically chosen improvement definition and thus a strong argument for the validity of the median split for dichotomization. However, the median split has drawbacks due to the loss of information, as small and large differences contribute equally to class membership. An assignment to either class would necessitate further investigation whether the gait parameter differences truly indicate clinically meaningful changes. In the future, the minimal clinically important difference (MCID) could be employed. This choice would also create an additional "non-responder" class, clustering patients showing only small gait changes. The introduction of an additional class would, however, necessitate a higher sample size. A different approach would be to use regression methods in order to avoid dichotomization of patients into different response classes. Other objective functional outcome assessments involving measures such as physical performance from accelerometry could be used to assess real-life physical performance and to group patients into differently responding groups [37]. It has also been shown that PROMs based on the Knee Injury and Osteoarthritis Outcome Score (KOOS) could be used for sub-grouping into non-, low-, and high-responders [25,26].

The prediction of positive or negative surgery outcome based on objective gait parameters was well feasible with all classification algorithms. For this data set, the decision tree performed best (89% accuracy) and allowed the direct interpretation of the prediction model. The final model included stride time and stride length as discriminative gait parameters. Patients with short stride times and long strides pre-operatively were prone to exhibit rather negative responses. This corresponds to findings of Berliner et al., who observed that patients with high pre-operative function were less likely to experience clinically meaningful improvements [23]. The high specificity indicates that patients with low probability of benefiting from TKA could be well predicted beforehand from a gait function evaluation using spatio-temporal gait parameters. Non-invasive treatment options (e.g. insoles, physical therapy, braces) could be exhausted for those patients before surgery. Additionally, joint preserving surgery for leg axis correction could be considered.

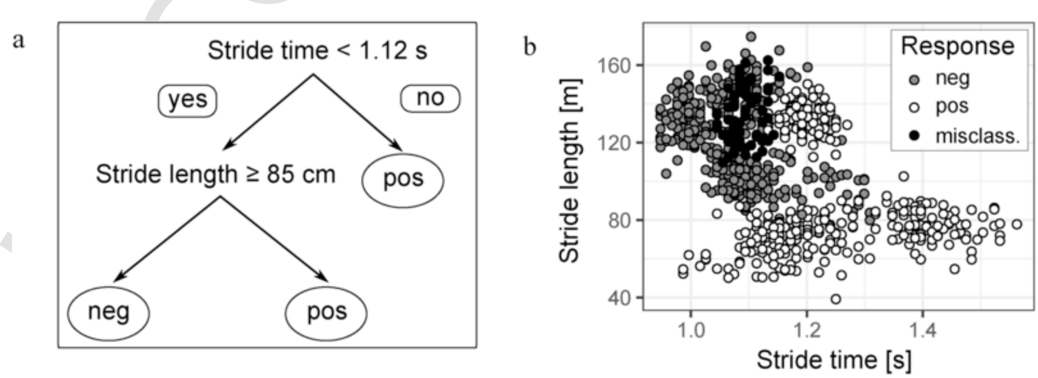

Fig. 2. (a) Decision tree trained on the whole data set. (b) Outcome prediction for all pre-operative strides based on the gait parameters stride length and stride time. The misclassifications occurred mainly due to stride time.

Our approach of using gait parameters allows the direct clinical interpretation of the results and a direct comparison with results from different motion capture systems that extract the same parameters. Generic features (e.g. frequency features or curve extrema) directly extracted from the raw inertial sensor signals could also be used for prediction. Such features instead of spatio-temporal gait parameters may further improve the classification accuracy for this specific application. However, transferability to other measurement systems (e.g. gait carpets) is limited, if differing raw signals are acquired.

We cannot exclude that between-day variability had an effect on our outcomes. This should be further evaluated for example by using a longitudinal study protocol. Such a protocol could also reveal information on the patient's rehabilitation process and the optimal time-frame for post-TKA gait analysis.

A limitation of this study is that possible joint deteriorations of the control participants were only assessed by self-reports. No radiographic imaging was performed which could have revealed potential deteriorations and might have led to the exclusion from a radiographic perspective.

The cause for patients to respond differently to the surgery remains as an open question. A limitation of this study was that we could not guarantee the same rehabilitation procedures for all patients over the whole time frame between surgery and post-operative gait assessment. All patients underwent rehabilitation, but we did not monitor the individual physical therapy schemes in detail. This has to be kept in mind as a major confounding variable which should be controlled in future studies. Other potential confounding factors (e.g. population characteristics such as gender [38]) influencing the results should be investigated using a larger cohort of patients. Future studies should also evaluate different clustering approaches to refine the response definition, include further gait parameters (also considering generic features) assessed in different clinical tests. Due to the ease of application, mobile gait analysis systems could also enable the study of the rehabilitation process in more detail in longitudinal study designs.

To conclude, the employed mobile gait analysis system comprising of two foot-worn sensors allowed the acquisition of spatio-temporal gait parameters that provide complementary information to traditionally used PROMs. We were able to objectively assess functional limitations before and after TKA in a population of knee OA patients. In the future, refined methods to identify responders and discriminate between different disease stages are envisioned based on more available data. The identification of patients who are likely to exhibit negative functional gait changes after TKA may allow a better informed pre-operative advice regarding surgery success, including altered treatment options (such as non-invasive treatments or joint preservative surgery), more detailed information on what to expect after surgery, and an improved post-surgery management, including adapted physiotherapeutic training.

Conflict of interest statement

None of the authors have any conflicts of interest with respect to this study.

Acknowledgment

This study was supported by funding from the Emerging Fields Initiative of the Friedrich-Alexander University Erlangen-Nürnberg (FAU) for the project EFIMoves [funding number: 2_Med_03]. Bjoern M. Eskofier gratefully acknowledges the support of the German Research Foundation (DFG) within the framework of the Heisenberg professorship programme [grant number ES 434/8-1].

References

[1] J.W. Bijlsma, F. Berenbaum, F.P. Lafeber, Osteoarthritis: an update with relevance for clinical practice, Lancet 377 (2011) 2115–2126, https://doi.org/10.1016/S0140-6736(11)60243-2.
[2] D.J. Hunter, Osteoarthritis, BMJ 332 (2006) 639–642, https://doi.org/10.1136/bmj.332.7542.639.
[3] C.F. Dillon, E.K. Rasch, Q. Gu, R. Hirsch, Prevalence of knee osteoarthritis in the United States: arthritis data from the third national health and nutrition examination survey 1991–94, J. Rheumatol. 33 (2006) 2271–2279.
[4] K. Turcot, Y. Sagawa, D. Fritschy, P. Hoffmeyer, D. Suvà, S. Armand, How gait and clinical outcomes contribute to patients' satisfaction three months following a total knee arthroplasty, J. Arthroplasty 28 (2013) 1297–1300, https://doi.org/10.1016/j.arth.2013.01.031.
[5] F.S. Hossain, S. Konan, S. Patel, E.C. Rodriguez-Merchan, F.S. Haddad, The assessment of outcome after total knee arthroplasty: are we there yet?, Bone Joint J. 97-B (2015) 3–9, https://doi.org/10.1302/0301-620X.97B1.34434.
[6] S.A.A.N. Bolink, B. Grimm, I.C. Heyligers, Patient-reported outcome measures versus inertial performance-based outcome measures: a prospective study in patients undergoing primary total knee arthroplasty, Knee 22 (2015) 618–623, https://doi.org/10.1016/j.knee.2015.04.002.
[7] L. Sosdian, F. Dobson, T. Wrigley, K. Paterson, K.L. Bennell, M. Dowsey, P. Choong, K. Allison, R. Hinman, Longitudinal changes in knee kinematics and moments following knee arthroplasty: a systematic review, Knee 21 (2014) 994–1008, https://doi.org/10.1016/j.knee.2014.09.009.
[8] J.A. McClelland, K.E. Webster, J.A. Feller, Gait analysis of patients following total knee replacement: a systematic review, Knee 14 (2007) 253–263, https://doi.org/10.1016/j.knee.2007.04.003.
[9] G.L. Hatfield, C.L. Hubley-Kozey, J.L. Astephen Wilson, M.J. Dunbar, The effect of total knee arthroplasty on knee joint kinematics and kinetics during gait, J. Arthroplasty 26 (2011) 309–318, https://doi.org/10.1016/j.arth.2010.03.021.
[10] B.M. Alice, A. Stéphane, S.J. Yoshisama, H. Pierre, S. Domizio, M. Hermes, T. Katia, Evolution of knee kinematics three months after total knee replacement, Gait Posture 41 (2015) 624–629, https://doi.org/10.1016/j.gaitpost.2015.01.010.
[11] M. Tali, J. Maaroos, Lower limbs function and pain relationships after unilateral total knee arthroplasty, Int. J. Rehabil. Res. 33 (2010) 264–267, https://doi.org/10.1097/MRR.0b013e3283352126.
[12] T. Calliess, R. Bocklage, R. Karkosch, M. Marschollek, H. Windhagen, M. Schulze, Clinical evaluation of a mobile sensor-based gait analysis method for outcome measurement after knee arthroplasty, Sensors 14 (2014) 15953–15964, https://doi.org/10.3390/s140915953.
[13] H. Zhang, S. Yan, C. Fang, X. Guo, K. Zhang, Clinical evaluation and gait characteristics before and after total knee arthroplasty based on a portable gait analyzer, Orthop. Surg. 8 (2016) 360–366, https://doi.org/10.1111/os.12270.
[14] J.C. Schlachetzki, J. Barth, F. Marxreiter, J. Gossler, Z. Kohl, S. Reinfelder, H. Gassner, K. Aminian, B.M. Eskofier, J. Winkler, J. Klucken, Wearable sensors objectively measure gait parameters in Parkinson's disease, PLOS ONE 12 (2017) 1–18, https://doi.org/10.1371/journal.pone.0183989.
[15] R.I. Spain, R.J. St. George, A. Salarian, M. Mancini, J.M. Wagner, F.B. Horak, D. Bourdette, Body-worn motion sensors detect balance and gait deficits in people with multiple sclerosis who have normal walking speed, Gait Posture 35 (2012) 573–578, https://doi.org/10.1016/j.gaitpost.2011.11.026.
[16] A. Rampp, J. Barth, S. Schülein, K.G. Gaßmann, J. Klucken, B.M. Eskofier, Inertial sensor based stride parameter calculation from gait sequences in geriatric patients, IEEE Trans. Biomed. Eng. 62 (2015) 1089–1097, https://doi.org/10.1109/TBME.2014.2368211.
[17] M. Börjesson, L. Weidenhielm, E. Mattsson, E. Olsson, Gait and clinical measurements in patients with knee osteoarthritis after surgery: a prospective 5-year follow-up study, Knee 12 (2005) 121–127, https://doi.org/10.1016/j.knee.2004.04.002.
[18] R. Senden, B. Grimm, K. Meijer, H. Savelberg, I.C. Heyligers, The importance to including objective functional outcomes in the clinical follow up of total knee arthroplasty patients, Knee 18 (2011) 306–311, https://doi.org/10.1016/j.knee.2010.07.008.
[19] S.A.A.N. Bolink, S.N. van Laarhoven, M. Lipperts, I.C. Heyligers, B. Grimm, Inertial sensor motion analysis of gait, sit-stand transfers and step-up transfers: differentiating knee patients from healthy controls, Physiol. Meas. 33 (2012) 1947–1958, https://doi.org/10.1088/0967-3334/33/11/1947.
[20] R. van der Straaten, L. De Baets, I. Jonkers, A. Timmermans, Mobile assessment of the lower limb kinematics in healthy persons and in persons with degenerative knee disorders: a systematic review, Gait Posture 59 (2017) 229–241, https://doi.org/10.1016/j.gaitpost.2017.10.005.
[21] J.M. Barden, C.A. Clermont, D. Kobsar, O. Beauchet, Accelerometer-based step regularity is lower in older adults with bilateral knee osteoarthritis, Front. Hum. Neurosci. 10 (2016) 625, https://doi.org/10.3389/fnhum.2016.00625.
[22] C.A. Clermont, J.M. Barden, Accelerometer-based determination of gait variability in older adults with knee osteoarthritis, Gait Posture 50 (2016) 126–130, https://doi.org/10.1016/j.gaitpost.2016.08.024.
[23] J.L. Berliner, D.J. Brodke, V. Chan, N.F. SooHoo, K.J. Bozic, Can preoperative patient-reported outcome measures be used to predict meaningful improvement in function after TKA?, Clin. Orthop. Relat. Res. 475 (2017) 149–157, https://doi.org/10.1007/s11999-016-4770-y.

[24] A.J. Smith, D.G. Lloyd, D.J. Wood, Pre-surgery knee joint loading patterns during walking predict the presence and severity of anterior knee pain after total knee arthroplasty, J. Orthop. Res. 22 (2004) 260–266, https://doi.org/10.1016/S0736-0266(03)00184-0.

[25] D. Kobsar, S.T. Osis, B.A. Hettinga, R. Ferber, Gait biomechanics and patient-reported function as predictors of response to a hip strengthening exercise intervention in patients with knee osteoarthritis, PLOS ONE 10 (2015) e0139923, https://doi.org/10.1371/journal.pone.0139923.

[26] D. Kobsar, S.T. Osis, J.E. Boyd, B.A. Hettinga, R. Ferber, Wearable sensors to predict improvement following an exercise intervention in patients with knee osteoarthritis, J. Neuroeng. Rehabil. 14 (2017) 94, https://doi.org/10.1186/s12984-017-0309-z.

[27] J. Barth, C. Oberndorfer, C. Pasluosta, S. Schülein, H. Gassner, S. Reinfelder, P. Kugler, D. Schuldhaus, J. Winkler, J. Klucken, B.M. Eskofier, Stride segmentation during free walk movements using multi-dimensional subsequence dynamic time warping on inertial sensor data, Sensors 15 (2015) 6419–6440, https://doi.org/10.3390/s150306419.

[28] J. Hannink, M. Ollenschläger, F. Kluge, N. Roth, J. Klucken, B.M. Eskofier, Benchmarking foot trajectory estimation methods for mobile gait analysis, Sensors 17 (2017) 1940, https://doi.org/10.3390/s17091940.

[29] F. Kluge, H. Gaßner, J. Hannink, C. Pasluosta, J. Klucken, B.M. Eskofier, Towards mobile gait analysis: concurrent validity and test-retest reliability of an inertial measurement system for the assessment of spatio-temporal gait parameters, Sensors 17 (2017) 1522, https://doi.org/10.3390/s17071522.

[30] J.C. Menant, J.R. Steele, H.B. Menz, B.J. Munro, S.R. Lord, Effects of walking surfaces and footwear on temporo-spatial gait parameters in young and older people, Gait Posture 29 (2009) 392–397, https://doi.org/10.1016/j.gaitpost.2008.10.057.

[31] F. Dobson, R. Hinman, E. Roos, J. Abbott, P. Stratford, A. Davis, R. Buchbinder, L. Snyder-Mackler, Y. Henrotin, J. Thumboo, P. Hansen, K. Bennell, OARSI recommended performance-based tests to assess physical function in people diagnosed with hip or knee osteoarthritis, Osteoarthr. Cartil. 21 (2013) 1042–1052, https://doi.org/10.1016/j.joca.2013.05.002.

[32] C. Leys, C. Ley, O. Klein, P. Bernard, L. Licata, Detecting outliers: do not use standard deviation around the mean, use absolute deviation around the median, J. Exp. Soc. Psychol. 49 (2013) 764–766, https://doi.org/10.1016/j.jesp.2013.03.013.

[33] T. Fawcett, An introduction to ROC analysis, Pattern Recognit. Lett. 27 (2006) 861–874, https://doi.org/10.1016/j.patrec.2005.10.010.

[34] E. Combrisson, K. Jerbi, Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy, J. Neurosci. Methods 250 (2015) 126–136, https://doi.org/10.1016/j.jneumeth.2015.01.010.

[35] S. Chen, J. Lach, B. Lo, G.Z. Yang, Toward pervasive gait analysis with wearable sensors: a systematic review, IEEE J. Biomed. Health Inform. 20 (2016) 1521–1537, https://doi.org/10.1109/JBHI.2016.2608720.

[36] P. Levinger, D.T.H. Lai, R.K. Begg, K.E. Webster, J.A. Feller, The application of support vector machines for detecting recovery from knee replacement surgery using spatio-temporal gait parameters, Gait Posture 29 (2009) 91–96, https://doi.org/10.1016/j.gaitpost.2008.07.004.

[37] M. Smuck, C. Tomkins-Lane, M.A. Ith, R. Jarosz, M.C.J. Kao, Physical performance analysis: a new approach to assessing free-living physical activity in musculoskeletal pain and mobility-limited populations, PLOS ONE 12 (2017) 1–16, https://doi.org/10.1371/journal.pone.0172804.

[38] A. Phinyomark, S.T. Osis, B.A. Hettinga, D. Kobsar, R. Ferber, Gender differences in gait kinematics for patients with knee osteoarthritis, BMC Musculoskelet. Disord. 17 (2016) 1–12, https://doi.org/10.1186/s12891-016-1013-z.