Discrimination of knee osteoarthritis patients from asymptomatic individuals based on pain sensitivity and knee vibroarthrographic recordings

Samani, Afshin; Andersen, Rasmus Elbæk; Arendt-Nielsen, Lars; Madeleine, Pascal Max

Published in:
Physiological Measurement

DOI (link to publication from Publisher):
10.1088/1361-6579/ab8857

Publication date:
2020

Document Version
Accepted author manuscript, peer reviewed version

Discrimination of knee osteoarthritis patients from asymptomatic individuals based on pain sensitivity and knee vibroarthrographic recordings

Discrimination of knee osteoarthritis patients from asymptomatic individuals based on pain sensitivity and knee vibroarthrographic recordings

Afshin Samani[*1], Rasmus Elbæk Andersen[1,2], Lars Arendt-Nielsen[2], Pascal Madeleine[1]

[1] Sport Sciences – Performance and Technology, Department of Health Science and Technology, School of Medicine, Aalborg University, Aalborg, Denmark

[2] SMI®, Department of Health Science and Technology, School of Medicine, Aalborg University, Aalborg, Denmark

Submission type: Original article

Address for correspondence: Associate Professor Afshin Samani, Sport Sciences - Performance and Technology, Department of Health Science and Technology, School of Medicine, Aalborg University, Niels Jernes vej 12, 9220 Aalborg East, Denmark. Tel: 00 45 99402411. Fax: 00 45 98154008. Email: afsamani@hst.aau.dk

ABSTRACT

Recording of knee vibroarthrographic (VAG) activity during activities of daily living (ADL) can contribute to diagnose knee osteoarthritis (KOA). However, classifying KOA patients based on knee VAG during ADL has been an elusive problem not related to knee pain. Therefore, the aims of this study was to classify KOA patients based on 1) VAG during ADL and 2) knee pain sensitivity and then compare their results.

The experimental procedure consisted of the recording of VAG signals during four ADLs (over-ground gait, stairs descent, stairs ascent and sit-to-stand) from eight patellar and peri-patellar locations in 20 KOA and 20 asymptomatic participants. Pressure pain thresholds (PPT) were obtained from eight locations around the knee joint to quantify pain sensitivity. A random forest classifier was utilized to identify KOA patients based on VAG signal features and PPTs. The most important features contributing to the classification accuracy were determined. The KOA patients participated in a second identical experimental session to examine the day-to-day reproducibility.

The participants were classified with accuracy of 90%, 70%, 64% and 82% during over-ground gait, stairs descent, stairs ascent and sit to stand, respectively. However, the accuracy of the classifier was reduced by about 10-25% due to a systematic bias in the extracted features across days. Features of the VAG signals in time and frequency domains as well as nonlinear features were found importantly contributing towards the classification accuracy. The VAG features extracted from the lateral side of the knee was found to be more informative than other locations. The classification based on PPT reached 77%. Medial and proximal knee PPT points contributed to the classification accuracy. This study showed that using multichannel VAG signals to identify KOA patients allows better accuracy than the use of PPTs. However, VAG setup must be standardized to avoid day-to-day bias.

Key Words: Accelerometer, random forest classification, cross-validation, pressure pain thresholds, algometry

1. Introduction

Knee osteoarthritis (KOA) is a very common disease resulting in millions of years living with disability among affected people around the world [1]. Often excessive biomechanical loading and inflammatory processes involving bone erosion are believed to be of the underlying factors in the development of KOA [2]. Vibroarthrography (VAG) of the knee is a technique recording the vibrations due to intra-and extra-articular movements and frictions during joint motion [3]. VAG is of interest as it could be a new low-cost non-invasive diagnostic tool to provide insight into the relationship between KOA and activities of daily living [4]. With recent advances in developing miniature sensors, VAG can be recoded in a multichannel setup and allow for investigating the spatial dependencies of the VAG signal and its relationship to the underlying joint structure and pain locations [5].

It has recently been shown that common VAG signal features are spatially distributed non-uniformly around the knee joint during knee flexion-extension movements [4]. Such a spatial heterogeneity of the VAG signal distribution is suggested to be the results of variations in the internal pressure distribution applied to the cartilage and synovial fluid [6].

Most studies on VAG signal features have been conducted during open chain movements not characterizing activities of daily living (ADL). Closed kinetic chain exercises have been reported to be effective in rehabilitation of KOA patients [7]. However, studying VAG signals during closed kinetic chain movements has been an elusive problem due to movement artefacts [4].

Even though VAG signals have been used in various studies to characterize KOA patients, it is not clear to what extent the discriminative power of VAG signal can improve the diagnosis beyond relatively simple clinical examinations probing altered sensory manifestations in KOA patients. Pain sensitization is a typical symptom of KOA, which is manifested as increased sensitivity to mechanical, thermal, electrical and chemical stimuli [8,9]. Pressure

algometry is a very common technique to assess pain sensitivity reflected in the pressure pain threshold (PPT) segmentally and extra-segmentally in KOA patients [10]. Similar to multichannel VAG recordings, multiple PPT assessments allows the spatial localization and calculation of knee topographical pressure pain sensitivity maps [10]. A recent study has shown that the assessment of PPTs in KOA patients is reliable over days [11] and enables the assessment of somatic structures sensitivity over time in KOA patients. Therefore, multiple PPT assessment can also potentially be utilized to differentiate the KOA patients from asymptomatic individuals.

The aims of this explorative clinical study were 1) to examine the feasibility of utilizing a multichannel VAG recorder to reveal the topographical distribution of VAG features and to classify the KOA patients from asymptomatic controls during ADLs, 2) to examine whether the classification accuracy of the KOA patients based on VAG signals can reach beyond the classification accuracy solely based on the participants pain sensitivity reflected in topographical PPT assessments and 3) determine the day-to-day reproducibility of the classification results. The fulfillment of the aforementioned aims makes the basis of the novel aspects of the study. In this paper, the method section describes the characteristics of participants and their inclusion and exclusion criteria, the experimental protocol and data recordings including VAG signals and PPTs, the data analysis approach and the adopted statistical methods to classify the participants. The results section presents the obtained results in terms of classification accuracy and important features contributing towards the classification accuracy. Additionally, the distribution of important features across the recorded anatomical locations and the day-to-day reproducibility of the results are outlined. In the discussion section, the results are interpreted and compared to relevant existing literature.

2. Related work

A previous study has presented a thorough systematic review of existing studies on VAG signals to diagnose KOA patients [12]. Previous studies have successfully utilized VAG signals from one single accelerometer to classify KOA patients and asymptomatic controls [6,13,14] but topographical mapping is now a viable solution using multi-channel recordings [15]. High accuracy in classification (>90%) of KOA patients have been achieved in some studies [16,17] but this has been mainly during activities with open kinetic chain [12]. Some recent studies has shown that VAG signal characteristics may alter in KOA patients during ADLs [15,18] but this has not shown whether that could be generalized and used to effectively discriminate KOA patients from asymptomatic individuals. The non-uniformity of the VAG signal features around the knee joint during ADLs has been indicated [15], but it is not clear whether specific locations around the knee were more informative to differentiate the KOA patients from asymptomatic controls based on VAG signals around the knee joint and which features of the VAG signal are more important for this purpose.

3. Methods

3.1. Design

The study involved two groups of participants, namely, a KOA patient group (11 males and 9 females) and an asymptomatic control group consisting of 20 asymptomatic participants (10 males and 10 females). The participants were the same as in a recent cross-sectional study only investigating VAG signal characteristics in KOA patients and controls during ADLs [15]. One subject from each of the groups did not finalize the recording on the second experimental day (see below), therefore, the analysis was performed on 19 subjects in each group. The study was conducted according to the ethical guidelines of the Helsinki Declaration and was approved by the North Denmark Region Committee on Health Research Ethics (VN-20160081). All participants provided written informed consent.

3.2. Participants

The patient group was recruited from a database at the Centre for Clinical and Basic Research (CCBR, Aalborg, Denmark) and the asymptomatic control group was recruited from the dwelling community. The diagnosis of KOA was in accordance to the American College of Rheumatology classification [19] and patients were clinically screened for inclusion. The KOA patients were included if they were aged 18-80, clinically diagnosed KOA with Kellgren-Lawrence grade ≥ 2, self-reported pain during normal walking and BMI <35, no use of painkillers in the 24 hours prior to experimentation. The participants in the asymptomatic control group were matched with the KOA patient group in terms of sex, body height (169.0 (10.8) vs. 169.8 (9.0) cm), body mass (81.5 (13.0) vs. 77.7 (9.9) Kg) and BMI (27.2 (3.2) vs. 28.1 (2.7) kg/cm^2) but controls were about four years older than the KOA patients (70.3 yrs. (5.9) vs. 66.2 yrs. (5.2)). The participants in the asymptomatic control group had neither a history of pain nor use of pain killers 24 h prior the test.

3.3. Experimental protocol

The participants in the KOA patient group took part in two experimental days whereas the asymptomatic control group had only one experimental day as the study also aimed at examining the reliability of the study outcomes in the patient group. The experiment consisted of VAG recordings during several relevant activities of daily living (ADL): (i) over-ground gait for 40 m, (ii) stairs descent (10 stairs), (iii) stairs ascent (10 stairs) and (iv) 5 repetitions of sit to stand movement as recommended for testing physical functions in KOA patients [20]. The ADLs (i-iv) were performed with a counterbalanced order and carried out at a self-chosen speed meaning that the participants were allowed to perform five repetitions of sit to stand in max of 60 s while their arms were kept alongside their body and not used to help the movement. The hand railing was not used during stair descent and ascent and the task was performed at the slowest pace which was comfortable enough for the participants while maintaining their balance. The subjects were barefoot except wearing their socks.

3.4. Vibroarthrographic recording

During each of the ADLs (i-iv) mentioned above, VAG recording was carried out using a custom-made device based on a Trentadue wireless multichannel recorder (OT Bioeletronica, Torino, Italy), a custom 16 channel accelerometers adaptor and micro machined accelerometers LIS344ALH (ST microelectronics, Geneva, Switzerland). The recording chain has a sensitivity of 600 mV/g and 0-1800 Hz linear transmission. The miniature size (8.5×7 mm) and light weighted (approx. 0.75 g) accelerometers were set to only record the acceleration in the orthogonal direction. The measurement setup is equipped with a band-pass hardware filter (10-500Hz) sampling the VAG signal at 2000 Hz and a gain of three. The data collection was carried out by a custom script (IOIVibcorder, Aalborg University, Aalborg, Denmark) implemented in MATLAB 2016a (The MathWorks, Inc, Natick, Massachusetts, United States).

The accelerometers were placed on the most painful knee in the KOA patient group (all right knee except one). In the asymptomatic control group, the accelerometers were placed on the right knee except one participant to keep the knee side balanced between the two groups. Accelerometers were attached to the skin with double sided tape on eight significant points on the knee joint. Four accelerometers were placed on the patella in square form with the side of 1-2 cm, one on the tibial tuberosity below the patella, two on the lateral and medial side of the knee with 1-2 cm from their respective epicondyle of femur towards the patella and finally, one over the quadriceps tendon [4] (see figure 1(b)). The recording package was placed in a belt bag fastened around the participants' waist and the wires were taped to the thigh without any obstruction to natural movements.

3.5. Pressure pain threshold

The PPTs were assessed while the participants lied in a supine position. A handheld algometer (type II, Somedic AB, Hoerby, Sweden) with a tip area of 1 $cm^2$was utilized to

register PPTs from eight anatomical locations [11] around the most painful knee in a randomized order (see figure 2(b)). The eight locations were: 2 cm distal to the inferomedial and inferolateral edge of patella, 3 cm lateral to the center of the lateral edge of patella, 2 cm proximal to the superolateral, superior and superomedial edge of patella, 3 cm medial to the center of the medial edge of patella and on the center of patella. The entire procedure for all participants was performed by the same examiner. The procedure involved applying a continuous pressure with an ascending pressure gradient of 30 kPa/s until the participants felt pain and pressed a stop button. The pressure threshold indicated the onset of pain sensation, was registered, and noted by the examiner as the PPT for the specific location. This procedure was repeated three times and the mean value was used for the analysis. A 1-min resting interval was considered between the repetitions to avoid temporal summation [21].

3.6. Data analysis

The VAG signals were converted into SI units (ms^{-2}) and digitally filtered using a band pass FIR filter using a Kaiser window, 10-500 Hz (1453-points, beta: 5.6533). Since the VAG signals captured during ADL were influenced by the mechanical impact of the movement pattern (i.e. cyclic pattern of the movement), the VAG signals were adaptively filtered to further reduce the effect of the common components on the VAG signals due to the movement pattern [6]. For each of the VAG signals recorded from a specific location, the rest

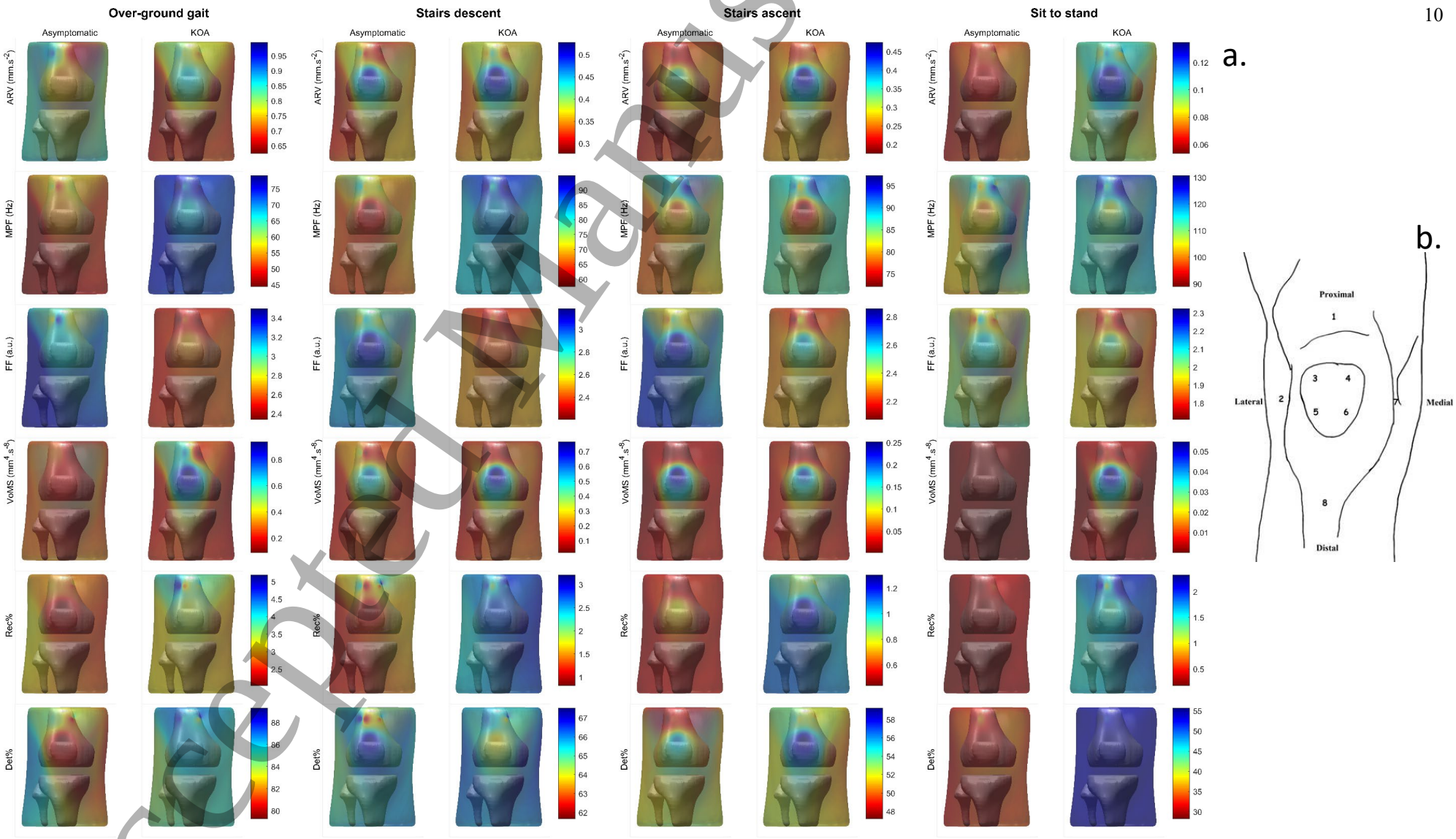


Figure 1 a) Average maps of the spatial interpolation of extracted vibroarthrographic features in 3D. Averaged rectified value (ARV, mm.s-2), mean power frequency (MPF, Hz), form factor (FF, a.u.), variance of means squared (VoMS, mm4.s-8), % of recurrence (%REC), and % of determinism (%DET) of the vibroarthrographic signals recorded using eight accelerometers (depicted to the right side of the figure) during activities of daily living (gait, stairs descent, stairs ascent and sit to stand) among patients with knee osteoarthritis (N = 19) and asymptomatic participants (N = 19). The interpolation was based on an inverse distance weighting interpolation method (inverse distance weighted interpolation [32]) and solely used for visualization purposes. b) the anatomical location of placing the accelerometers.

of the VAG signals were used to render the reference input in a recursive least square adaptive filter (δ=0.2 and λ=0.97) [22]. The reference input was obtained based on the synchronized averaging of the VAG signals to improve the signal to noise ratio of the common component due to the movement pattern [23]. Additionally, independent component analysis was performed to remove components causing unstable baseline [24]. The first and last second of the VAG signals were removed to avoid any transient effect at the beginning and end of each activity. The VAG signals were analyzed in epochs of four seconds during the ADLs [14]. Six features of the signals, i.e., averaged rectified values (ARV), mean power frequency (MPF), variance of means squared (VoMS), form factor (FF), the % of determinism and recurrence (%DET and %REC) have been suggested reflecting the characteristics of the VAG signal thoroughly [12]. ARV represents the signal amplitude, MPF is to reflect the central tendency of the frequency spectrum, VoMS and FF indicate the relative and absolute variability of the VAG signals [13] and finally (%DET and %REC) are to reflect the regularity and deterministic nature of the VAG signals [25]. Recurrent analysis was performed on the Z-score of the VAG signals and it was based on portraying the signal in a multidimensional space known as embedded space [26]. %REC quantifies how much the trajectories of the signal in the embedded space recur (returns to the vicinity of past points). %DET determines the fraction of the recurrent points which constitute a deterministic pattern. The feature of recurrent analysis have been shown to be reflecting the dynamics of the underlying system [25]. To construct the embedding space, the embedding dimension, delay and tolerance values were determined based on the global false nearest neighbor approach, the drop of auto correlation function below 0.2 and %REC minimization optimization method, respectively [26]. As reported previously [4], an embedding dimension of five, the delay of 19 ms and the tolerance of 0.2839 were found to be appropriate. As the duration of the recordings were different across subjects (self-chosen pace), out of the

calculated features equal number of epochs in the middle of the activity were chosen for conducting the statistical analysis. This resulted in eight, five, two and two epochs during the over-ground gait, the sit to stand movement, the stairs descent and the stairs ascent, respectively.

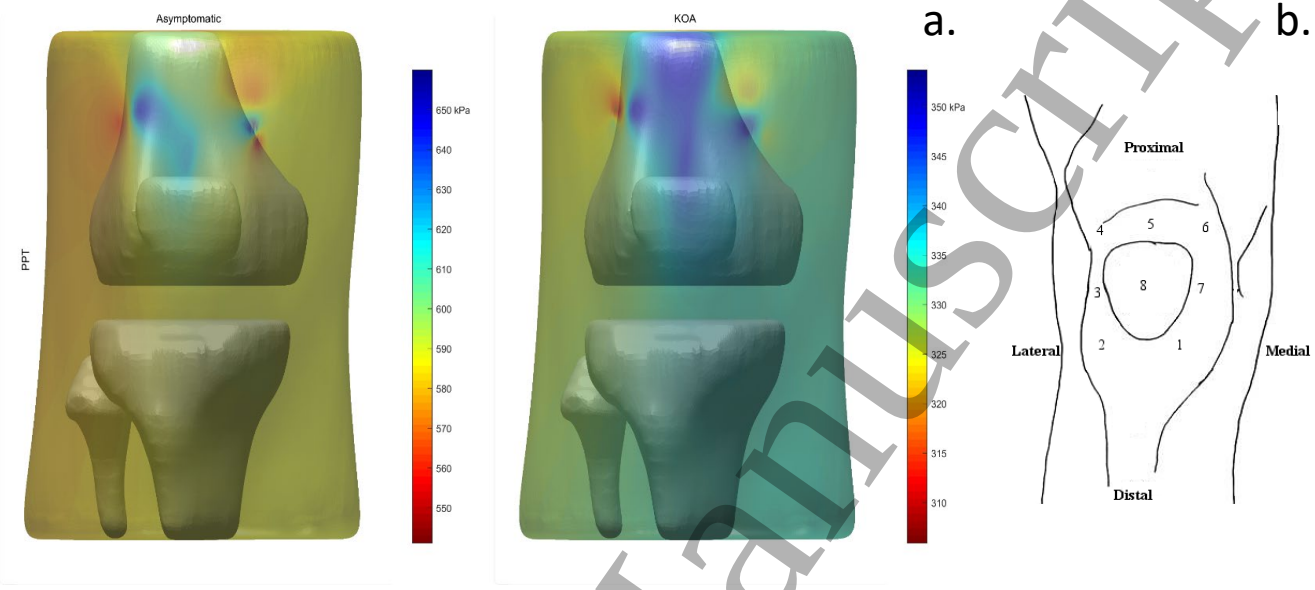


Figure 2 Average maps of the spatial interpolation of pressure pain threshold (PPT, kPa) in 3D. a) PPT maps are depicted for patients with knee osteoarthritis (KOA; N = 19) and asymptomatic participants (N = 19). The interpolation was based on an inverse distance weighting interpolation method (inverse distance weighted interpolation [32]) and solely used for visualization purposes. b) the anatomical location of obtained PPT points.

3.7. Statistical analysis

A random forest classifier (RFC) was utilized to discriminate between KOA patients and asymptomatic individuals based on the extracted features of VAG signals. The RFC was chosen due to its resistance to overfitting to the training dataset [27] and its invariance to the scaling of the features [28] used in this study. The minimum leaf node size was set to 15 based on a search method to avoid overfitting of the RFC [29]. A set of 100 decision trees were included in the construct of RFC. To choose the features for splitting at the nodes of the decision trees, a curvature test was performed between the features and the class labels, i.e., patient and asymptomatic [30]. The curvature test examines the association of the features and the class label. The importance of the features towards the classification accuracy was

determined based a permutation test of the data points not used in the train phase of the trees [31]. The classification error of the RFC was tested by a leave-one-person-out approach to use one participant as the test dataset and the rest of the participants as the training dataset. This procedure was repeated so that all participants used as the test dataset once. Additionally, to examine the reproducibility of the results across days, the error of RFC was calculated with the dataset collected from KOA patients in the second experimental day. As outlined in section 3.6, the VAG recording from each subject was divided into multiple epochs, since the ultimate aim was to classify the participants and not the epochs, each participant was classified based the majority of the assigned labels of his/her epochs. A similar approach was adopted to differentiate between the subjects based on the PPTs with the minimum leaf node size of three as in this case as the number of observations in the dataset was the same as the number of participants and no epoch could be defined for the PPTs.

A Mann-Whitney U test was performed to compare the VAG signal features with an important contribution towards the classification performance across the participant groups and across experimental days in the KOA group. A similar approach was performed for the important PPTs for the classification based on the PPTs. If not specified otherwise, the results were presented as the mean (SD).

4. Results

4.1 Classification accuracy

The classification error for the training, the testing datasets and the second experimental day dataset of the KOA patients were provided (Table 1). The test error was lowest for over-ground gait (16%) and highest for stair ascent (38%). The classification error in the second experimental day increased for all ADLs except for stair descent but the increased error in the second day was more marked for stair ascent (50%). When the participants were classified

based on the majority of their epochs, the classification errors of 10%, 30%, 36% and 18% were found for over-ground gait, stairs descent, stairs ascent and sit to stand, respectively. However, the classification error in the second experimental day was increased to 37%, 21%, 58% and 25% for over-ground gait, stairs descent, stairs ascent and sit to stand, respectively. The sensitivity, specificity and the area under the receiver operating characteristics of the RFC in all tested ADLs were obtained (Figure 3). The test error based on the PPTs was 23% (42%) and the in the second experimental day the error was quite consistently about 25%. The sensitivity, specificity and the area under the receiver operating characteristics of the RFC based on the PPT points were also obtained (Figure 4).

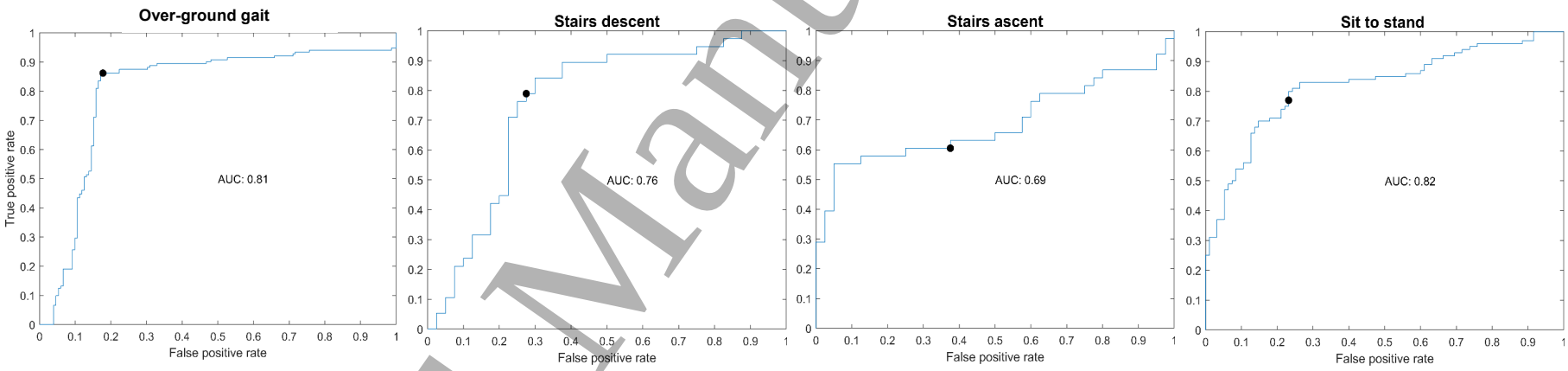
4.2 Important VAG features and PPT locations

The VAG features extracted from the lateral side of the knee made an important contribution to the classification in all the ADLs (Table 1). Additionally, the features extracted from the medial and distal side of the knee were not found significant in any of the ADLs. Nonlinear features were only found significant in stair decent and sit to stand. The spatial distribution of the extracted VAG signal features are depicted in Figure 1(a). Of note, the maps depicted in this figure visualizes the overall spatial distribution of VAGs of the underlying tissue due to intra-and extra-articular movements as well as frictions during joint motion and are not specific to bone articular surfaces. For illustration of the spatial distribution of VAG features, the VAG features were interpolated based on an inverse distance weighted interpolation approach [32].

The detailed results of the comparison between the participant groups in terms of important features in the classification can be found in Table 2. Generally, the KOA patients were characterized by higher ARV, MPF, VoMS, %REC, %DET and lower FF where they found to be important features for classification of the groups. The important features that exhibited a significant bias across the experimental days are presented in Table 3. Wherever a bias

across days was found, ARV, MPF and VOMS were found to slightly drop in the second day and FF was found to slightly increase.

The detailed results of the comparison between the participant groups in terms of important PPT points contributed towards the classification of the participant groups are presented in Table 4. Increased pain sensitivity, i.e., lower PPT in KOA patients was verified. Particularly, the contribution of the PPT points on the medial and proximal side of the knee were found important for the classification. As expected spatial heterogeneity of PPT maps was also observed (figure 2(a)). The maps depicted in this figure visualizes the overall spatial distribution of mechanical sensitivity to pain in the underlying tissue and are not specific to bone articular surfaces. For illustration of the spatial distribution of PPT maps, the PPT assessments were interpolated based on an inverse distance weighted interpolation approach [32].



Sensitivity	Specificity
86.2	82.2

Sensitivity	Specificity
78.9	72.5

Sensitivity	Specificity
60.5	62.5

Sensitivity	Specificity
77.0	76.8

Figure 3 The receiver operating characteristics for the classification of participants groups based on vibroarthrographic features (the group of knee osteoarthritis participants was set as the positive class). The area under the curve (AUC) is displayed on the figure and the sensitivity and specificity of classification during activities of daily living (over-ground gait, stairs descent, stairs ascent and sit to stand) are presented below the figure. The round black point on the figure is where the discrimination threshold was set to 0.5 and the sensitivity and specificity were computed at this point.

Table 1 The mean (SD) of the error rate of the classification of the random forest classification method in the train dataset, testing dataset and the second experimental day for knee osteoarthritis (KOA) patients. Furthermore, the important features for the classification of the groups, i.e., symptomatic KOA and asymptomatic participants during activities of daily living (gait, stairs descent, stairs ascent and sit to stand) among patients with knee osteoarthritis.

	Training error (%)	Test error (%)	Second day error (%)	Important features
Gait	14.4 (0.8)	16.1 (31.0)	32.9	VAG2_MPF, VAG2_FF, VAG5_MPF
Stair descent	14.1 (1.8)	25.6 (36.0)	15.8	VAG2_ARV, VAG1_FF, VAG2_%REC, VAG1_MPF
Stair ascent	17.8 (2.4)	38.5 (43.6)	50.0	VAG2_ARV, VAG6_ARV, VAG4_FF, VAG2_FF
Sit to stand	15.9 (1.1)	23.1 (31.6)	30.0	VAG4_ARV, VAG4_%DET, VAG2_%REC, VAG4_VoMS

Table 2 The results of the Mann-Whitney U test comparing the symptomatic and asymptomatic groups in term of the important features contributing to the classification performance. The Median [25th -75th percentile] are reported in each case.

	Feature	Mann-Whitney	KOA	Asymptomatic
Gait	VAG2_MPF (Hz)	U= 14611, p< 0.001	66.2 [56.8-90.7]	40.9 [34.7-49.6]
	VAG2_FF (a.u.)	U= 31908, p< 0.001	2.5 [2.0-2.8]	3.4 [3.1-3.8]
	VAG5_MPF (Hz)	U= 16847, p< 0.001	67.0 [51.8-82.5]	43.4 [35.8-54.4]
Stair descent	VAG2_ARV ($mm.s^{-2}$)	U= 1502, p= 0.4	0.3 [0.1-0.4]	0.2 [0.1-0.3]
	VAG1_FF (a.u.)	U= 1959, p< 0.001	2.3 [1.9-2.9]	3.2 [2.4-3.9]
	VAG2_%REC (%)	U= 1310, p= 0.007	0.7 [0.2-3.2]	0.3 [0.1-0.5]
	VAG1_MPF (Hz)	U= 1201, p< 0.001	79.7 [59.9-102.7]	51.6 [35.4-74.9]
Stair ascent	VAG2_ARV ($mm.s^{-2}$)	U= 1368, p= 0.03	0.2 [0.1-0.3]	0.2 [0.1-0.3]
	VAG6_ARV ($mm.s^{-2}$)	U= 1324, p= 0.01	0.3 [0.2-0.3]	0.2 [0.1-0.3]
	VAG4_FF (a.u.)	U= 1796, p= 0.03	1.9 [1.8-2.0]	2.3 [1.8-2.9]
	VAG2_FF (a.u.)	U= 1830, p= 0.01	2.2 [1.8-2.6]	2.7 [2.0-3.2]
Sit to stand	VAG4_ARV ($mm.s^{-2}$)	U= 6748, p< 0.001	0.1 [0.1-0.1]	0.1 [0.0-0.1]
	VAG4_%DET (%)	U= 6537, p< 0.001	51.1 [31.3-72.0]	22.9 [14.9-38.4]
	VAG2_%REC (%)	U= 6707, p< 0.001	0.4 [0.1-1.4]	0.1 [0.0-0.2]
	VAG4_VoMS ($mm^4.s^{-8}$)	U= 6801, p< 0.001	2.11e-3 [3.58e-4 - 1.93e-2]	1.68e-4 [3.08e-5 - 6.49e-4]

Table 3 The results of the Mann-Whitney U test comparing the experimental days in term of the important features contributing to the classification performance. The Median [25th -75th percentile] are reported in each case.

	Feature	Mann-Whitney	KOA Day 1	KOA Day 2
Gait	VAG2_MPF (Hz)	U= 26163, p< 0.001	66.2 [56.8-90.7]	59.1 [46.0-72.2]
	VAG2_FF (a.u.)	U= 20127, p< 0.001	2.5 [2.0-2.8]	2.7 [2.4-3.1]
	VAG5_MPF (Hz)	U= 24965, p= 0.02	67.0 [51.8-82.5]	60.6 [47.2-74.3]
Stair descent	VAG2_ARV ($mm.s^{-2}$)	U= 1675, p= 0.03	0.3 [0.1-0.4]	0.2 [0.1-0.3]
	VAG1_FF (a.u.)	U= 1358, p= 0.2	2.3 [1.9-2.9]	2.5 [2.0-2.9]
	VAG2_%REC (%)	U= 1436, p= 0.7	0.7 [0.2-3.2]	0.8 [0.3-3.4]
	VAG1_MPF (Hz)	U= 1610, p= 0.1	79.7 [59.9-102.7]	67.2 [53.2-95.5]
Stair ascent	VAG2_ARV ($mm.s^{-2}$)	U= 1685, p= 0.02	0.2 [0.1-0.3]	0.1 [0.1-0.2]
	VAG6_ARV ($mm.s^{-2}$)	U= 1575, p= 0.2	0.3 [0.2-0.3]	0.2 [0.2-0.3]
	VAG4_FF (a.u.)	U= 1327, p= 0.1	1.9 [1.8-2.0]	2.1 [1.8-2.5]
	VAG2_FF (a.u.)	U= 1452, p= 0.9	2.2 [1.8-2.6]	2.3 [1.8-2.8]
Sit to stand	VAG4_ARV ($mm.s^{-2}$)	U= 11238, p= 0.004	0.1 [0.1-0.1]	0.1 [0.1-0.1]
	VAG4_%DET (%)	U= 10024, p= 0.9	51.1 [31.3-72.0]	53.5 [31.3-74.0]
	VAG2_%REC (%)	U= 9999, p= 0.9	0.4 [0.1-1.4]	0.3 [0.1-2.1]
	VAG4_VoMS ($mm^4.s^{-8}$)	U= 10993, p= 0.02	2.1e-3 [3.58e-4 - 1.93e-2]	6.8e-4 [1.05e-4 - 7.83e-3]

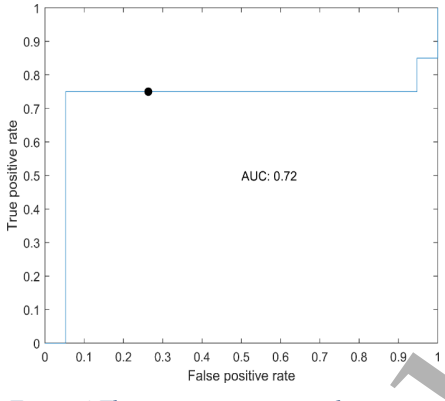
Table 4 The results of the Mann-Whitney U test comparing the symptomatic and asymptomatic groups in term of the important location of registering pressure pain thresholds (PPT) contributing to the classification performance. The Median [25th -75th percentile] are reported in each case

Feature	Mann-Whitney	KOA	Asymptomatic
PPT7 (kPa)	U= 505, p< 0.001	281.5 [227.5-365.0]	482.0 [431.3-753.0]
PPT4 (kPa)	U= 510.5, p< 0.001	280.0 [175.0-383.5]	509.0 [438.3-658.8]
PPT6 (kPa)	U= 496.5, p= 0.001	302.5 [188.0-401.5]	464.0 [379.8-633.8]
PPT1 (kPa)	U= 497, p= 0.001	368.5 [212.5-445.0]	519.0 [405.3-810.0]

5. Discussion

This study showed that multichannel topographical VAG signals captured during ADLs could discriminate between the asymptomatic control group and KOA patients with an accuracy ranging from 64% to 90% (100 minus the error mean) depending on the type of ADL. However, the accuracy dropped by about 10-25% when the KOA patients were tested again on a second experimental day, suggesting the importance of the placement of the accelerometers on the affected knee.

The information obtained from the lateral side of the knee made a very marked contribution to the classification accuracy likely related to the mild to moderate severity of pain in the KOA patients in this study.



Sensitivity	Specificity
75.0	73.7

Figure 4 The receiver operating characteristics for the classification of participants groups based on pressure pain threshold (the group of knee osteoarthritis participants was set as the positive class). The area under the curve (AUC) is displayed on the figure and the sensitivity and specificity of classification during activities of daily living (gait, stairs descent, stairs ascent and sit to stand) are presented below the figure. The round black point on the figure is where the discrimination threshold was set to 0.5 and the sensitivity and specificity were computed at this point.

The classification based on the PPT reached about 77% accuracy and kept its consistency in the second experimental day by about 75% accuracy. This highlights the classification based on the VAG signals in certain ADLs can perform better than the classification based on the pressure pain thresholds, however, the VAG classification consistency across days needs improvement.

In contrast to VAG signal features, PPT points on the medial and proximal side made an important contribution to the classification accuracy. This contrast in sensory manifestation of KOA and the features of mechanical vibrations of the underlying tissue may highlight the

complex structure of the knee joint and the interconnection between the medial to the lateral side of the knee. The discussion presents first our findings in terms of classification accuracy. Then, VAG features are reviewed in relation to the classification results. This section is followed by the evaluation of the investigated ADLs.

5.1 Classification accuracy

Even though the accuracy of the classification of the participants based on the majority of their VAG epochs can reach up to 90% accuracy during over-ground gait, the accuracy of VAG epochs classification ranged from 61% to 84% when a leave-one-out testing approach was adopted. Additionally, a large variability (reflected by a high SD in the classification error) was found. The classification accuracy of the VAG epochs in some of the ADLs was lower than what has been reported (90%) elsewhere [12]. However, most of the previous studies have been performed on knee movements during open-kinetic chain activities and used an identical dataset (Calgary group [13]) to develop their algorithms which may be a source of bias in their results even though using an identical dataset allows for inter-study comparison [12].

Recent studies have emphasized the importance of movements with closed kinetic chain and highlighted their importance in loading of the patellofemoral joint [33,34]. Due to contribution of movement artefacts to the VAG signal characteristics, analyzing VAG in closed-kinetic chain movement poses a challenge. In this study, the effect of movement artefacts was reduced by the application of adaptive filtering, independent component analysis and multichannel recording of the VAG signals [6]. Although a previous study has investigated the VAG signal during sit to stand activity, the study only performs an inferential statistics to report higher VAG signal energy in certain frequency bands in KOA patients [18,35]. Similarly, in a previous study using the same dataset as the current one, the VAG signal characteristics in KOA and asymptomatic participants were studied [15],

however, the study has been limited to performing a statistical inference to find significant differences between the two groups. Developing statistical models for predictions is essential to discriminate between the groups and testing generalizability of the results for diagnoses of the KOA. Thus, in this respect, the current study is among the very few studies investigating VAG signal features in movements with closed kinetic chain and using the signal features to discriminate between the KOA patients and the asymptomatic group.

Even though the obtained accuracy in the current study is not as high as reported in the literature, all the epochs of the signal captured in the timeline of the recording were classified in this study. Ultimately, as the aim was to classify the participants and not the epochs, the participants were classified based on the majority of the assigned labels of his/her epochs. Adopting this approach improved the classification accuracy up to about 90% in over-ground gait even though due to very few (two) epochs during stair ascent and stair descent, not much of improvement was observed during these two ADLs.

The classification accuracy in the second experimental day when the RFC was trained based on the first day dropped markedly. As examined statistically, the VAG signal characteristics exhibited a significant bias on the second day compared with the first experimental day. This suggests that the researcher consistency in setting up the accelerometers, inherent variations in KOA symptoms across days and/or diurnal variation in the water content of soft tissue around the knee joint [36] resulted in observing such a bias resulting in a poorer classification accuracy.

The classification accuracy based on the PPT reached about 77% accuracy and was quite consistent across the experimental days. The PPT assessments have previously been reported to be reliable in this population despite a tendency towards increased PPTs on the second experimental day [11]. Even though the classification based on the VAG signal features can reach higher classification accuracy as mentioned above, PPT assessments are more

straightforward to perform. However, it requires for a trained practitioner to carry out the procedure whereas the VAG signal recordings could potentially be embedded into an ambulatory recording system.

5.2 Important features

In this study, relevant features of VAG signals were extracted as these features have been commonly utilized in previous studies [12]. In addition, features of VAG signals based on the recurrence map analysis were extracted as these features reflect the nonlinear characteristics of the signals [26]. These features have been shown to be relevant to quantify the heterogeneity of VAG signals around the knee during unloaded extension-flexion in an asymptomatic group of subjects [4]. In the current study, a combination of extracted features of VAG signal were importantly contributing to the classification accuracy but the influence of extracted features was dependent on the location of the recording and the type of the ADL. This partly stands in contrast to previously reported results where only VAG ARV and %REC were found to differentiate the groups [15]. Most particularly, a lower %REC in the KOA group has been reported previously, whereas in the current study, a higher %REC was found in the KOA group in some recording locations. Apart from the preprocessing procedure which has been completely modified in the current study, the previous results corresponded to a statistical inference based on a linear mixed model which compared to a RFC is more rigid in modeling the dataset variability, most particularly, a lack of modeling nonlinear separability between the groups [37]. Additionally, the previous results present the main effect of the participant group on the extracted features whereas the current results show the importance of localized features of VAG in a non-homogenous spatial distribution to classify the participant groups. In the aforementioned study, a significant interaction between participant group and the location of the recording has been reported. The current study sheds light on such an interaction and finds the specific locations of the recording which seem to be

more discriminative across the participant group. For example, location #2 on the lateral side of the knee turned out to be an important site of information in all tested ADLs. Additionally, all the important features were obtained from the lateral, proximal and patellar location and in this list no important feature has been obtained from the medial or distal side of the knee. This seems to be in contrast to clinical findings which report the medial compartment of the knee as often most frequently affected by KOA [38]. However, the localized pain sensitivity is found to be different among patients with mild to moderate KOA (similar to the current KOA group) compared with patients with severe cases of KOA [9]. The categorization of KOA patient into mild to moderate or severe has been based on the reported pain on a visual analogue scale (VAS) in the last 24 h before the experiment where VAS≥6 has been the decisive threshold for categorization into strong/server and mild to moderate [39]. Based on such a criterion, most patients (12 out 20) in the current study should have been categorized into mild to moderate group and the highest pain sensitivity was observed on the superolateral side of the knee [11]. The importance of the VAG signal features on the top of patella, proximal and lateral side of the knee is likely related to the patellar cartilage damage often reported in KOA patients [40].

Out of the extracted features only ARV and VoMS were presented in absolute scale and therefore may be sensitive to a normalization procedure. Even though some previous studies have normalized the VAG signal to its amplitude range [e.g. 13], we opted out normalizing these two features because previous results have shown the sensitivity of these features to the interaction between load, movement type and the location of recording [4,15]. Since the RFC is invariant to the scaling of the used features, we believe that the not-normalized scales of ARV and VoMS are not likely to be a major source of concern.

PPT assessments on the medial and proximal side of the knee contributed importantly to the classification. This may seem in contrast to the contribution of the VAG signals extracted

from the lateral side being important features for classification based on the VAG signal features. However, the knee joint construct is interconnected, for example, the anterior and posterior cruciate ligaments connect the medial to lateral side of the knee and could contribute to establish a link between the recorded VAG signals on the lateral side and the pain sensitivity on the medial side of the painful knee. In support of that, a history of anterior cruciate ligament injury has been associated with an increased risk of KOA development [41]. All in all, the present findings suggest that KOA have two different signatures conveying the VAG or PPT information.

5.3 Role of ADL type

Among the captured ADLs, the classification accuracy has generally been highest during gait. This could potentially be related to the consistency of the movement pattern during gait compared with the other ADLs tested in this study. Supporting such a premise, stride-to stride variability of the knee motion in KOA patients has been shown not to be significantly different from an asymptomatic control group [42]. A more plausible explanation may be related to the duration of recording as a relatively shorter recoding during stair descent, stair ascent and sit to stand compared with the over-ground gait results in a reduced ratio between the number of observations to the features of VAG signals. This ratio is important to achieve a robust performance of a multivariate analysis [43]. However, it is trivial to imagine that a longer recording during stair descent and stair ascent was quite difficult for KOA patients to carry out and further safety measures must have been implemented in the setup.

Even though we have included the commonly used VAG features in the analysis, one may assume that feature extraction can be even further optimized if novel methods such as deep learning is used in that front [44,45]. However, deep learning usually encounters some limitations including low interpretability, high computational costs, and the need for large

datasets [46]. Increasing the number of participants could potentially improve the generalizability of the RFC.

5.4 Computational complexity

The bottleneck of the computational cost in the proposed approach was the recurrent map analysis, which in the current implementation, the computational order of growth for %DET algorithm was of the order $O(N^2)$ for a time-series with N samples. Given that the feature set is available, the average execution time, including the testing and training the RFC the elapsed times were as follows. The calculations took 2.8, 1.4, 1.4 and 2.3 s when the original dataset was used for classification and finding the important features and 0.7, 0.6, 0.6 and 0.7 s when only the selected important features were used for classification in over-ground gait, stairs descent, stairs ascent and sit to stand, respectively. These times were 1.4 and 0.6 when the PPT assessments were used for classification. These time were obtained on a laptop running with an Intel(R) Core (TM) i5-5300U@ 2.3 GHz and 8 GB RAM and the analysis was performed on MATLAB 2019b. Given that such analysis is expected to be performed offline, the computational cost of the proposed approach is not of major concern.

5.5 Conclusion

In conclusion, this is the first study investigating the discrimination of the multichannel topographical knee VAG signal between KOA patients and asymptomatic controls during ADLs. The classification accuracy to label the subjects as KOA patient reaches up to 90% accuracy during gait but with a large day-to-day variation. The VAG classification accuracy was higher than the classification based on the pressure pain thresholds but this method showed lower day-to-day variation.

Conflict of interest statement

The authors did not receive financial support or other benefits from commercial sources for the work reported in this manuscript, or any other financial support that could create a potential conflict of interest or the appearance of a conflict of interest concerning the work.

Acknowledgments

This study was part of the individualized osteoarthritis interventions (IOI) project and has received funding from Innovation Fund Denmark (grant nr. 40-2014-3). The authors are grateful to CCBR (Aalborg, Denmark) for the recruitment of participants. The authors are also thankful to Neeraja Srimurugan Pratheep for performing PPT assessments.

References

1. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: A systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012;380:2163–96.
2. Wallace IJ, Worthington S, Felson DT, Jurmain RD, Wren KT, Maijanen H, et al. Knee osteoarthritis has doubled in prevalence since the mid-20th century. Proc Natl Acad Sci U S A. 2017;14(35):9332–9336.
3. Walters CF. THE VALUE OF JOINT AUSCULTATION. Lancet. 1929;213:920–1.
4. Andersen RE, Arendt-Nielsen L, Madeleine P. Knee joint vibroarthrography of asymptomatic subjects during loaded flexion-extension movements. Med Biol Eng Comput. 2018;56(12):2301–12.
5. Mascaro B, Prior J, Shark LK, Selfe J, Cole P, Goodacre J. Exploratory study of a non-invasive method based on acoustic emission for assessing the dynamic integrity of knee joints. Med Eng Phys. 2009;31(8):1013–22.
6. Wu Y, Chen P, Luo X, Huang H, Liao L, Yao Y, et al. Quantification of knee vibroarthrographic signal irregularity associated with patellofemoral joint cartilage pathology based on entropy and envelope amplitude measures. Comput Methods Programs Biomed. 2016;130:1–12.
7. Cardoso R, Porto P, Burin A, Daitx R, Dohnert M. Ground or Swimming Pool Exercises for Women with Knee Osteoarthritis? A Double-blind Randomized Clinical Trial. J Adv Med Med Res. 2017;23(10):1–13.
8. Suokas AK, Walsh DA, McWilliams DF, Condon L, Moreton B, Wylde V, et al. Quantitative sensory testing in painful osteoarthritis: A systematic review and meta-analysis. Osteoarthr Cartil. 2012;20(10):1075–85.
9. Arendt-Nielsen L, Nie H, Laursen MB, Laursen BS, Madeleine P, Simonsen OH, et al. Sensitization in patients with painful knee osteoarthritis. Pain. 2010;149:573–81.
10. Alburquerque-Sendín F, Madeleine P, Fernández-De-Las-Peñas C, Camargo PR, Salvini TF. Spotlight on topographical pressure pain sensitivity maps: A review. J Pain Res. 2018;11:215.
11. Srimurugan Pratheep N, Madeleine P, Arendt-Nielsen L. Relative and absolute test-retest reliabilities of pressure pain threshold in patients with knee osteoarthritis. Scand J Pain. 2018;18(2):229–36.
12. Andersen RE, Arendt-Nielsen L, Madeleine P. A Review of Engineering Aspects of Vibroarthography of the Knee Joint. Crit Rev Phys Rehabil Med. 2016;28(1_2):13–32.
13. Rangayyan RM, Wu Y. Analysis of vibroarthrographic signals with features related to signal variability and radial-basis functions. Ann Biomed Eng. 2009;37(1):156–63.
14. Wu Y, Yang S, Zheng F, Cai S, Lu M, Wu M. Removal of artifacts in knee joint vibroarthrographic signals using ensemble empirical mode decomposition and detrended fluctuation analysis. Physiol Meas. 2014;35(3):429.
15. Madeleine P, Andersen RE, Larsen JB, Arendt-Nielsen L, Samani A. Wireless multichannel vibroarthrographic recordings for the assessment of knee osteoarthritis during three activities of daily living. Clin Biomech. 2020;72:16–23.
16. Rangayyan RM, Oloumi F, Wu Y, Cai S. Fractal analysis of knee-joint vibroarthrographic signals via power spectral analysis. Biomed Signal Process Control. 2013;8:23–9.
17. Kim KS, Seo JH, Kang JU, Song CG. An enhanced algorithm for knee joint sound classification using feature extraction based on time-frequency analysis. Comput Methods Programs Biomed. 2009;94:198–206.

18. Kalo K, Niederer D, Sus R, Sohrabi K, Banzer W, Groß V, et al. The detection of knee joint sounds at defined loads by means of vibroarthrography. Clin Biomech. 2020;74:1–7.
19. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis. 1957;16(4):494.
20. Dobson F, Hinman RS, Roos EM, Abbott JH, Stratford P, Davis AM, et al. OARSI recommended performance-based tests to assess physical function in people diagnosed with hip or knee osteoarthritis. Osteoarthr Cartil. 2013;21(8):1042–52.
21. Binderup AT, Arendt-Nielsen L, Madeleine P. Pressure pain threshold mapping of the trapezius muscle reveals heterogeneity in the distribution of muscular hyperalgesia after eccentric exercise. Eur J Pain. 2010;14(7):705–12.
22. Diniz PSR. Adaptive filtering: Algorithms and practical implementation. Adaptive Filtering: Algorithms and Practical Implementation. New York: Springer; 2013.
23. Rangayyan RM. Introduction to Biomedical Signals. In: Biomedical Signal Analysis. 2nd ed. New Jersey: John Wiley & Sons Inc; 2010.
24. Hyvärinen A, Oja E. Independent component analysis: Algorithms and applications. Neural Networks. 2000;13:411–30.
25. Samani A, Srinivasan D, Mathiassen SE, Madeleine P. Nonlinear metrics assessing motor variability in a standardized pipetting task: Between-and within-subject variance components. J Electromyogr Kinesiol. 2015;25(3):557–64.
26. Jr CLW, Zbilut JP. Recurrence quantification analysis of nonlinear dynamical systems. Tutorials Contemp nonlinear methods Behav Sci. 2005;26–94.
27. Kleinberg EM. An overtraining-resistant stochastic modeling method for pattern recognition. Ann Stat. 1996;24(6):2319–49.
28. Hastie T, Tibshirani R, Friedman J. Elements of Statistical Learning 2nd ed. Springer. Springer; 2009.
29. Hutter F, Lücke J, Schmidt-Thieme L. Beyond Manual Tuning of Hyperparameters. KI - Künstliche Intelligenz. 2015;29(4):329–37.
30. Loh WY, Shin YS. Split selection methods for classification trees. Stat Sin. 1997;7:815–40.
31. Archer KJ, Kimes R V. Empirical characterization of random forest variable importance measures. Comput Stat Data Anal. 2008;52(4):2249–60.
32. Franke R, Nielson G. Smooth interpolation of large sets of scattered data. Int J Numer Methods Eng. 1980;
33. Adouni M, Shirazi-Adl A. Knee joint biomechanics in closed-kinetic-chain exercises. Comput Methods Biomech Biomed Engin. 2009;12(6):661–70.
34. Baczkowicz D, Kręcisz K, Borysiuk Z. Analysis of patellofemoral arthrokinematic motion quality in open and closed kinetic chains using vibroarthrography. BMC Musculoskelet Disord. 2019;20(1):48.
35. Tanaka N, Hoshiyama M. Vibroarthrography in patients with knee arthropathy. J Back Musculoskelet Rehabil. 2012;25:117–22.
36. Waterton JC, Solloway S, Foster JE, Keen MC, Gandy S, Middleton BJ, et al. Diurnal variation in the femoral articular cartilage of the knee in young adult humans. Magn Reson Med. 2000;43(1):126.
37. Denil M, Matheson D, De Freitas N. Narrowing the gap: Random forests in theory and in practice. In: 31st International Conference on Machine Learning, ICML 2014. 2014.
38. Kerrigan DC, Lelas JL, Goggins J, Merriman GJ, Kaplan RJ, Felson DT. Effectiveness of a lateral-wedge insole on knee varus torque in patients with knee osteoarthritis. Arch Phys Med Rehabil. 2002;83:889–93.
39. Collins SL, Moore RA, McQuay HJ. The visual analogue pain intensity scale: What is

moderate pain in millimetres? Pain. 1997;72(1–2):95–7.
40. Dong B, Kong Y, Zhang L, Qiang Y. Severity and distribution of cartilage damage and bone marrow edema in the patellofemoral and tibiofemoral joints in knee osteoarthritis determined by MRI. Exp Ther Med. 2017;13(5):2079–84.
41. Vina ER, Kwoh CK. Epidemiology of osteoarthritis: literature update. Curr Opin Rheumatol. 2018;30(2):160–7.
42. Lewek MD, Scholz J, Rudolph KS, Snyder-Mackler L. Stride-to-stride variability of knee motion in patients with knee osteoarthritis. Gait Posture. 2006;23(4):505–11.
43. Hair JF, Anderson RE, Tatham RL, Black WC. Multivariate data analysis. 5th ed. Prentice Hall, Englewood Cliffs, NJ; 1998. 730–787 p.
44. Hochreiter S, Schmidhuber JJ. Long short-term memory. Neural Comput. 1997;9(8):1–32.
45. Reddy BK, Delen D. Predicting hospital readmission for lupus patients: An RNN-LSTM-based deep-learning methodology. Comput Biol Med. 2018;101:199–209.
46. Ching T, Himmelstein DS, Beaulieu-Jones BK, Kalinin AA, Do BT, Way GP, et al. Opportunities and obstacles for deep learning in biology and medicine. J R Soc Interface. 2018;15(141):20170387.