



Association between central sensitization and gait in chronic low back pain: Insights from a machine learning approach



Xiaoping Zheng^{a,*}, Michiel F. Reneman^b, Jone Ansuategui Echeita^b, Rita HR. Schiphorst Preuper^b, Herbert Kruitbosch^c, Egbert Otten^a, Claudine JC. Lamoth^a

^a University of Groningen, University Medical Center Groningen, Department of Human Movement Sciences, Groningen, Netherlands

^b University of Groningen, University Medical Center Groningen, Center for Rehabilitation, Groningen, the Netherlands

^c University of Groningen, Center for Information Technology, Data Science Team, Groningen, the Netherlands

ARTICLE INFO

Keywords:

Low back pain
Central sensitization
Supervised machine learning
Gait
Daily life

ABSTRACT

Background: Central sensitization (CS) is often present in patients with chronic low back pain (CLBP). Gait impairments due to CLBP have been extensively reported; however, the association between CS and gait is unknown. The present study examined the association between CS and CLBP on gait during activities of daily living.

Method: Forty-two patients with CLBP were included. CS was assessed through the Central Sensitization Inventory (CSI), and patients were divided in a low and high CS group (23 CLBP- and 19 CLBP+, respectively). Patients wore a tri-axial accelerometer device for one week. From the acceleration signals, gait cycles were extracted and 36 gait outcomes representing quantitative and qualitative characteristics of gait were calculated. A Random Forest was trained to classify CLBP- and CLBP+ based on the gait outcomes. The maximum Youden index was computed to measure the diagnostic test's ability and SHapley Additive exPlanations (SHAP) indexed the gait outcomes' importance to the classification model.

Results: The Random Forest accurately (84.4%) classified the CLBP- and CLBP+. Youden index was 0.65, and SHAP revealed that the gait outcomes' important to the classification model were related to gait smoothness, stride frequency variability, stride length variability, stride regularity, predictability, and stability.

Conclusions: CLBP- and CLBP+ patients had different motor control strategies. Patients in the CLBP- group presented with a more "loose control", with higher gait smoothness and stability, while CLBP+ patients presented with a "tight control", with a more regular, less variable, and more predictable gait pattern.

1. Introduction

Chronic low back pain (CLBP) is one of the most prevalent chronic musculoskeletal pains [1]. It is responsible for high treatment costs, sick leave and individual suffering and it represents a significant socioeconomic burden [2]. For 85%–90% of patients with CLBP, the relation between pathoanatomical and clinical presentations is absent [3] and, therefore, it is classified as nonspecific CLBP [4]. In CLBP, and other chronic musculoskeletal disorders, central sensitization (CS) might be present (reviewed in Ref. [5]). CS is defined as "increased responsiveness of nociceptive neurons in the central nervous system to their normal or subthreshold afferent input" [6] and manifests as mechanical hypersensitivity, allodynia and hyperalgesia [7]. A considerable number of people need treatment for CLBP. Although the overall efficacy of CLBP rehabilitation programs is positive, but the effect sizes are modest [8].

Correctly recognizing the physical and psychosocial factors perpetuating pain and physical disability of patients with CLBP remains a challenge [9]. Altered motor control of patients with CLBP could possibly contribute to the persistence of CLBP [10]. Altered motor control could affect daily-living activities, as patients with CLBP often exhibit altered movement patterns and motor control strategies; probably to avoid painful movement, such as walking [11]. Many clinicians may intuitively identify "abnormal" gait patterns in patients with CLBP, but identification and objectifying of specific "abnormal" gait outcomes is challenging. During walking, it is suggested that patients often adopt a "protective guarding" or "splinting" strategy [12] to avoid painful movements of the spine. These adaptations may lead to a slower and less flexible gait pattern [13]. Evidence for this, however, is ambiguous. Studies between patients with CLBP and healthy controls, observed inconsistent evidence regarding preferred walking velocity [13,14],

* Corresponding author.

E-mail address: x.zheng@umcg.nl (X. Zheng).

<https://doi.org/10.1016/j.combiomed.2022.105329>

Received 20 October 2021; Received in revised form 14 February 2022; Accepted 14 February 2022

Available online 16 February 2022

0010-4825/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

stride length [15,16], and stride-to-stride variability [17,18].

A possible explanation for these inconsistencies might be an unknown heterogeneity within the samples, such as the presence of CS. CS could plausibly be related to the inconsistent results, because the presence of high CS levels is associated with long-lasting chronic pain [19] and movement may be changed due to pain. Also, general gait outcomes such as walking speed and stride length, might not be sensitive enough to detect small differences between patients with low or high levels of CS. In addition to stride related parameters, gait outcomes that reflect gait quality in terms of regularity, synchronization, smoothness, local stability, and predictability, are sensitive to detect differences in gait performance. These gait outcomes were successfully used to detect the differences between age groups [20], older adults with and without fall risk [21], and patients with and without Parkinson's disease [22]. Even though the effects of CLBP on gait have been frequently investigated in controlled laboratory studies, there are no studies about the relationship between CS levels and gait performance under daily-living environment circumstances.

Advances in wearable technology and machine learning approaches offer new opportunities in gait data collection and analysis. Wearable technology allows researchers to record patients' physical activities in unobserved, daily-living environments over extended periods of time. This data can reflect the real gait performance of the patients, since being observed may change the performance of patients under the controlled laboratory environment [23]. The successful employment of machine learning approaches in gait analysis makes it possible to extract the most informative gait outcomes from the accelerometer sensor data [20]. If patients with low and high levels of CS walk differently, machine learning approaches will be able to successfully recognize these differences and can classify patients with low and high CS level based on their gait outcomes. Many gait outcomes are not independent and interact with each other, such as gait speed and step regularity. Machine learning approaches such as Random Forest (RF), are able to process high dimensional and non-linear data structures and take the interrelation and interaction of the gait outcomes into consideration [20].

Therefore, the aim of this study was to analyze whether and how the presence of CS is related to differences in gait performance of patients with CLBP during daily life by using a machine learning approach. It was hypothesized that patients with CLBP and higher CS levels show differences in daily life gait performance, compared with those with lower CS levels.

2. Methods

2.1. Patients

This study included patients with primary CLBP who were recruited from the outpatient Pain Rehabilitation Department of the Center for Rehabilitation of the University Medical Center Groningen (CvR-UMCG). Primary CLBP is defined as low back pain persistent for more than three months, with pain not being the result of any other diagnosis. The patients were selected according to the following inclusion criteria: (a) age between 18 and 65 years old at the time of recruitment; (b) admitted to the interdisciplinary pain rehabilitation program; (c) could follow instructions; (d) signed informed consent. Additionally, patients were excluded if they: (a) had a specific diagnosis that would better account for the symptoms (e.g. cancer, inflammatory diseases and/or spinal fractures); (b) had neuralgia and/or radicular pain in the legs; (c) were pregnant; (d) in an acute phase of pain.

The study was approved by the Medical Research Ethics Committee of the University Medical Center Groningen (METc 2016/702) and conducted according to the principles expressed in the Declaration of Helsinki. The data used in this paper was derived from a larger study, of which protocol details were described elsewhere [19].

2.2. Data collection

Demographics were collected and standard clinical test were applied as part of the usual care of CLBP patients that are referred to the outpatient Pain Rehabilitation Department of the Center for Rehabilitation. Assessments included: Visual Analogue Scale for pain intensity (VAS Pain; 0–10), the Dictionary of Occupational Titles (DOT, the Pain Disability Index (PDI; 0–70), the physical functioning subscale of the Rand36 questionnaire (Rand36-PF; 0–100), the Pain Catastrophizing Scale (PCS, 0–52), the Injustice Experience Questionnaire (IEQ, 0–48), and the Brief Symptom Inventory (BSI global severity index t-score (GSIT)) (see Table 3).

Central sensitization (CS). The presence of CS-related manifestations was assessed with section A of the Central Sensitization Inventory (CSI) [24]. Section A has 25-items to assess the presence of common CS-related symptoms. Scores can range from 0 to 100 where a higher scoring represents a higher level of CS. A score lower than 40 indicates lower CS levels (CLBP- group) and a score of 40–100 is interpreted as higher CS levels (CLBP + group) [25].

Accelerometer data. The accelerometer data were collected between 2017 and 2019. Patients were instructed to wear a tri-axial accelerometer (ActiGraph GT3X, Actigraph Corporation, Pensacola, FL) at all times for about one week, excluding sleeping or bathing times. The accelerometer was worn at the front right hip of the patient (at the anterior superior iliac spine). Assuming a standing and upright position, the Y-axis pointed to the ground (vertical direction, V), Z-axis faced the walking direction (anteroposterior direction, AP), and the X-axis was perpendicular to the walking direction, pointing from a patient's right to left (mediolateral direction, ML). These directions are approximate only. The sampling frequency of the accelerometer was set to 100 Hz and the dynamic range was ± 6 gravity.

2.3. Data processing and analysis

2.3.1. Raw data segmentation

Accelerometer data of each patient was segmented into 24 h span data segments (from 12:00 p.m. to next day 11:59 a.m.) to represent the activities during the days. Because the measurement started at 12:00 p.m., to make full use of the data, the 24 h span was between 12:00 p.m. until next day 12:00 p.m. Data that did not completely covered this 24 h span was discarded from the analysis. Because of technical errors or personal reasons, a full week of data could not be collected from all patients. To compare the data between different patients fairly, 4 segments (representing 4 days) of each patient were included in the analysis. Therefore, 7 patients who had less than 4 segments, were excluded. From patients with more than 4 segments, 4 segments were randomly sampled. Fig. 1a graphically shows the process of the raw data segmentation.

2.3.2. Walking bouts extraction

The accelerometer data of the 4 segments were first smoothed by a low-pass filter with a 2nd order Butterworth and a 20 Hz cut-off frequency. Subsequently, potential walking events were detected by the Fast Fourier Transform (FFT) based method [26], which identified periods with 0.5–3.0 Hz power spectrum values. To remove false walking events from the potential walking periods, the zero-cross method [27] was employed. If the time interval between any two adjacent walking events was shorter than 2 s, these two walking events were merged into one walking bout. Finally, the walking bouts in each segment were extracted and their gait outcomes were calculated. Fig. 1b presents the walking bouts as the yellow vertical bars in the rectangle.

2.3.3. Gait outcomes

All walking bouts in one 24 h segment were used to determine the total duration of walking, the total number of steps, the maximum duration of a walking bout and the maximum number of steps of a

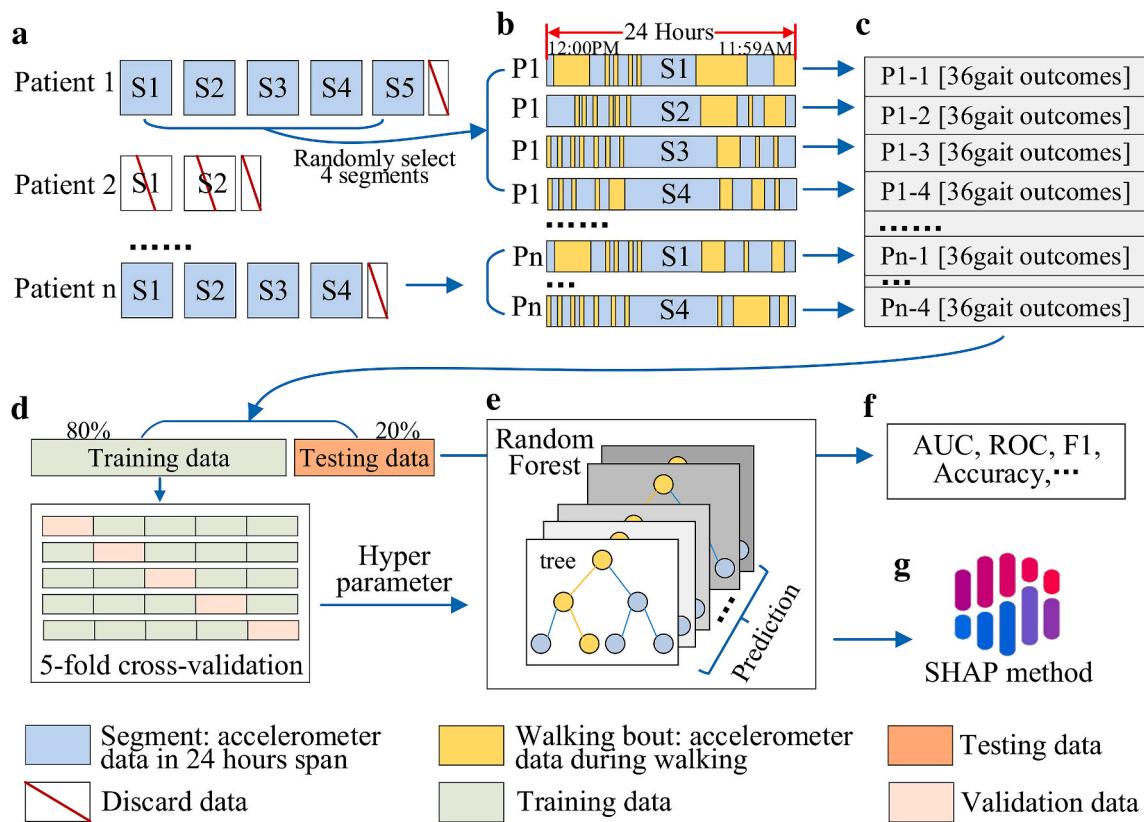


Fig. 1. The data processing and analysis: (a) raw data segmentation, (b) walking bouts extraction, (c) gait outcome vectors, (d) training and testing data preparation, (e) Random Forest classifier, (f) accuracy evaluation, (g) feature importance.

walking bout. Subsequently, all walking bouts exceeding 10 s were selected and cut into non-overlapping 10 s windows [28]. From the segment, each 10 s window was used to calculate different gait outcomes, and these values were averaged over all 10 s windows in the segment representing the patient's gait performance on that day.

Gait outcomes were divided into two categories, quantitative and qualitative gait outcomes. From one segment, we obtained one gait outcome vector, including 36 gait outcomes, based on the walking bouts (see Fig. 1c). The detailed descriptions of the quantitative and qualitative gait outcomes are presented in Table 1 and Table 2 – for extended explanation of variables see Ref. [29].

Pearson-coefficient was calculated to examine relationship of gait outcomes between weekdays and weekend. The Pearson-coefficient ranges from -1 to 1, where 1 represents a perfect correlation.

The Mann-Whitney *U* test was used to statistically test the differences between CLBP- and CLBP + groups for demographics and CSI scores. To separate CLBP- and CLBP + groups by gait outcomes, RF was used.

2.3.4. Random Forest classifier

RF is considered as the optimal machine learning classification approach for the present data, because it performs well with (a) nonlinear and linear data; (b) high dimensional data; and (c) unbalanced and small datasets [30]. Apart from this, a comparison of different machine learning classifiers was performed to help to select RF as the best classifier for this study (details in Appendix A).

The input data of this approach was $\langle S, L \rangle$. S represents the gait outcome vectors of all patients and L was its corresponding label. The definition of S is: $S = \{s_1, s_2, \dots, s_i, \dots, s_m\}$ and $s_i = [d_1, \dots, d_k]$, where s_i represents a gait outcome vector i and m is the number of all gait outcome vectors, d represents a gait outcome and $k = 36$. $L = l_1, \dots, l_m$, where $l \in \{\text{CLBP} -, \text{CLBP} +\}$.

RF is constructed in four steps. Step one: Randomly sample n gait

Table 1
Quantitative gait outcomes.

Catalog	Gait characteristic	Description and method
Pace	Total duration of walking in the day Total number of steps in the day Maximum duration of a walking bout Maximum number of steps of a walking bout Walking speed (WS; mean, variability)	The accumulated time (in seconds) of the walking bouts in one segment. The accumulated steps of walking bouts in one segment. Duration (in seconds) of longest walking bout in one segment. Maximum number of steps of one walking bout in one segment. $WS = D/T$, where D is the distance (in meters) and equals to the accumulated of step length; T is the corresponding time (in seconds).
	Stride length (SL; mean, variability)	$SL = 2\sqrt{2lh - h^2}$, where h is the change in height (in meters), l equals leg length (in meters). h was calculated by a double integration of the accelerometer signal in vertical direction. SL is the sum of the adjacent two step lengths.
	Stride time (ST; mean, variability) Stride frequency (SF; mean, variability-V/ML/AP)	$ST = n/f$, where f is the sample frequency (in Hertz) and n is the number of samples per dominant period derived from autocorrelation. $SF = f/n$.
	Root mean square of the variability of the amplitude of accelerations (RMS)	$RMS = \sqrt{\frac{1}{n}(x^2 + y^2 + z^2)}$, where x, y, z represent the accelerometer signal (in meters per second squared) in x, y, z axis and n is the number of samples.

outcome vectors from S and n corresponding labels from L , with replacement. These new set of gait outcome vectors and labels are called S_b and L_b . In S_b , s_i may appear more than one time or not appear. Step

Table 2
Qualitative gait outcomes.

Catalog	Gait characteristic	Description and method
Regularity	Stride regularity (SR; V, ML, AP, All)	SR is computed by using the unbiased autocorrelation coefficient: $Ad(m) = \frac{1}{N - m } \sum_{i=1}^{N- m } Acc(i).Acc(i + m)$, where is the sample acceleration signal, $Acc(i)$ the number of samples, and N the number of time lag. The first peak of m is $Ad(m)$ and it represents the stride regularity. Higher values (maximum 1.0) reflect repeatable patterns between strides.
	Gait symmetry index (GSI)	GSI quantifies the ratio of the first and second peak of the Ad_m , as Ad_1/Ad_2 . It is a measure of the degree of symmetry of the left and right lower limbs during walking.
Smoothness	Index of harmonicity (IH; V, ML, AP, All)	$IH = \frac{P_0}{\sum_{i=0}^6 P_i}$. It is the ratio of the power spectral density of the fundamental frequency P_0 and the sum of the power spectral density of the first six frequency P_i . IH quantifies gait smoothness., with higher values representing a smoother (max 1.0) gait pattern.
	Harmonic ratio (HR; V, ML, AP)	$HR = \frac{\sum P_a}{\sum P_b}$. In VT and AP directions, $\sum P_a$ = the sum of even power spectral and $\sum P_b$ = the sum of odd power spectral. In ML direction, P_a is odd and P_b is even. It reflects the rhythmicity of the walking patterns. Higher values mean more rhythmic.
Predictability	Sample entropy (Sen; V, ML, AP)	$Sen = -\ln\left(\frac{A}{B}\right)$, with $A = d[Acc_{m+1}(i), Acc_{m+1}(j)]\langle r, B = d[Acc_m(i), Acc_m(j)]\langle r$. $Acc_m(i)$ means the accelerometer signal vector from time i to $m + i - 1$. $d[\cdot]$ is the Chebyshev distance, and r was set to 0.3. Sen quantifies the predictability of a time series. Smaller values (minimum 0) indicate better synchronization between acceleration signals.
Stability	Maximal Lyapunov exponent (max LyE; V, ML, AP)	Max LyE, as calculated by the Rosenstein algorithm, quantifies the local stability of trunk acceleration patterns. The fitting window length was $60/100^*f$, where f is the sample frequency, and the embedding dimension was set to 7. The overall max LyE were calculated and normalized by per stride. Higher values represent greater sensitivity to local perturbations.
	Maximal Lyapunov exponent normalized per stride by time (max LyE per stride; V, ML, AP)	Max LyE, as calculated by the Rosenstein algorithm, quantifies the local stability of trunk acceleration patterns. The fitting window length was $60/100^*f$, where f is the sample frequency, and the embedding dimension was set to 7. The overall max LyE were calculated and normalized by per stride. Higher values represent greater sensitivity to local perturbations.

two: In S_b , randomly sample j ($j \leq k$) gait outcomes from s . Therefore, $s'_i = d'_1, \dots, d'_j$ and $S'_b = s'_1, \dots, s'_n$. Step three: Training a decision tree f_b on S'_b, L_b . Step four: Repeat steps one to three 1000 times and combine the decision trees into an ensemble, called RF, that predicts by voting (see Fig. 2).

Before training RF, 80% of patients were randomly selected and their 4 corresponding gait outcome vectors were used as the training data. The gait outcome vectors of the remaining 20% of patients were used as the testing data. To avoid overfitting of the hyperparameters, a 5-fold cross-validation approach was used to estimate them, as shown in Fig. 1d. Four folds were used to train the model and the rest fold was used to estimate the performance of the current hyperparameters in RF. The performance reported by the 5-fold cross-validation was the average of the values computed in the 5 splits. After the best hyperparameters were determined, the testing dataset was used to evaluate the

generalizability of the model.

2.3.5. Accuracy evaluation

Accuracy, sensitivity, specificity, precision, F1-score, and maximum Youden index were calculated to evaluate the performance of the classification (Fig. 1f). In this study, CLBP+ was considered as the positive case and CLBP- was the negative case. Correct predictions of CLBP+ and CLBP- patients are called true positives (TP) and true negatives (TN), respectively. Incorrect classifications of CLBP- patients as CLBP+ or of CLBP+ patients as CLBP-, are called false positives (FP) and false negatives (FN) respectively.

Accuracy was the proportion of all the correct classification results.

$$\text{accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

Sensitivity represents the proportion of positive cases that are correctly assigned (true positive rate).

$$\text{sensitivity} = \frac{TP}{TP + FN} \quad (2)$$

Specificity refers to the rate of correctly predicted negative cases in all negative cases (true negative rate).

$$\text{specificity} = \frac{TN}{TN + FP} \quad (3)$$

Precision is the ratio of the correctly predicted positive cases in all predicted positive cases.

$$\text{precision} = \frac{TP}{TP + FP} \quad (4)$$

F1-score is the harmonic mean (average) of the precision and sensitivity.

$$F1 = \frac{2 \times \text{precision} \times \text{sensitivity}}{\text{precision} + \text{sensitivity}} \quad (5)$$

The receiver operating characteristic (ROC) curve was calculated to evaluate the performance of RF. The Y-axis of this curve represents the true positive rate (sensitivity) and the X-axis means false positive rate (1-specificity). The overall classification performance of RF was evaluated by the area under the ROC curve (AUC). A classification model with a larger AUC value has a higher correct rate, and $AUC = 1$ represents perfect performance. The maximum Youden index was computed to measure the diagnostic test's ability.

$$J(c) = \text{Max}\{\text{sensitivity}(c) - (1 - \text{specificity}(c))\} \quad (6)$$

where c is the cut-point. When the value J is maximum, the corresponding c is the optimal cut-point.

2.3.6. Feature importance

SHapley Additive exPlanations (SHAP) [31] was used to assess the gait outcomes' importance to the classification model. SHAP connects optimal credit allocation with local explanations using the classic Shapley values from game theory. Shapley values, \emptyset_i , explains the importance of gait outcome i for RF and is defined as:

$$\emptyset_i = \frac{1}{|N|!} \sum_{\{i\} \in s \text{ and } s \subseteq N} (|s| - 1)!(|N| - |s|)! [R(s) - R(s - \{i\})] \quad (7)$$

where N is the size of the full set of gait outcomes, s is the subset that includes i in N , and $R(\cdot)$ is the accuracy of RF of the input gait outcomes. Since computing the exact Shapley values is computationally expensive, SHAP uses a tree explainer to exploit the information stored in the tree structure to calculate the SHAP values which are highly approximate Shapley values. Therefore, higher SHAP values represent higher impact to classify CLBP- and CLBP+ groups.

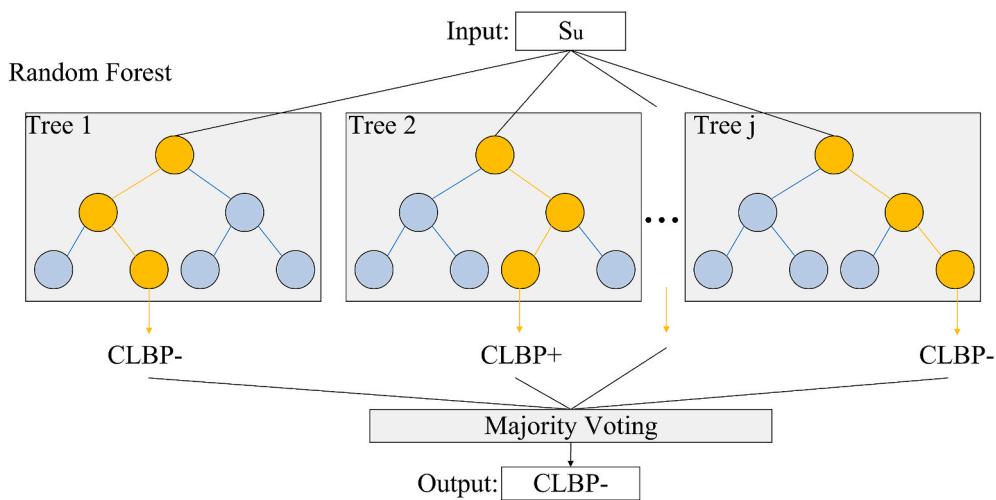


Fig. 2. Architecture of the Random Forest classifier.

3. Results

Demographic characteristics are provided in Table 3. Out of a total of 60 patients, 11 were excluded because essential parts of their dataset were incomplete (CSI scores or/and accelerometry data), 7 were

Table 3
Patient characteristics (n = 42).

	CLBP- (n = 23)	CLBP+ (n = 19)	All (n = 42)	P-Value
Gender	15 W/8 M	12 W/7 M	27 W/15 M	
Age, years	40.8 ± 12.8	38.1 ± 12.7	39.6 ± 12.6	
Height, cm	173.5 ± 10.6	175.7 ± 8.8	174.5 ± 9.8	
Weight, kg	87 ± 17.7	85.4 ± 15.1	86.3 ± 16.4	
Body mass index, kg/m ²	28.9 ± 5.3	27.7 ± 4.4	28.3 ± 4.9	
Central Sensitization Inventory (0–100)	31 ± 4.8	48.7 ± 8.7	39.0 ± 11.2	<0.0001
Time since pain onset (years)	4.5 ± 6.1	3.5 ± 3.1	4.1 ± 4.9	
Educational Level	17S/6H	10S/9H	26S/15H	
Physical demands at work (DOT; Se/Li/Me/He)	3/11/8/1	4/7/7/1	7/18/15/2	
Patient-reported Pain Intensity (VAS, 0–10)	5.5 ± 2	5.2 ± 1.8	5.4 ± 1.9	
Disability (PDI, 0–70)	33.6 ± 11.2	26.8 ± 11.9	31.0 ± 11.7	
Work Ability (WAS, 0–10)	4.5 ± 2.3	4.9 ± 2.8	4.6 ± 2.5	
Physical Functioning (Rand36-PF, 0–100)	49.8 ± 22.3	63.3 ± 16.1	54.7 ± 21.1	
Catastrophizing (PCS, 0–52)	16.3 ± 8.9	20.3 ± 11.1	18.1 ± 10	
Injustice (IEQ, 0–48)	15.2 ± 8.9	18.5 ± 8.5	16.7 ± 8.8	
Psychological traits Screening (BSI, t-score)	34.4 ± 4.9	41.5 ± 5.8	37.6 ± 6.4	= 0.01

Except gender, all results represent mean ± standard deviation. CLBP-, CLBP+: Patients with chronic low back pain with low (−) and high (+) central sensitization levels. W: Women; M: Men. H: Higher education; S: Secondary education. Se: Sedentary; Li: Light; Me: Medium; He: Heavy. DOT: Dictionary of Occupational Titles. VAS: Visual Analogue Scale. PDI: Pain Disability Index. WAS: Work Ability Score. Rand36-PF: Rand 36-Physical Functioning subscale. PCS: Pain Catastrophizing Scale. IEQ: Injustice Experience Questionnaire. BSI: Brief Symptom Inventory.

excluded because they had less than 4 segments data (3 had 1 segment, 2 had 2 segments, and 2 had 3 segments). Therefore, 42 patients were included in the data analysis. Differences between CLBP+ and CLBP-group characteristics (Table 3) were not statistically significant ($p > 0.05$), with exception of CSI score ($p < 0.001$) and BSI ($p = 0.01$).

Because 42 patients (23 CLBP- and 19 CLBP+) were included, and for every patient 4 segments were randomly selected, the total accelerometer data segments were 168. Therefore, the scales of training and testing dataset were 136 and 32. The mean Pearson-coefficient between workdays and weekend was 0.983, indicating almost perfect correlation.

Testing data were used to evaluate the generalizability of RF and the confusion matrix is shown in Fig. 3. From the confusion matrix, accuracy, sensitivity, specificity, precision, and the F1-score were calculated to evaluate the performance metrics of the model. RF achieved an accurate classification-result (84.4% accuracy), and the sensitivity and specificity were 75.0% and 93% respectively. The precision was 92% and the F1-score was 82.6%. The ROC curve is presented in Fig. 4 showing that RF achieved a 0.83 AUC and the maximum Youden index was 0.69.

The importance of the gait outcomes for RF is shown in Fig. 5. Based on the SHAP values, the 10 gait outcomes (above the red line in Fig. 5) were considered as important to the classification model. For the gait outcomes below the red line, the SHAP values were too low. Important gait outcomes are IH-V, SF variability-ML/AP, SR-ML, Max LyE-V/ML, Sen-AP, Max LyE per stride-V, HR-ML and SL variability.

Fig. 6 shows the violin-box plot of the 10 important gait outcomes. Violin-box plot is a hybrid of a kernel density plot and a box plot, and the dots show the individuals data. A box plot contains a set of whiskers, a box and a horizontal line in the middle of the box, representing the minimum, maximum, first quartile, third quartile and median of the data respectively. From this figure, it is easy to distinguish the differences of the median between groups. It shows that CLBP- group has higher IH-V, HR-ML (better smoothness); higher SF-variance-ML, SF-variance-AP, SL-variance (higher variability); lower SR-ML (lesser regularity), lower Max LyE-V, Max LyE-per-stride-V, slightly lower Max LyE-ML (better stability); and slightly higher Sen-AP (lesser predictability). Although the differences of medians between 2 groups in Sen-AP and Max LyE-ML are small, their distributions are different. In Sen-AP, data of CLBP- had a wider distribution and CLBP + shows more data at the bottom. In the Max LyE-ML, data of CLBP- is concentrated around the median, while CLBP + has a wide distribution and a lower peak. For other gait outcomes, the distributions are also different. In IH-V, distributions of CLBP- and CLBP + all showed a bimodal distribution, but the peaks of distribution are different. In SF Variability-ML and SF Variability-AP, CLBP + has a larger peak at the bottom while CLBP- has a

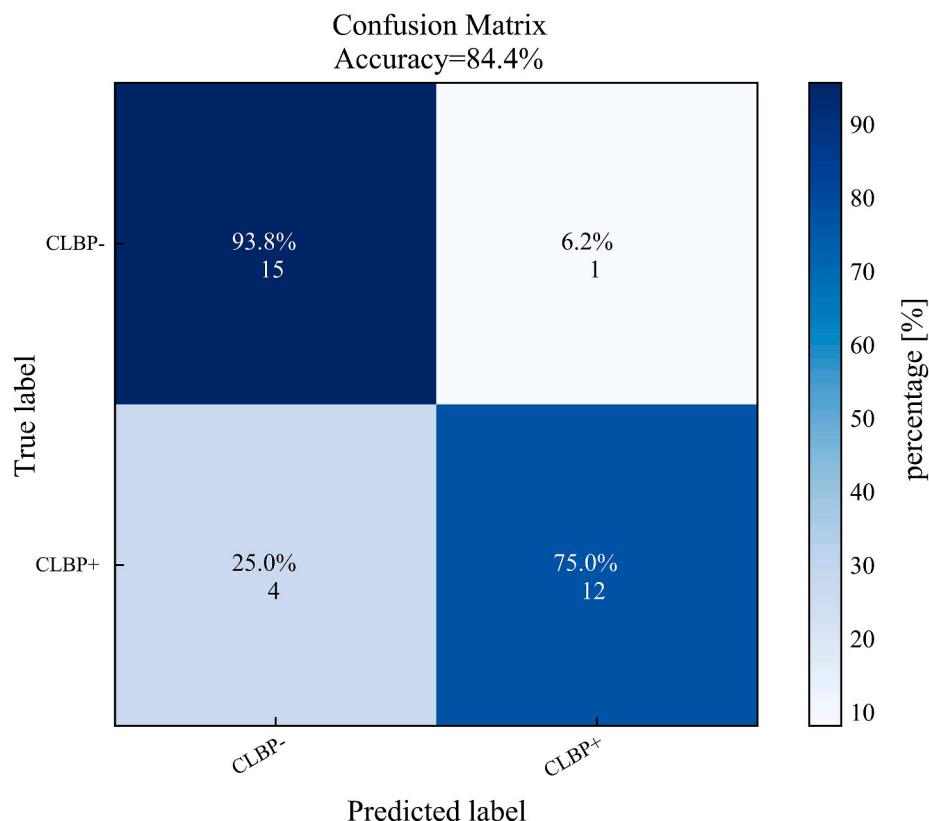


Fig. 3. Classification results for Random Forest, and the mean accuracy was 84.4%. CLBP-, CLBP+: Patients with chronic low back pain with lower (–) and higher (+) central sensitization levels.

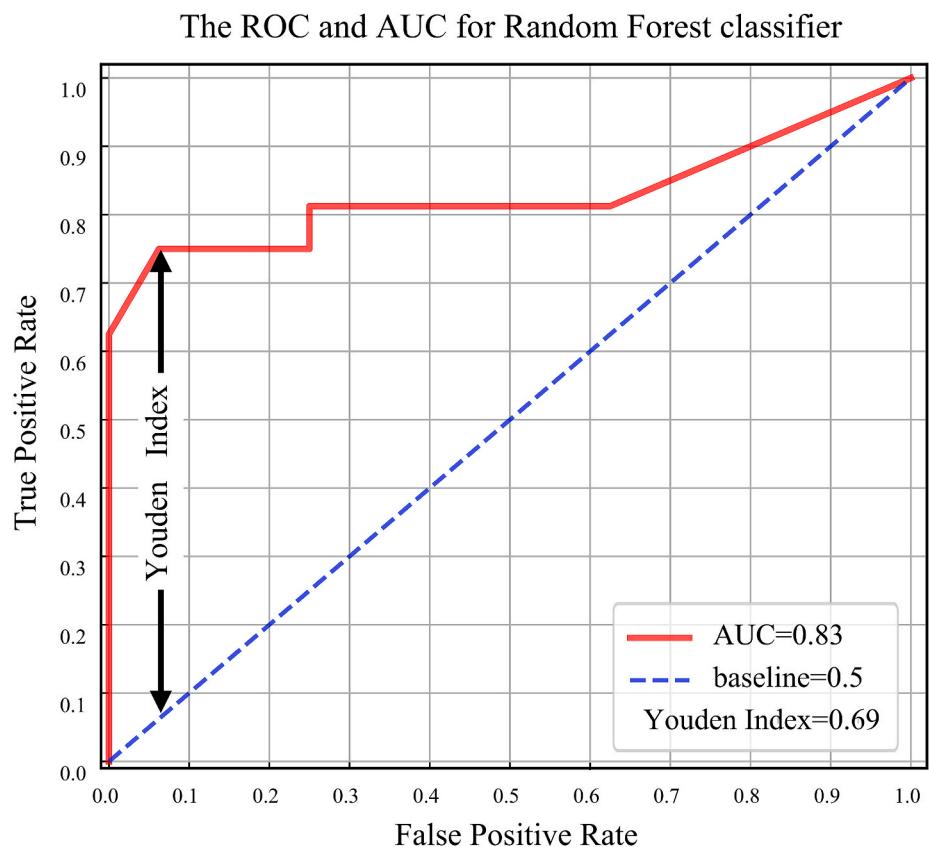


Fig. 4. The receiver operating characteristic (ROC) curve (in red) for Random Forest classifier. AUC: area under the curve.

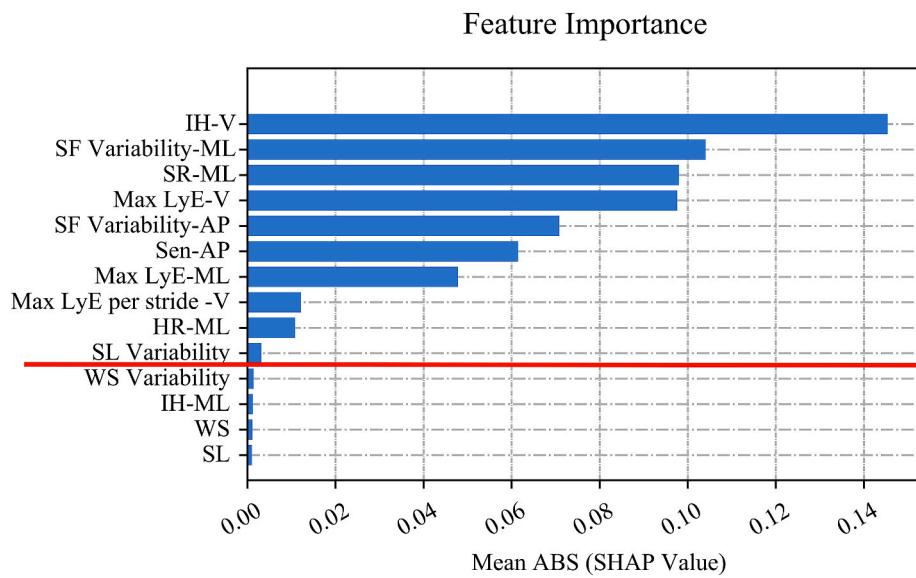


Fig. 5. Features importance of Random Forest classifier. The 10 gait outcomes above the red line are: index of harmonicity in vertical direction (IH-V), variability of stride frequency in mediolateral/antero-posterior direction (SF variability-ML/AP), stride regularity in mediolateral direction (SR-ML), Maximal Lyapunov exponent in vertical/mediolateral direction (Max LyE-V/ML), sample entropy in anteroposterior direction (Sen-AP), Max LyE-V: Maximal Lyapunov exponent per stride in vertical direction, harmonic ratio in mediolateral direction (HR-ML) and variability of stride length (SL variability). The remaining gait outcomes below the red line are: WS variability: variability of walking speed, IH-ML: index of harmonicity in mediolateral direction, WS: mean walking speed and SL: mean stride length. ABS: absolute value. SHAP: SHapley Additive exPlanations.

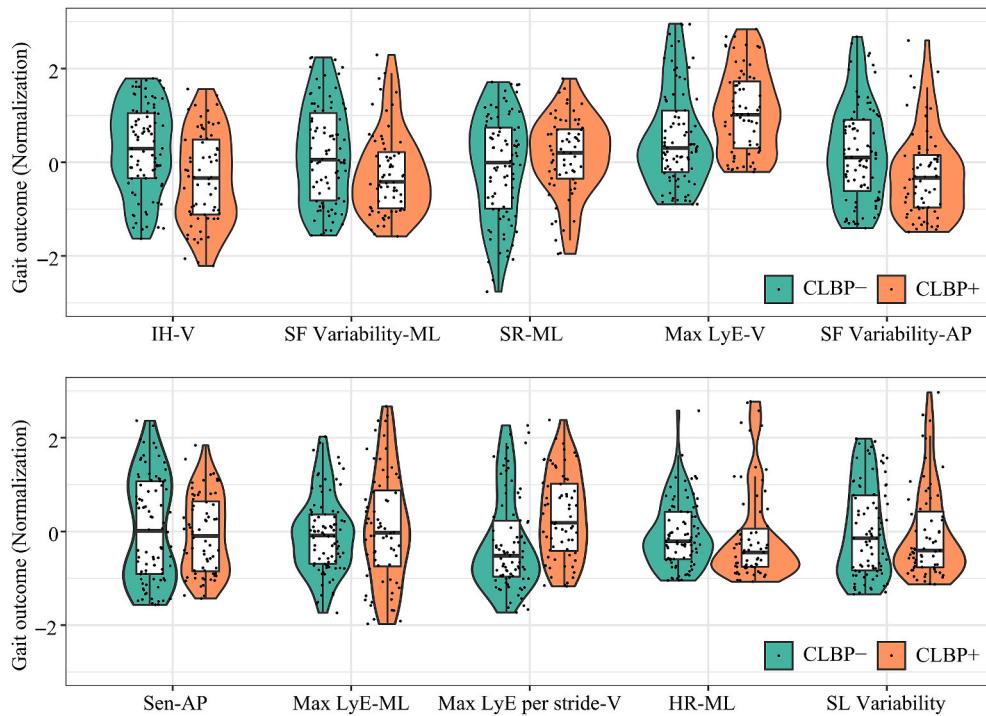


Fig. 6. Violin-box plot for the 10 gait outcomes. Dots show the individuals data. CLBP-, CLBP+: Patients with chronic low back pain with low (-) and high (+) CS levels. IH-V: index of harmonicity in vertical direction, SF variability-ML/AP: variability of stride frequency in mediolateral/antero-posterior direction, SR-ML: stride regularity in mediolateral direction, Max LyE-V/ML: Maximal Lyapunov exponent in vertical/mediolateral direction, Sen-AP: sample entropy in anteroposterior direction and HR-ML: harmonic ratio in mediolateral direction.

wide range distribution. Similarly, in SR-ML, CLBP + has a concentrating distribution while the peak of CLBP- is lower. In Max LyE-V and Max LyE per stride -V, CLBP- shows a log-normal distribution while CLBP + shows a wider distribution. In HR-ML and SL Variability, the distributions are similar but CLBP + has more outliers.

4. Discussion

The aim of this study was to analyze whether and how the presence of CS was related to differences in gait performance of patients with CLBP during daily life by using a machine learning approach. Based on quantitative and qualitative gait outcomes, using a RF, the two groups (CLBP- and CLBP+) could be classified with a high accuracy. The classification results indicated that CLBP- patients walk differently from

CLBP + patients. Furthermore, the SHAP values showed that the differences between CLBP- and CLBP + groups were present in gait outcomes that represented smoothness, stability, predictability, regularity, and variability.

In the present study, we addressed the walking measurement of patients with CLBP in a daily-living environment. Walking in a controlled laboratory or during a clinical assessment is different from self-initiated gait, during activities of daily living. Walking in daily life, might be subject to environmental perturbations, quick changes while performing a task, and often involves the performance of several actions at the same time [32], e.g. walking when carrying a cup of coffee. These influences on gait are not present in controlled studies and are not captured by conventional gait outcomes that average outcomes over stride cycles, such as mean step length, mean step time, and number of steps.

Therefore, the present study included gait outcomes that take into account the interdependency of gait cycles and how gait cycles evolve over time, e.g., using sample entropy as a measure of predictability of the gait pattern, the maximal Lyapunov exponent as quantification of local stability and correlation-based measures [33].

The accuracy value of 84.4% shows that RF has a high classification accuracy. The specificity scores of RF reveals that 93% of the samples (15 samples, true negative) are correctly classified as member of the CLBP- without a high CS level, but it misses 7% (1 sample, false positive). The sensitivity scores show that 75% of the samples (12 samples, true positive) of the CLBP + group were assigned to this group, however 25% were wrongly classified as belonging to the CLBP- group (4 samples, false negative). Decreasing the possibility of false positive will increase the possibility of false negative, and vice versa. The F1-score was calculated to take false positive and false negative into consideration at the same time by computing their harmonic mean. The high F1-score (82.6%) of RF implies that the model has a good and balanced performance. The Youden Index (0.69) was higher than 0.5 which means that RF has a diagnostic test's ability to balance sensitivity and specificity. The AUC indicates that RF has a 83% chance to distinguish CLBP+ and CLBP- correctly. Based on these performance measures of RF, this study leads us to concluded that the CLBP- and CLBP + had different gait patterns, and that the gait outcomes important to the classification model identified by SHAP are trustworthy.

In the present study RF was applied for classification, among the many available machine learning approaches, such as K-nearest neighbors (KNN), Naive Bayes (NB), Artificial Neural Network (ANN), Support Vector Machine (SVM). In general, machine learning approaches can take the interaction of gait outcomes into consideration. KNN and NB are instance-based learning approaches which imply they do not learn from training data [34]. Our choice for RF was based on the results of a previous study that compared RF, ANN, and SVM to classify different age groups on similar gait outcomes. The results of this study showed that all approaches had a good overall classification accuracy [20]. Moreover, for the current dataset, our preliminary empirical work in which we compared the performance of different machine learning classifiers, showed that RF and ANN had the best performance compared to SV, NB and KNN (details were in Appendix A). A drawback of ANN is that it requires a large data set to find the optimal activation function and avoid overfitting [35]. With a limited scale of dataset, both SVM and RF are good choices. Considering the clinical aim of the study, namely to investigate the relationship between CLBP, CS, and gait patterns, it is important that the results of the machine learning can be translated into meaningful outcomes that can support clinical decision making. SVM can deal with non-linear data by using kernel functions; however, choosing an appropriate kernel function could be difficult for clinicians. Additionally, it implicitly maps gait outcomes to a high-dimensional features space. This mapping changes the structure of gait outcomes and makes it hard to explain which gait outcomes contribute most to the classification model. Similarly, ANN uses various of activation functions (e.g., Tanh, Sigmoid), and makes the interactions of the gait outcomes invisible. On the contrary, RF is an ensemble of decision trees. Decision trees can incorporate gait outcomes interactions naturally in the classification process. For example, a decision tree with depth 2 from a RF, with the father node IH-V and the son node Sen-AP, can describe an interactive gait pattern: if IH-V >* and Sen-AP >*, the data belong to CLBP-. Because RF includes multiple decision trees it can capture the complex interaction of gait outcomes with good accuracy. Each tree is built based on a random subset of gait outcomes and the samples in the dataset can be repeatedly selected when training. Consequently, it can help to reduce the chance of overfitting and provide a generalized model. RF can incorporate gait outcomes interactions naturally in the classification process. SHAP can use this information that stores in the tree structure to disclose which gait outcomes are different between CLBP- and CLBP + groups. These differences in terms of gait regularity, smoothness, and stability are meaningful to the clinicians.

In this study, SHAP was used to evaluate the importance of each gait outcome, instead of the conventionally used Gini impurity and information entropy. The value of Gini impurity is based on the tree structure in RF and information entropy reflects the level of "information" of a gait outcome. Gait outcomes are interrelated and interact in a complex nonlinear manner [33]. SHAP is based on the game theory and evaluates the contribution of each gait outcome to the classification accuracy by computing all possible combinations between gait outcomes. Therefore, SHAP provides a good method to explain the importance of gait outcomes to RF. The SHAP values suggest that the differences between CLBP- and CLBP + groups are reflected in smoothness, stability, predictability, regularity, and variability of gait. Compared with CLBP-group, CLBP + group exhibited lower smoothness and local stability of gait, while the CLBP + group exhibited a more regular, less variable, and more predictable gait pattern.

Gait patterns of patients with CLBP, are usually compared with the gait pattern of healthy persons. To the best of our knowledge, this is the first study in patients with CLBP that addresses the difference in gait pattern between two CLBP groups based on low and high CS levels, which makes a direct comparison with other studies intricate. The results of different gait patterns between low and high CS levels support the notion that within the heterogenous CLBP group, different motor control strategies are adopted. Two motor control strategies on a continuum have been suggested with "tight control" and "loose control" at each end, and normal trunk control in the middle [36].

The gait patterns of CLBP + group might suggest that patients with CLBP + adopt a more "tight control". The "tight control" involves increased trunk muscle activation and enhanced muscle co-contraction, might enhance control over trunk posture and movement [36]. Increased muscle activation and enhanced co-contraction would help individuals to maintain the stability of lumbar spine [37] by restricting the movement amplitude of lumbar spine. However, in a complex daily-living environment, this strategy might impair patients' ability to maintain balance during walking because of the unstable surfaces and environmental perturbations [38], and therefore has a lower gait stability (compared with CLBP- patients). Increased co-contraction would reduce the demand for the intricate control of the sequences of muscle activation. It might avoid the potential error raised by inaccurate sensory feedback of CLBP [36]. This might allow patients to control their trunks' movement precisely [39] and, therefore, result in a lower variability and a higher regularity of gait of CLBP + patients. Our results might infer thus that the CLBP + group exhibited a more "tight control". Therefore, the lower stability and variability, higher regularity and predictability in gait of the CLBP + group could be the result of the adoption of "tight control".

The gait patterns of CLBP- group, on the other hand, might be explained by "loose control" strategy. The "loose control" that involves reduced muscle excitability, might reduce the control over trunk movements [36]. The spine of which each spinal unit has 6° of freedom, is controlled by its surrounding musculature. Reduced muscular excitability, leads to a reduced control over the spinal muscle, to larger amplitude movements, and to more movement variability during repeated tasks [36]. The increased variability in gait of CLBP- group might support this finding. Additionally, increased variability would lead to a lower regularity in gait which was also found in the CLBP-group. Apart from this, increased motor variability might probably prevent muscle fatigue [40] because it allows sharing the load between different structures or tissues. Motor variability makes it possible to explore new pain-free motor control solutions [41]. This is a possible explanation for the higher smoothness in gait of CLBP- patients, because it allowed them to flexibly adapt to the complex daily-living environment by using different movement solutions. So, the higher variability and smoothness, and lower regularity in gait patterns might hint that CLBP- group adopted a more "loose control" compared with CLBP + group.

Although the "tight control" adapted strategy might have short-term

benefits, it may also contribute to a higher level of CS. The “tight control” present in CLBP + patients presumably increase muscle activation and co-contraction, and lead to larger forces acting on the spine and higher spinal loading. Moreover, it has been shown that even when patients are at rest, muscle co-contraction can be continuous [42]. These long-lasting peripheral noxious stimuli might explain the development and/or persistence of CS [43]. Additionally, it has been reported that a “tight control” strategy relates to negative pain cognitions [44], a psychological process that also might contribute to the higher CS scores of the CLBP + group.

Clinically, the gait outcomes identified as important to the classifier, may assist clinicians in providing them with a more accurate understanding of the gait performance of patients with CLBP, with low or high CS levels, and with an explicit operationalization of the observed “abnormal” gait pattern of patients with chronic pain. Whether “abnormal” should be interpreted as a functional or a dysfunctional motor control strategy in the short or long term, remains to be studied. RF and SHAP used in this study have presented a novel way to identify interacting features, and therefore, can be used for further studies. The presented accurate classification could become meaningful if this would lead to effective treatment approaches. The differences in gait patterns of CLBP- and CLBP + groups could be the results of the different motor control adapted strategies and the different motor control adapted strategies could be the causes, consequences, or both, of differences in CS levels on patients with CLBP. While this cross-sectional study has objectified a relation between CS and gait outcomes, the causality of this relation is unknown. Follow-up studies would benefit from a longitudinal design with multiple measurements to help further unraveling of this relation, as well as the relation to disability.

In line with most studies on walking and CLBP, we used cross-sectional data, thus we are not allowed to infer causality between motor control changes, CS and CLBP. Some patients had analgesic or anti-inflammatory treatment at the beginning of the study, and how these medicines interact with CS and gait outcomes is unknown. Moreover, we labeled the groups based on CSI score and the cut-off values from a previous study [25]. It should also be noted that a gold standard measure to diagnose CS is unavailable. The CSI is regarded as an indirect measure of CS, because higher scores are associated with the presence of CS syndromes [25]. In addition to gait assessment, it would be interesting to explore differences in physical activities between CLBP- and CLBP + groups, because several studies reported that relationship between CLBP and physical activity levels is heterogeneous [45].

5. Conclusion

The present study analyzed gait data during daily living of CLBP patients with low and high CS levels. RF and SHAP were applied for classification and for assessing the contribution of gait outcomes to the model. This analytic approach demonstrated that RF has the ability to

accurately classify subgroups of patients with CLBP and low or high CS levels based on differences in gait outcomes. The results of SHAP showed that the differences of gait outcomes between low and high CS levels were in gait regularity, variability, predictability, smoothness, and stability. This may imply that patients with low and high CS levels adopted different motor control strategies. Patients with CLBP and low CS level (CLBP-) use a “loose control” and, therefore, exhibited more smoothness and stability in gait patterns. Patients with CLBP and high CS level (CLBP+) adopted a “tight control” and showed a more regular, less variable and more predictable gait pattern.

The results of this study may contribute to a better understanding of gait characteristics in patients with CLBP, its association with CS, and may in the future assist in better-personalized rehabilitation interventions [46].

Ethics approval and consent to participate

The study was approved by the Medical Research Ethics Committee of the University Medical Center Groningen (METc 2016/702) and conducted according to the principles expressed in the Declaration of Helsinki.

Authors' contributions

XZ, MR, EO, and CL developed the conception and design of the study. JAE and RSP acquired the data from participants. XZ analyzed the data and wrote the paper under supervision of MR, OB, and CL. HK reviewed the code. All authors reviewed and commented on the manuscript. All authors approved the final manuscript.

Research data

The raw data of all 42 participants and 168 acceleration segments are not publicly available. RF was available from the package scikit-learn 1.0.1 (<https://scikit-learn.org/stable/>) and the source code of SHAP is available from shap package (<https://github.com/slundberg/shap>) based on Python 3.8.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

XZ was supported by China Scholarship Council-University of Groningen Scholarship [Grant No.201906410084].

Appendix A

We did empirical work of comparing often applied approaches for classification: Random Forest (RF), Artificial Neural Network (ANN), Support Vector Machine (SVM), Naive Bayes (NB), and K-Nearest Neighbors (KNN). As can be seen in the Table below, RF and ANN had the best performance compared to SV, NB, and KNN. RF performed better in precision and specificity, while ANN performed better in sensitivity.

Table
Classification performance comparison

	RF	ANN	SVM	NB	KNN (n = 3)
Accuracy	84.4%	81.2%	68.8%	62.5%	62.5%
Sensitivity	75%	93%	56%	50%	50%
Specificity	93%	68.8%	81.3%	75%	75%
Precision	92%	75%	75%	66.7%	66.7%
F1-score	82.6%	83%	64.1%	57%	57%
AUC	0.83	0.85	0.67	0.62	0.62

References

- [1] M.S. Thiese, K.T. Hegmann, E.M. Wood, A. Garg, J.S. Moore, J. Kapellusch, et al., Prevalence of low back pain by anatomic location and intensity in an occupational population, *BMC Musculoskel. Disord.* 15 (2014) 283.
- [2] S. Dagenais, J. Caro, S. Haldeman, A systematic review of low back pain cost of illness studies in the United States and internationally, *Spine J.* 8 (2008) 8–20.
- [3] J. Hartvigsen, M.J. Hancock, A. Kongsted, Q. Louw, M.L. Ferreira, S. Genevay, et al., What low back pain is and why we need to pay attention, *Lancet* 391 (2018) 2356–2367.
- [4] O. Airaksinen, J.I. Brox, C. Cedraschi, J. Hildebrandt, J. Kläber-Moffett, F. Kovacs, et al., European guidelines for the management of chronic nonspecific low back pain, *Eur. Spine J.* 15S2 (2006) s192.
- [5] J. Nijs, S.Z. George, D.J. Clauw, C. Fernández-de-las-Peñas, E. Kosek, K. Ickmans, M. Curatolo, et al., Central sensitisation in chronic pain conditions: latest discoveries and their potential for precision medicine, *The Lancet Rheumatology* 3 (5) (2021) e383–e392.
- [6] M.W. van Tulder, B.W. Koes, L.M. Bouter, A cost-of-illness study of back pain in The Netherlands, *Pain* 62 (1995) 233–240.
- [7] R.H. Jenkinson, Central Sensitization, 2019, pp. 45–48. Pain.
- [8] N.E. Foster, J.R. Anema, D. Cherkin, R. Chou, S.P. Cohen, D.P. Gross, et al., Prevention and treatment of low back pain: evidence, challenges, and promising directions, *Lancet* 391 (2018) 2368–2383.
- [9] C. Maher, M. Underwood, R. Buchbinder, Non-specific low back pain, *Lancet* 389 (2017) 736–747.
- [10] M. Götsze, M. Ernst, M. Koch, R. Blickhan, Influence of chronic back pain on kinematic reactions to unpredictable arm pulls, *Clin. BioMech.* 30 (2015) 290–295.
- [11] T. Giesecke, R.H. Gracely, M.A.B. Grant, A. Nachemson, F. Petzke, D.A. Williams, et al., Evidence of augmented central pain processing in idiopathic chronic low back pain, *Arthritis Rheum.: Off. J. Am. Coll. Rheumatol.* 50 (2004) 613–623.
- [12] D.K. Ahern, M.J. Follick, J.R. Council, N. Laser-Wolston, H. Litchman, Comparison of lumbar paravertebral EMG patterns in chronic low back pain patients and non-patient controls, *Pain* 34 (1988) 153–160.
- [13] C.J.C. Lamoth, O.G. Meijer, A. Daffertshofer, P.I. Wuisman, P.J. Beek, Effects of chronic low back pain on trunk coordination and back muscle activity during walking: changes in motor control, *Eur. Spine J.* 15 (2006) 23–40.
- [14] G. Christe, F. Kade, B.M. Jolles, J. Favre, Chronic low back pain patients walk with locally altered spinal kinematics, *J. Biomech.* 60 (2017) 211–218.
- [15] C.J.C. Lamoth, J.F. Stins, M. Pont, F. Kerckhoff, P.J. Beek, Effects of attention on the control of locomotion in individuals with chronic low back pain, *J. NeuroEng. Rehabil.* 5 (2008) 1–8.
- [16] S.P. Gombatto, T. Brock, A. DeLork, G. Jones, E. Madden, C. Rinere, Lumbar spine kinematics during walking in people with and people without low back pain, *Gait Posture* 42 (2015) 539–544.
- [17] D. Hamacher, D. Hamacher, F. Herold, L. Schega, Are there differences in the dual-task walking variability of minimum toe clearance in chronic low back pain patients and healthy controls? *Gait Posture* 49 (2016) 97–101.
- [18] R. Müller, T. Ertelt, R. Blickhan, Low back pain affects trunk as well as lower limb movements during walking and running, *J. Biomech.* 48 (2015) 1009–1014.
- [19] J.A. Echeita, H.R.S. Preuper, R. Dekker, I. Stuive, H. Timmerman, A.P. Wolff, et al., Central Sensitisation and functioning in patients with chronic low back pain: protocol for a cross-sectional and cohort study, *BMJ Open* 10 (2020), e031592.
- [20] Y. Zhou, R. Romijnders, C. Hansen, J. van Campen, W. Maetzler, T. Hortobágyi, et al., The detection of age groups by dynamic gait outcomes using machine learning approaches, *Sci. Rep.* 10 (2020) 1–12.
- [21] K.S. van Schooten, M. Pijnappels, S.M. Rispens, P.J. Elders, P. Lips, J.H. van Dieën, Ambulatory fall-risk assessment: amount and quality of daily-life gait predict falls in older adults, *J. Gerontol. A Biol. Med. Sci.* 70 (2015) 608–615.
- [22] S. Del Din, A. Godfrey, B. Galna, S. Lord, L. Rochester, Free-living gait characteristics in ageing and Parkinson's disease: impact of environment and ambulatory bout length, *J. NeuroEng. Rehabil.* 13 (1) (2016) 1–12.
- [23] J. Vickers, A. Reed, R. Decker, B.P. Conrad, M. Olegario-Nebel, H.K. Vincent, Effect of investigator observation on gait parameters in individuals with and without chronic low back pain, *Gait Posture* 53 (2017) 35–40.
- [24] C.P. van Wilgen, M. Meeus, F. Descheemaeker, B. Cagnie, Central sensitization inventory – nederlandse consensusveraling, Available: <http://www.paininmotion.be/storage/app/media//materials/CSI-consensusveraling.pdf>, 2013. (Accessed 10 December 2020).
- [25] R. Neblett, H. Cohen, Y.H. Choi, M.M. Hartzell, M. Williams, T.G. Mayer, R. J. Gatchel, The Central Sensitization Inventory (CSI): establishing clinically significant values for identifying central sensitivity syndromes in an outpatient chronic pain sample, *J. Pain* 14 (2013) 438–445.
- [26] N. Ichinoseki-Sekine, Y. Kuwae, Y. Higashi, T. Fujimoto, M. Sekine, T. Tamura, Improving the accuracy of pedometer used by the elderly with the FFT algorithm, *Med. Sci. Sports Exerc.* 38 (2006) 1674–1681.
- [27] J. Qian, L. Pei, J. Ma, R. Ying, P. Liu, Vector graph assisted pedestrian dead reckoning using an unconstrained smartphone, *Sensors* 15 (2015) 5032–5057.
- [28] K.S. Van Schooten, M. Pijnappels, S.M. Rispens, P.J. Elders, P. Lips, A. Daffertshofer, et al., Daily-life gait quality as predictor of falls in older people: a 1-year prospective cohort study, *PLoS One* 11 (2016), e0158623.
- [29] L.H.J. Kikkert, N. Vuillerme, J.P. van Campen, B.A. Appels, T. Hortobágyi, C.J. C. Lamoth, Gait characteristics and their discriminative power in geriatric patients with and without cognitive impairment, *J. NeuroEng. Rehabil.* 14 (2017) 1–10.
- [30] G. Blau, E. Scornet, A random forest guided tour, *Test* 25 (2016) 197–227.
- [31] S.M. Lundberg, S.I. Lee, A unified approach to interpreting model predictions, *Adv. Neural Inf. Process. Syst.* (2017) 4765–4774.
- [32] C.J.C. Lamoth, E. Ainsworth, W. Polomski, H. Houdijk, Variability and stability analysis of walking of transfemoral amputees, *Med. Eng. Phys.* 32 (2010) 1009–1014.
- [33] L. Quach, A.M. Galica, R.N. Jones, E. Procter-Gray, B. Manor, M.T. Hannan, L. A. Lipsitz, The nonlinear relationship between gait speed and falls: the maintenance of balance, independent living, intellect, and zest in the elderly of Boston study, *J. Am. Geriatr. Soc.* 59 (6) (2011) 1069–1073.
- [34] B. Pogorelc, Z. Bosnić, M. Gams, Automatic recognition of gait-related health problems in the elderly using machine learning, *Multimed. Tool. Appl.* 58 (2) (2012) 333–354.
- [35] E. Halilaj, A. Rajagopal, M. Fiterau, J.L. Hicks, T.J. Hastie, S.L. Delp, Machine learning in human movement biomechanics: best practices, common pitfalls, and new opportunities, *J. Biomed.* 81 (2018) 1–11.
- [36] J.H. Van Dieën, N.P. Reeves, G. Kawchuk, L.R. Van Dillen, P.W. Hodges, Motor control changes in low back pain: divergence in presentations and mechanisms, *J. Orthop. Sports Phys. Ther.* 49 (2019) 370–379.
- [37] J. Cholewicki, A.P. Simons, A. Radebold, Effects of external trunk loads on lumbar spine stability, *J. Biomed.* 33 (2000) 1377–1385.
- [38] N.W. Mok, P.W. Hodges, Movement of the lumbar spine is critical for maintenance of postural recovery following support surface perturbation, *Exp. Brain Res.* 231 (2013) 305–313.
- [39] W. Van den Hoorn, S.M. Bruijn, O.G. Meijer, P.W. Hodges, J.H. Van Dieën, Mechanical coupling between transverse plane pelvis and thorax rotations during gait is higher in people with low back pain, *J. Biomed.* 45 (2012) 342–347.
- [40] P. Madeleine, On functional motor adaptations: from the quantification of motor strategies to the prevention of musculoskeletal disorders in the neck–shoulder region, *Acta Physiol.* 199 (2010) 1–46.
- [41] M.L. Meier, A. Vrana, P. Schweinhardt, Low back pain: the potential contribution of supraspinal motor control and proprioception, *Neuroscientist* 25 (2019) 583–596.
- [42] I.A. Schinkel, B.C. Nairn, J.D. Drake, Investigation of trunk muscle co-contraction and its association with low back pain development during prolonged sitting, *J. Electromyogr. Kinesiol.* 23 (2013) 778–786.
- [43] J. Nijs, S.Z. George, D.J. Clauw, et al., Central sensitisation in chronic pain conditions: latest discoveries and their potential for precision medicine, *The Lancet Rheumatology* 3 (5) (2021) e383–e392.
- [44] G.B. Ross, P.J. Sheahan, B. Mahoney, Pain catastrophizing moderates changes in spinal control in response to noxiously induced low back pain, *J. Biomed.* 58 (2017) 64–70.
- [45] R. Parker, E. Bergman, A. Mntambo, S. Stubbs, M. Wills, Levels of PA in people with chronic pain, *S. Afr. J. Physiother.* 73 (1) (2017) 1–7.
- [46] J.H. Van Dieën, N.P. Reeves, G. Kawchuk, L.R. Van Dillen, Analysis of motor control in patients with low back pain: a key to personalized care? *J. Orthop. Sports Phys. Ther.* 49 (2019) 380–388.