Effect of level of Sensitization on gait in Chronic low back pain: insights from a machine learning approach

Xiaoping Zheng1*, Michiel Reneman2, Jone Ansuatgeui Echeita2, HR Schiphorst Preuper2, Herbert Kruitbosch3, Egbert Otten1, Claudine Lamoth1

* Correspondence: x.zheng@umcg.nl

1. University of Groningen, University Medical Center Groningen, Department of Human Movement Sciences, Groningen, Netherlands

2. University of Groningen, University Medical Center Groningen, Center for Rehabilitation, Groningen, the Netherlands

3. University of Groningen, Center for Information Technology, Data science team, Groningen, the Netherlands
Abstract

Background: Despite central sensitization (CS) often presents in patients with chronic low back pain (CLBP), there is a lack of quantitative and qualitative analysis of the effect of CS on gait performance. The study aimed to investigate the daily-living gait performance of patients with CLBP with low and high CS levels (CLBP- and CLBP+, respectively).

Method: Forty-two patients with CLBP were included. CS was assessed by Central Sensitization Inventory (CSI). Patients were classified according to low or high CSI score (23 CLBP- and 19 CLBP+). Patients wore a 3D accelerometer for about one week. From each patient, 4 days of accelerometer-data were randomly selected. For each day data, continuous gait cycles were extracted by using a Fast Fourier Transform-based and a zero-cross method. For all gait cycles in one day, 36 gait outcomes representing variables related to pace, regularity, smoothness, local stability and predictability of gait were calculated. A Random Forest classifier was trained to classify CLBP- and CLBP+ groups based on gait outcomes and SHapley Additive exPlanations (SHAP) method was used to explain the differences between groups in gait outcomes.

Results: The Random Forest classifier could accurately recognize the CLBP- and CLBP+ groups (accuracy = 84.4%, F1-score = 82.6%). SHAP reported that the most differences between CLBP- and CLBP+ groups were: index of harmonicity-vertical and harmonic ratio-mediolateral (gait smoothness), stride frequency variability-mediolateral/anteroposterior, stride length variability (gait variability), stride regularity-mediolateral (gait regularity), maximal Lyapunov exponent-
vertical/mediolateral and maximal Lyapunov exponent per stride-vertical (gait stability), and sample entropy-anteroposterior (gait predictability).

**Conclusions:** CLBP- and CLBP+ presented different motor control strategies. CLBP- presented a more “loose” control, including higher gait smoothness and stability. CLBP+ presented a more “tight” control, including a more regular, less variable and more predictable gait pattern.

**Keywords:** Low back pain, Central sensitization, Supervised machine learning, Gait, Daily life.
**Background**

Chronic low back pain (CLBP) is one of the most prevalent chronic musculoskeletal pains in the world [1]. It is responsible for high treatment costs, sick leave and individual suffering and it represents a significant socioeconomic burden [2]. For 85% to 90% of patients with CLBP, the relation between pathoanatomical and clinical presentations is weak [3] and, therefore, it is classified as nonspecific CLBP [4]. In CLBP, and other chronic musculoskeletal disorders, central sensitization (CS) might be present [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, 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 in order to avoid painful movement, such as walking [11]. Walking is one of the abilities most affected by CLBP. Many clinicians may intuitively identify “abnormal” gait patterns in patients with CLBP, but identification and objectifying of specific “abnormal” gait features 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 not ubiquitous. Studies between patients with CLBP and healthy controls, observed inconsistent evidence regarding preferred walking velocity [13] [16], stride length [14] [17], and stride-to-stride variability [15] [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 features such as walking speed and stride length, might not be sensitive enough to detect small differences between patients with high and low 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]. Despite the fact that 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 methods 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 the controlled laboratory environment, while being observed, may change the performance of patients [23]. The successful employment of machine learning methods in gait analysis makes it possible to extract the most informative gait outcomes from the accelerometer sensor data [20]. If patients with lower and higher levels of CS walk differently, the machine learning methods will be able to successfully recognize these differences by their gait outcomes. Many gait outcomes are not independent and interact with each other, such as gait speed and step regularity. Machine learning methods such as Random Forest approach, 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 method. It was hypothesized that patients with CLBP and higher levels of CS show differences in daily life gait performance, compared with those with lower levels.
Methods

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.

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].

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 0-100 where a higher scoring represents a higher level of CS. A score lower 40 indicates low CS level (CLBP- group) and a score of 40-100 is interpreted as high CS level (CLBP+ group) [25].

Mann-Whitney U test was used to statistically test the differences between CLBP- and CLBP+ groups for demographics and CSI scores.

Accelerometer data. The accelerometer data was 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). The sampling frequency of the accelerometer was set to 100 Hz and the dynamic range was ±6 gravity.

Data processing and analysis

Raw data segmentation

Each patient’s accelerometer data was segmented into 24 hours span data segments (from 12:00 P.M. to next day 11:59 A.M) to represent the activities during the days. Data which was not completely covering this 24-hour span was discarded from the analysis. Due to technical or personal reasons of the patient, not all patients were collected a full week of data. In order to compare the data between different patients fairly, 4 segments (representing 4 days) of each patient were included. Therefore, 7 patients who had less than 4 segments, were excluded and from patients with more than 4 segments, 4 segments were randomly sampled. Figure 1a graphically shows the process of the raw data segmentation.
Figure 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 bouts extraction**

The accelerometer data of the 4 segments was first smoothed by a low-pass filter with a 2nd order Butterworth, filtered with a 20 Hz cut-off frequency, such that only frequencies lower than 20 Hz remained. 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 seconds, 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. Figure 1b presents the walking bouts as the yellow vertical bars in the rectangle.

**Gait outcomes**

All walking bouts in one 24-hour 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 walking bout. And then, all walking bouts exceeding 10 seconds were selected and cut into non-overlapping 10-second windows [28]. From the segment, each 10-second window was used to calculate different gait outcomes, and these values were averaged over all 10-second windows in the segment representing the patient’s gait performance on that day.

The 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 Figure 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 reference [29].
## Table 1. Quantitative Gait Outcomes.

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

## Table 2. Qualitative Gait Outcomes.

| Catalog | Gait characteristic | Description and method |
|---------|---------------------|------------------------|
|         | Stride regularity (SR; V, ML, AP, All) | SR is computed by using the unbiased autocorrelation coefficient: $Ad(m) = \frac{1}{N-\left|m\right|} \sum_{i=1}^{N-\left|m\right|} Acc(i) \cdot Acc(i + m)$, where $Acc(i)$ is the sample acceleration signal, $N$ the number of samples, and $m$ the number of time lag. The first peak of $Ad(m)$ is $Ad_1$ and it represents the stride regularity. Higher values (maximum 1.0) reflect repeatable patterns between strides. 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. |
| Regularity | Gait symmetry index (GSI) | $IH = \frac{P_h}{\sum_{i=1}^{N} P_i}$. 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. |
| Smoothness | Index of harmonicity (IH; V, ML, AP, All) | $HR = \frac{\sum_{i=1}^{N} P_a}{\sum_{i=1}^{N} P_h}$ In VT and AP directions, $\sum P_a$ = the sum of even power spectral and $\sum P_h$ = 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) |
|----------------|----------------------------------|
| Stability      | Maximal Lyapunov exponent (max LyE; V, ML, AP) | Maximal Lyapunov exponent normalized per stride by time (max LyE per stride; V, ML, AP) |

**Random Forest Classifier**

To separate CLBP- and CLBP+ groups by gait outcomes, a Random Forest classifier was used. The Random Forest classifier is considered as the optimal machine learning classification method for the present data, because it performs well with (a) nonlinear and linear data; (b) high dimensional data, obsoleting dimensionality reduction; and (c) unbalanced and small data sets [30].

The input data of this method was \(< S, L >\). \( S \) represents the gait outcome vectors of patients and \( L \) was its corresponding label. The definition of \( S \) is: \( S = \{ s_1, s_2, ..., s_l, ..., s_m \} \) and \( s_i = [d_1, ..., d_k] \), 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 \( k = 36. L = l_1, ..., l_m \), where \( l \in \{ CLBP-, CLBP+ \} \).

The Random Forest is constructed in four steps. Step one: Randomly sample \( n \) gait 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 two: In $S_b$, randomly sample $j$ ($j \leq k$) gait outcomes from $s$. Therefore, $s'_i = d'_1, \ldots, d'_j$ and $S'_b = s'_1, \ldots, s'_n$. 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 random forest, that predicts by voting (see Figure 2).

Before training the Random Forest classifier, 80% of patients were randomly selected and their 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. In order to avoid overfitting of the hyperparameters, a 5-fold cross-validation method was used to estimate them. As shown in Figure 1d, the training data was randomly split into 5 folds. Four folds were used to train the model and the rest fold was used to estimate the performance of the current hyperparameters in the Random Forest classifier. 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 whole

![Random Forest Diagram](image-url)
training data and the hyperparameters were used to fit the final Random Forest classifier and the testing data set was used to evaluate the generalizability of the model.

Accuracy evaluation

Accuracy, sensitivity, specificity, precision, and F1-score were calculated to evaluate the performance of the classification (see Figure 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.

\[
\text{accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}
\]

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

\[
\text{sensitivity} = \frac{TP}{TP + FN} \tag{2}
\]

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

\[
\text{specificity} = \frac{TN}{TN + FP} \tag{3}
\]

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

\[
\text{precision} = \frac{TP}{TP + FP} \tag{4}
\]

F1-score is the harmonic mean (average) of the precision and sensitivity.
The receiver operating characteristic (ROC) curve was calculated to evaluate the performance of the Random Forest classifier. The Y-axis of this curve represents the true positive rate (sensitivity) and the X-axis means false positive rate (1-specificity). Therefore, the overall classification performance of the Random Forest classifier 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 means perfect performance. Since the random forest can also output the percentage of trees that vote CLBP+, AUC represents the probability that the random forest outputs a lower percentage for a negative sample than for a positive sample.

**Feature importance**

SHapley Additive exPlanations (SHAP) [31] was used to explain the gait features importance to the classification model. SHAP connects optimal credit allocation with local explanations using the classic Shapley values from game theory [32]. Shapley values, $\phi_i$, explains the importance of gait outcome $i$ for Random Forest classifier and is defined as:

$$\phi_i = \frac{1}{|N|!} \sum_{(i) \subseteq s \subseteq N} ([|s| - 1]! ([N] - |s|)! [R(s) - R(s - \{i\})]$$

(6)

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 Random Forest classifier accuracy of the input gait outcomes. Since

$$F1 = \frac{2 \times \text{precision} \times \text{sensitivity}}{\text{precision} + \text{sensitivity}}$$

(5)
computing the exact Shapley values is computationally expensive, SHAP method 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.

**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 excluded since 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 for 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 < 0.0001) and BSI (p = 0.01).

Testing data was used to evaluate the generalizability of the Random Forest classifier and the confusion matrix is shown in Figure 3. From the confusion matrix, accuracy, sensitivity, specificity, precision, and the F1-score were calculated to evaluate the performance metric of the model. The Random Forest classifier achieved an accurate classification-result (84.4% accuracy) and the sensitivity and specificity were 75.0% and 93% respectively. The precision is 92% and the F1-score is 82.6%. The ROC curve is presented in Figure 4 showing that the Random Forest classifier achieved a 0.83 AUC.
Table 3 Patient characteristics (n=42).

|                              | CLBP- (n=23) | CLBP+ (n=19) | All (n=42) | P-Value |
|------------------------------|--------------|--------------|------------|---------|
| Gender                       | 15W / 8M     | 12W / 7M     | 27W / 15M  |         |
| 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/m²       | 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.
Figure 3. Classification results for random forest, and the mean accuracy is 84.4%.

CLBP-; CLBP+: Patients with chronic low back pain with low (-) and high (+) central sensitization levels.

Figure 4. The receiver operating characteristic (ROC) curve (in red) for Random Forest classifier. AUC: area under the curve.
The importance of the gait outcomes for the Random Forest classifier is shown in Figure 5. In this study, only the top 10 gait outcomes (above the red line in Figure 5) were considered as important since the mean absolute SHAP values of other gait outcomes are really low. They 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.

### Feature Importance

![Feature Importance](image)

- **IH-V**: Index of harmonicity in vertical direction.
- **SF Variability-ML**: Variability of stride frequency in mediolateral/anteroposterior direction.
- **SR-ML**: Stride regularity in mediolateral direction.
- **Max LyE-V**: Maximal Lyapunov exponent in vertical/mediolateral direction.
- **SF Variability-AP**: Variability of stride frequency in anteroposterior direction.
- **Sen-AP**: Sample entropy in anteroposterior direction.
- **Max LyE-ML**: Maximal Lyapunov exponent in mediolateral direction.
- **Max LyE per stride-V**: Maximal Lyapunov exponent per stride in vertical direction.
- **HR-ML**: Harmonic ratio in mediolateral direction.
- **SL Variability**: Variability of stride length.
- **WS Variability**: Variability of walking speed.
- **IH-ML**: Index of harmonicity in mediolateral direction.
- **WS**: Walking speed.
- **SL**: Stride length.

Figure 5. Features importance of Random Forest classifier. Top 10 gait outcomes above the red are: index of harmonicity in vertical direction (IH-V), variability of stride frequency in mediolateral/anteroposterior 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 rest 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.

Figure 6 shows the violin-box plot of the Top 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 each gait outcomes. It shows that CLBP- group has higher IH-V, HR-ML (better smoothness); higher SF-variance-ML, SF-variance-AP, SL-variance (lesser 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- has a wider distribution and CLBP+ shows more data at the bottom. In the Max LyE-ML, data of CLBP- is concentrated on 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 bimodality distribution but the peaks of distribution are totally different. In SF Variability-ML and SF Variability-AP, CLBP+ has a larger peak at the bottom while CLBP- has a 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.

Figure 6. Violin-box plot for the top 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/ anteroposterior 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 ration in mediolateral direction.
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 method. Based on quantitative and qualitative gait outcomes, using a Random Forest method, the two groups (CLBP- and CLBP+) could be classified with an accuracy of 84.4%. The classification results indicated that CLBP- patients walk differently from CLBP+. 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 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 [33], e.g. walking when carrying a cup of coffee. These influences on gait are not present in controlled studies, and are not captured by averaged based conventional gait outcomes that average outcomes over stride cycles, such as mean, step length, stride or step time, and number of steps. Therefore, in the present study we included gait outcomes that take into account how gait cycles evolve over time, e.g. the interdependency of gait cycles, 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 that take into account how gait cycles evolve over time [34].

The Random Forest method differs from conventional statistical methods, such as the T-test which does not consider the interaction of gait outcomes and simply evaluates the differences of each gait outcome one by one. The Random Forest method is an ensemble of decision trees and incorporating gait outcomes interactions naturally in the classification process. For example, a decision tree with depth 2 from the Random Forest, 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-. Since the Random Forest includes multiple decision trees and each tree is built based on a random subset of gait outcomes, Random Forest can capture the complex interaction of gait outcomes. Additionally, it can help to reduce the chance of overfitting to training data. Therefore, the Random Forest improves predictive accuracy and it can provide a generalized model of the difference between CLBP- and CLBP+ groups.

The SHAP tree explainer can provide good explanations for the Random Forest [31]. It fastens the calculation of the SHAP values by exploiting the information stored in the tree structure which already captured the gait outcomes interactions. The top 10 SHAP values suggest that the differences between CLBP- and CLBP+ groups are gait outcomes which represent smoothness, stability, predictability, regularity and variability in gait. Compared with CLBP- group, CLBP+ group exhibited worse
smoothness and local stability in gait. Moreover, the CLBP+ group exhibited a more regular, less variable and more predictable gait pattern.

Most studies on walking in patients with CLBP are compared with control participants with no back pain. 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 level, which makes a direct comparison with other studies intricate. However, the results of different gait patterns between low and high CS level 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 [35].

There is considerable evidence on the observed changes in muscle activation of patients with CLBP [36]. Tight control which involves increased trunk muscle activation and enhanced muscle co-contraction, might enhance control over trunk posture and movement [35]. Increased muscle activation and enhanced co-contraction would help individuals to maintain the stability of lumbar spine [37], which is an unstable structure. However, this strategy might impair patients’ ability to maintain balance in a complex daily-living environment where unstable surfaces and environmental perturbations occur [38]. 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 [35]. This might allow patients to control their trunks’ movement precisely and result in lower variability of movement patterns [39]. Our results might infer that CLBP+ group exhibited a more “tight” control. Tight control allows them to control their movement more strictly with results in a decreased motor variability, higher regularity and predictability of the gait pattern compared with CLBP- group.

The CLBP- group on the other hand might use another ‘loose control’ strategy. The “Loose control”, which involves reduced muscle excitability, might reduce the control over trunk movements [35]. The spine of which each spinal unit has 6 degrees 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 [35]. Increased motor variability might probably prevent muscle fatigue [40] since it allows sharing the load between different structures or tissues. Moreover, motor variability makes it possible to explore new pain-free motor control solutions [41]. The results of the present study hint at a more ‘loose control’ in the CLBP- group, which increased motor variability, which might allow them to flexibly adapt to the complex daily-living environment using different movement solutions. Additionally, irrespective of the larger variability local stability and smoothness of the gait pattern was higher in this group than in the CLBP+ group.
Although both motor control adapted strategies might have beneficial effects for the short term, they might lead to negative long-term consequences. Increased muscle activation and co-contraction in tight control would increase the forces acting on the spine and it would lead to higher spinal loading. Moreover, even when patients are at rest, the co-contraction of muscles is continuous [42]. These might result in accumulation of waste products, muscle fatigue, and intervertebral disc degeneration [40][43]. Regarding loose control, the reduced control of lumbar spine may eventually increase the tissue strains, with subsequent increases in spinal loading, and pain [41]. Thus, eventually both strategies affect mechanical loading on lumbar tissues. The loading might be the source of nociceptive input and might contribute to the CS since the load may sufficient to excite sensitized afferents [35].

Clinically, the important gait outcomes of this study may assist clinicians in providing a more accurate understanding of the gait performance of patients with CLBP, with low or high CS levels, and a more 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. The approaches used in this study have presented a novel way to identify interacting feature, and therefore, can be used for further studies. Clinically, the present accurate subclassification could become meaningful if this would lead to effective treatment approaches. While this cross-sectional study has objectified a relation between CS and gait features, the direction 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. 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+, because several studies reported that relationship between CLBP and physical activity levels is heterogeneous [44].

**Conclusion**

The present study analyzed gait data during daily living of CLBP patients with low and high CS levels. A Random Forest method and the SHAP method were applied for classification and identification the contribution of gait outcomes to the model. This analytic approach demonstrated that Random Forest method 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 method showed the differences between low and high CS levels were in gait regularity, variability, predictability, smoothness and stability. The differences in gait outcomes may infer that patients with
low and high CS levels adopted different motor control strategies. Patients with CLBP and low CS level (CLBP-) may use a more loose control and, therefore, exhibited more smoothness and stability in gait patterns. Patients with CLBP and high CS level (CLBP+) may adopt a more 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 [45].

**List of abbreviations**

CLBP: chronic low back pain  
CS: central sensitization  
CSI: Central Sensitization Inventory  
CLBP-: chronic low back pain with low central sensitization  
CLBP+: chronic low back pain with high central sensitization  
V: vertical direction  
ML: mediolateral direction  
AP: anteroposterior direction  
WS: walking speed  
SL: Stride length  
ST: Stride time
SF: Stride frequency
RMS: Root mean square of the variability of the amplitude of accelerations
SR: Stride regularity
GSI: Gait symmetry index
IH: Index of harmonicity
HR: Harmonic ratio
Sen: Sample entropy
max LyE: Maximal Lyapunov exponent
max LyE per stride: Maximal Lyapunov exponent normalized per stride by time
TP: true positives
TN: true negative
FP: false positives
FN: false negative
ROC: receiver operating characteristic
AUC: area under the receiver operating characteristic curve
SHAP: SHapley Additive exPlanations

Declarations

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.
Consent for publication
Not applicable.

Availability of data and materials
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Funding
XZ was supported by China Scholarship Council-University of Groningen Scholarship under Grant No.201906410084.

Authors’ contributions
XZ, MR, EO and CL developed the idea. JAE and HRSP collected 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.

Acknowledgements
Not applicable.
Reference

[1] Thiese, MS, Hegmann KT, Wood EM, Garg A, Moore JS, Kapellusch J, et al. Prevalence of low back pain by anatomic location and intensity in an occupational population. BMC musculoskeletal disorders, 2014;15:283.

[2] Dagenais S, Caro J, Haldeman S. A systematic review of low back pain cost of illness studies in the United States and internationally. The spine journal. 2008;8:8-20.

[3] Hartvigsen J, Hancock MJ, Kongsted A, Louw Q, Ferreira ML, Genevay S, et al. What low back pain is and why we need to pay attention. The Lancet. 2018;391:2356-2367.

[4] Airaksinen O, Brox JI, Cedraschi C, Hildebrandt J, Klaber-Moffett J, Kovacs F, et al. European guidelines for the management of chronic nonspecific low back pain. European spine journal. 2006;15 Suppl 2: s192.

[5] Nijs J, Van Wilgen CP, Van Oosterwijck J, van Ittersum M, Meeus M. How to explain central sensitization to patients with ‘unexplained’chronic musculoskeletal pain: practice guidelines. Manual therapy. 2011;16:413-418.

[6] van Tulder MW, Koes BW, Bouter LM. A cost-of-illness study of back pain in The Netherlands. Pain. 1995;62:233-240.

[7] Jenkinson RH. Central Sensitization. Top Pain. 2019. doi:10.1007/978-3-319-99124-5_11.

[8] Foster N E, Anema J R, Cherkin D, Chou R, Cohen SP, Gross DP, et al. Prevention and treatment of low back pain: evidence, challenges, and promising directions. The Lancet. 2018;391:2368-2383.

[9] Maher C, Underwood M, Buchbinder R. Non-specific low back pain. The Lancet.
Götze M, Ernst M, Koch M, Blickhan R. Influence of chronic back pain on kinematic reactions to unpredictable arm pulls. Clinical Biomechanics. 2015;30:290-295.

Giesecke T, Gracely RH, Grant MAB, Nachemson A, Petzke F, Williams DA, et al. Evidence of augmented central pain processing in idiopathic chronic low back pain. Arthritis & Rheumatism: Official Journal of the American College of Rheumatology. 2004;50:613-623.

Ahern DK, Follick MJ, Council JR, Laser-Wolston N, Litchman H. Comparison of lumbar paravertebral EMG patterns in chronic low back pain patients and non-patient controls. Pain. 1988;34:153-160.

Lamoth CJC, Meijer OG, Daffertshofer A, Wuisman PI, Beek PJ. Effects of chronic low back pain on trunk coordination and back muscle activity during walking: changes in motor control. European Spine Journal. 2006;15:23-40.

Lamoth CJC, Stins JF, Pont M, Kerckhoff F, Beek PJ. Effects of attention on the control of locomotion in individuals with chronic low back pain. Journal of Neuroengineering and Rehabilitation. 2008;5:1-8.

Hamacher D, Hamacher D, Herold F, Schega L. Are there differences in the dual-task walking variability of minimum toe clearance in chronic low back pain patients and healthy controls? Gait & Posture. 2016; 49:97-101.

Christe G, Kade F, Jolles BM, Favre J. Chronic low back pain patients walk with locally altered spinal kinematics. Journal of Biomechanics. 2017;60:211-218.

Gombatto SP, Brock T, DeLork A, Jones G, Madden E, Rinere C. Lumbar spine kinematics during walking in people with and people without low back pain. Gait & Posture. 2015;42:539-
[18] Müller R, Ertelt T, Blickhan R. Low back pain affects trunk as well as lower limb movements during walking and running. Journal of Biomechanics. 2015;48:1009-1014.

[19] Echeita JA, Preuper HRS, Dekker R, Stuive I, Timmerman H, Wolff AP, et al. Central Sensitisation and functioning in patients with chronic low back pain: protocol for a cross-sectional and cohort study. BMJ Open. 2020;10:e031592.

[20] Zhou Y, Romijnders R, Hansen C, van Campen J, Maetzler W, Hortobágyi T, et al. The detection of age groups by dynamic gait outcomes using machine learning approaches. Scientific Reports. 2020;10:1-12.

[21] van Schooten KS, Pijnappels M, Rispens SM, Elders PJ, Lips P, van Dieen JH. Ambulatory fall-risk assessment: amount and quality of daily-life gait predict falls in older adults. Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences. 2015;70:608-615.

[22] Del Din S, Godfrey A, Galna B, Lord S, Rochester L. Daily-living gait characteristics in ageing and Parkinson’s disease: impact of environment and ambulatory bout length. Journal of Neuroengineering and Rehabilitation. 2016;13:46.

[23] Vickers J, Reed A, Decker R, Conrad BP, Olegario-Nebel M, Vincent HK. Effect of investigator observation on gait parameters in individuals with and without chronic low back pain. Gait & Posture. 2017;53:35-40.

[24] van Wilgen CP, Meeus M, Descheemaeker F, Cagnie B. Central Sensitization Inventory – Nederlandse consensusvertaling. 2013. Available: http://www.paininmotion.be/storage/app/media//materials/CSI-consensusvertaling.pdf. Accessed 10 Dec 2020.
Neblett R, Cohen H, Choi Y H, Hartzell MM, Williams M, Mayer TG, & Gatchel RJ. The Central Sensitization Inventory (CSI): establishing clinically significant values for identifying central sensitivity syndromes in an outpatient chronic pain sample. The Journal of Pain. 2013;14:438-445.

Ichinoseki-Sekine N, Kuwae Y, Higashi Y, Fujimoto T, Sekine M, Tamura T. Improving the accuracy of pedometer used by the elderly with the FFT algorithm. Medicine & Science in Sports & Exercise. 2006; 38:1674-1681.

Qian J, Pei L, Ma J, Ying R, Liu P. Vector graph assisted pedestrian dead reckoning using an unconstrained smartphone. Sensors. 2015;15:5032-5057.

Van Schooten KS, Pijnappels M, Rispen SM, Elders PJ, Lips P, Daffertshofer A, et al. Daily-life gait quality as predictor of falls in older people: a 1-year prospective cohort study. PLoS One. 2016;11: e0158623.

Kikkert LHJ, Vuillerme N, van Campen JP, Appels BA, Hortobágyi T, & Lamoth CJC. Gait characteristics and their discriminative power in geriatric patients with and without cognitive impairment. Journal of neuroengineering and rehabilitation. 2017;14:1-10.

Liaw A, Wiener M. Classification and regression by randomForest. R News. 2002;2:18-22.

Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Advances In Neural Information Processing Systems. 2017:4765-4774.

Shapley LS. A value for n-person games. Contributions to the Theory of Games. 1953;2:307-317.

Hillel I, Gazit E, Nieuwoer A, Avanzino L, Rochester L, Cereatti A, & et al. Is every-day walking in older adults more analogous to dual-task walking or to usual walking? Elucidating
the gaps between gait performance in the lab and during 24/7 monitoring. European review of aging and physical activity. 2019;16:1-12.

[34] Lamoth CJC, Ainsworth E, Polomski W, Houdijk H. Variability and stability analysis of walking of transfemoral amputees. Medical Engineering & Physics. 2010;32:1009-1014.

[35] Van Dieën JH, Reeves NP, Kawchuk G, Van Dillen LR, & Hodges PW. Motor control changes in low back pain: divergence in presentations and mechanisms. Journal of Orthopaedic & Sports Physical Therapy. 2019;49:370-379.

[36] Prins MR, Griffioen M, Veeger TT, Kiers H, Meijer OG, van der Wurff P, et al. Evidence of splinting in low back pain? A systematic review of perturbation studies. European Spine Journal. 2018;27:40-59.

[37] Cholewicki J, Simons AP, Radebold A. Effects of external trunk loads on lumbar spine stability. Journal of biomechanics. 2000;33:1377-85.

[38] Mok NW, Hodges PW. Movement of the lumbar spine is critical for maintenance of postural recovery following support surface perturbation. Experimental brain research. 2013;231:305-13.

[39] Van den Hoorn W, Bruijn SM, Meijer OG, Hodges PW & Van Dieën JH. Mechanical coupling between transverse plane pelvis and thorax rotations during gait is higher in people with low back pain. Journal of biomechanics. 2012;45,342-347.

[40] Madeleine P. On functional motor adaptations: from the quantification of motor strategies to the prevention of musculoskeletal disorders in the neck–shoulder region. Acta Physiologica. 2010;199:1-46.

[41] Meier ML, Vrana A, Schweinhardt P. Low back pain: the potential contribution of supraspinal
motor control and proprioception. The Neuroscientist. 2019;25:583-96.

[42] Schinkel IA, Nairn BC, Drake JD. Investigation of trunk muscle co-contraction and its association with low back pain development during prolonged sitting. Journal of Electromyography and Kinesiology. 2013;23:778-86.

[43] Paul CP, Schoorl T, Zuiderbaan HA, Doulabi BZ, van der Veen AJ, van de Ven PM, et al. Dynamic and static overloading induce early degenerative processes in caprine lumbar intervertebral discs. PloS one. 2013;8.e62411.

[44] Parker R, Bergman E, Mntambo A, Stubbs S, & Wills M. Levels of PA in people with chronic pain. The South African Journal of Physiotherapy. 2017;73:1.

[45] Van Dieën J H, Reeves N P, Kawchuk G, Van Dillen LR. Analysis of motor control in patients with low back pain: a key to personalized care?. Journal of Orthopaedic & Sports Physical Therapy. 2019;49:380-388.