Wearable sensor devices can automatically identify the ON-OFF status of patients with Parkinson's disease through an interpretable machine learning model

Xiaolong Wu[1,2†], Lin Ma[3†], Penghu Wei[1,2], Yongzhi Shan[1,2], Piu Chan[4], Kailiang Wang[1,2]* and Guoguang Zhao[1,2,5]*

[1]Department of Neurosurgery, Xuanwu Hospital of Capital Medical University, Beijing, China, [2]International Neuroscience Institute (China-INI), Beijing, China, [3]Department of Neurorehabilitation, Rehabilitation Medicine of Capital Medical University, China Rehabilitation Research Centre, Beijing, China, [4]Department of Neurology and Neurobiology, Xuanwu Hospital of Capital Medical University, Beijing, China, [5]Beijing Municipal Geriatric Medical Research Center, Beijing, China

Introduction: Accurately and objectively quantifying the clinical features of Parkinson's disease (PD) is crucial for assisting in diagnosis and guiding the formulation of treatment plans. Therefore, based on the data on multi-site motor features, this study aimed to develop an interpretable machine learning (ML) model for classifying the "OFF" and "ON" status of patients with PD, as well as to explore the motor features that are most associated with changes in clinical symptoms.

Methods: We employed a support vector machine with a recursive feature elimination (SVM-RFE) algorithm to select promising motion features. Subsequently, 12 ML models were constructed based on these features, and we identified the model with the best classification performance. Then, we used the SHapley Additive exPlanations (SHAP) and the Local Interpretable Model agnostic Explanations (LIME) methods to explain the model and rank the importance of those motor features.

Results: A total of 96 patients were finally included in this study. The naive Bayes (NB) model had the highest classification performance (AUC = 0.956; sensitivity = 0.8947, 95% CI 0.6686−0.9870; accuracy = 0.8421, 95% CI 0.6875−0.9398). Based on the NB model, we analyzed the importance of eight motor features toward the classification results using the SHAP algorithm. The Gait: range of motion (RoM) Shank left (L) (degrees) [Mean] might be the most important motor feature for all classification horizons.

Conclusion: The symptoms of PD could be objectively quantified. By utilizing suitable motor features to construct ML models, it became possible to intelligently identify whether patients with PD were in the "ON" or "OFF" status. The variations in these motor features were significantly correlated with improvement rates in patients' quality of life. In the future, they might act as objective digital biomarkers to elucidate the changes in symptoms observed in

patients with PD and might be used to assist in the diagnosis and treatment of patients with PD.

Introduction

Parkinson's disease (PD) is a chronic degenerative disease of the central nervous system and is characterized by the degeneration or loss of dopaminergic neurons in the substantia nigra and the appearance of Lewy bodies. The clinical features of PD include bradykinesia, rest tremor, muscular rigidity, and postural impairment (1–3). As a result, the objective quantitative assessment of these clinical features plays a crucial role in diagnosis and guiding the formulation of treatment plans. The Movement Disorder Society-Unified Parkinson's Disease Rating Scale (MDS-UPDRS) is commonly used to measure the severity of patients with PD (4). The MDS-UPDRS-Part III is widely used to assess movement disorder in patients with PD. It is a semi-quantitative measurement and consists of 18 items (4). However, some limitations might be found in the clinical assessment by the MDS-UPDRS-Part III. First, it is not an objective, quantifiable evaluation method, and the evaluation process requires a specific physician and a lot of time. Second, the evaluation results are affected by the doctor's experience, the cognitive performance of the patients, and the surrounding environment; thus, the accuracy and objectivity of the evaluation results are limited (5–7). Finally, the symptoms of patients with PD fluctuate significantly, which poses great challenges for clinical assessment (8).

In recent years, the use of different sensor devices for the quantitative evaluation of motor ability in patients with PD has been increasingly explored. Many studies supported the reliability of motor data collected by wearable sensor devices (WSD) to assist in the diagnosis of patients with PD and the assessment of disease progression and to guide clinical practice (9–14). Some studies have combined machine learning (ML) or deep learning (DL) algorithms with WSD to estimate MDS-UPDRS-Part III and assist in the diagnosis of PD (15–17). Although it brought great opportunities and potential for an intelligent evaluation of PD, some challenges might affect the accuracy of their results (15, 16, 18). First, there is still no consensus on what motor features are most relevant to the changes in clinical symptoms (between the "OFF" and "ON" status, or between patients with PD and healthy elderly people). In addition, some studies had small sample sizes, and the WSDs were placed only on a single site to collect patient motor features. Moreover, some WSDs were cumbersome and inconvenient to wear, causing inconvenience to patients with PD during use. Finally, although the classification performance of multiple predictive models proved to be promising, their interpretability for classifying the "OFF" and "ON" status of patients with PD is still limited (19). These factors hinder the application of WSDs in objectively quantifying the clinical features of patients with PD. The question is whether we can maximize the accurate quantification of clinical features in patients with PD using simpler wearing methods and fewer motor features.

Therefore, based on the data on multi-site motor features, this study aimed to develop an interpretable ML model for classifying the "OFF" and "ON" status of patients with PD, as well as to explore the motor features most associated with changes in clinical symptoms. In addition, we used SHapley Additive exPlanations (SHAP) and Local Interpretable Model agnostic Explanations (LIME) methods (20, 21) to explain the model and rank the importance of those motor features.

Materials and methods

Study design

This was a retrospective observational study. It was approved by the Institutional Ethics Committee of Xuanwu Hospital and performed according to the principles of the Declaration of Helsinki. All personal information was made anonymous before analysis. We retrospectively analyzed the following clinical information: personal and medical history, the history of drug substance intake, current medication usage, MDS-UPDRS-Part III (4), the results of motor assessment, Hoehn and Yahr scale (HY) (22), and Activities of Daily Living (ADL, assessed by the Barthel Index) (23). We applied the levodopa challenge test to assess the responsiveness of patients with PD to levodopa medication and to distinguish between patients in the "ON" and "OFF" status (24, 25). The "OFF" status was defined as being off dopamine agonists for 72 h and off antiparkinsonian drugs for 12 h. The "ON" status was the best statement after taking antiparkinsonian medications (∼2 h after taking the medicine). The MDS-UPDRS-Part III scores and motor assessment were evaluated at 8:00 AM during the "OFF" status. After taking the medicine, MDS-UPDRS-Part III was recorded at 1, 2, 3, and 4 h, and the best performance was selected for statistical analyses. MDS-UPDRS-Part III improvement rate = (MDS-UPDRS-Part III score "OFF" – MDS-UPDRS-Part III score "ON")/MDS-UPDRS-Part III score "OFF" × 100%. ADL improvement rate = (ADL score in the "ON" status – ADL score in the "OFF" status)/ADL score in the "OFF" status × 100%.

Patients

From January 2019 to December 2020, data on patients with PD who were recruited from the neurology clinic of Xuanwu Hospital, Beijing, China, were retrospectively collected. According to the Movement Disorder Society Clinical Diagnostic Criteria for PD (26), they were diagnosed with clinically established PD by a movement disorders specialist. The inclusion and exclusion criteria are shown in Supplementary Table S1.

Tools for motor assessment

Six Opal™ Movement Monitors (APDM, Inc., Portland, OR, United States) were placed on the bilateral wrists and ankles, the anterior sternum, and the lower back (Figure 1A). Each movement monitor included a three-axis accelerometer, a three-axis gyro, a three-axis magnetometer, and a temperature sensor. It was secured to the patient using optional straps that could be connected to the host via wireless communication. For detailed information about Opal™ Movement Monitors and gait data, please visit the official website (APDM, Inc., Portland, OR, United States, https://apdm.com/wearable-sensors/).

Motor assessment procedures

We utilized the Instrumented Long Walk (IWalk) Test protocol for gait data collection. The IWalk analysis algorithms automatically process recorded movement data and provide objective measures related to gait and turning. For more detailed information, please visit the official website (APDM, Inc., Portland, OR, United States, https://apdm.com/wearable-sensors/). The patients were asked to wear comfortable clothing and walking shoes that did not bind or impede their movement in any significant way. After signing the informed consent form, they wore six Opal™ Movement Monitors and walked on a straight 10 m walkway at a comfortable pace. All patients walked from the start point to the end of the 10 m walkway and then turned around to return to the starting point (Figure 1B). All patients were required to walk for at least 1 min. These recorded parameters could be transmitted to the computing center in real time by wireless transmission technology for three-dimensional movement posture reconstruction, and then the gait, posture balance, arm swing, movement coordination, etc. could be evaluated.

Statistical analysis

The R software (version 4.1.3; R Core Team) was used to perform all statistical analyses. If there were missing values in the data included in this study, the k-nearest neighbors interpolation method was used for imputation (27). Continuous quantitative variables are expressed as means, standard deviations, medians, and interquartile ranges (IQRs). Categorical variables are expressed as total numbers and percentages. The support vector machine – recursive feature elimination (SVM-RFE) algorithm was used to filter the motor features with a five-fold cross validation and five re-samplings (28). We used the Wilcoxon signed-rank test to compare the differences in motor features between the "OFF" and "ON" status. The correlation between the ADL improvement rate and the improvement rate of motor features was explored using the Pearson correlation test.

Classification model construction

A total of 12 ML models, namely, Adaboost, LogitBoost, XGBoost, logistic regression (LR), random forest (RF), support vector machine (SVM), neural network (NN), k-nearest neighbors (KNN), decision tree C5.0, naive Bayes (NB), gradient boosting machine (GBM), and multilayer perceptron (MLP), were used to develop the classification models. The 10-fold cross-validation and RandomSearch for hyperparameters were used for training each ML model. We calculated the area under the receiver operating characteristic curve, sensitivity, specificity, accuracy, positive predictive values, negative predictive values, recall, and F1 score. The area under the curve (AUC) was used as the main index to evaluate the classification performance of each model. If the AUC value was the same or similar, the sensitivity, specificity, accuracy, positive predictive values, negative predictive values, recall, and F1 score were referred.

Interpretation tool for the model

The interpretation of the ML model of the best classification performance was performed using the SHAP and LIME methods (20, 21). SHAP is a common method to analyze the contribution and influence of each motor feature toward the overall performance of the classification model. The SHAP value is calculated to show how important each motor feature is for the target variable, either positively or negatively. The SHAP and LIME methods were used to explore the contributions of each variable to the overall performance of the classification model and the classification of an instance, respectively.
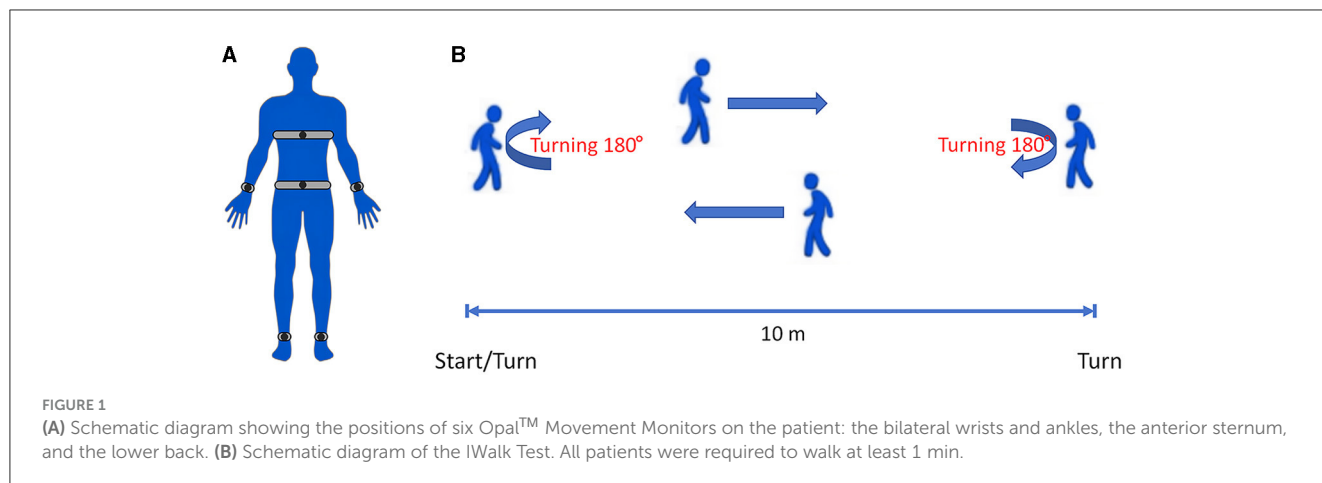
Results

Patient characteristics

From January 2019 to December 2020, a total of 110 patients with PD were screened, and based on the inclusion and exclusion criteria, 96 patients were ultimately included in this study. The reasons for excluding 14 patients were as follows: some patients were unable to complete the IWalk Test task owing to heart and orthopedic diseases; some patients had other neurological or psychiatric disorders; some had concurrent other neurological or psychiatric disorders; and a significant amount of information was missing for some patients. Supplementary Figure S1 presents the patient screening process. The demographic and clinical information of patients with PD are presented in Table 1. We randomly divided a total of 192 patients (the "OFF" and "ON" status of 96 patients) into two parts: 80% (training dataset, $n = 154$) of the subjects were used to train the classification model, and 20% (validating dataset, $n = 38$) were used to validate. Then, based on the training dataset, we used the SVM-RFE algorithm to find eight potential motor features for developing the classification model.

Machine learning model construction and evaluation

Based on the training dataset, Adaboost, LogitBoost, XGBoost, LR, RF, SVM, NN, KNN, decision tree C5.0, NB, GBM, and MLP classification models were constructed, and the AUCs of the validating datasets were 0.898, 0.802, 0.927, 0.934, 0.909, 0.95,

FIGURE 1
(A) Schematic diagram showing the positions of six Opal^TM Movement Monitors on the patient: the bilateral wrists and ankles, the anterior sternum, and the lower back. (B) Schematic diagram of the IWalk Test. All patients were required to walk at least 1 min.

TABLE 1 Demographic and clinical information of patients with PD.

Characteristics	Baseline
Demographic	
Patient numbers	96
Age, Median [IQR]	62 [10.25]
Gender, male [%]	41 [42.71%]
Education, high school or higher [%]	57 [59.38%]
Clinical	
Disease duration, Median [IQR]	9 [6]
HY stage (the "OFF" status)	
I/II/III/IV/V	0/53/35/6/2
MDS-UPDRS-III (the "OFF" status), Median [IQR]	57 [26]
MDS-UPDRS-III (the "ON" status), Median [IQR]	30.5 [20]
MDS-UPDRS-III improvement rate, Median [IQR]	47.3% [26.19%]
LEDD (mg), Median [IQR]	300 [75]

LEDD, levodopa-equivalent daily dose; SD, standard deviation; IQR, interquartile range; HY stage, Hoehn & Yahr stage.

0.953, 0.945, 0.88, 0.956, 0.9, and 0.917, respectively (Table 2 and Figure 2). After a comprehensive comparison, the NB model had the highest classification performance (AUC = 0.956; sensitivity = 0.8947, 95% CI 0.6686–0.9870; accuracy = 0.8421, 95% CI 0.6875–0.9398; positive predictive values = 0.8095, 95% CI 0.5809–0.9455; negative predictive values = 0.8824, 95% CI 0.6356–0.9854; recall = 0.8947; F1 score = 0.85).

Interpretation of the NB model with the SHAP and LIME methods

Based on the NB model, we analyzed the importance of eight motor features toward the classification results using the SHAP algorithm. The motor feature importance ranking is shown in Figure 3A. The Gait: range of motion (RoM) Shank left (L) (degrees) [Mean] might be the most important motor feature for all classification horizons, followed by the Gait: Stride Length L
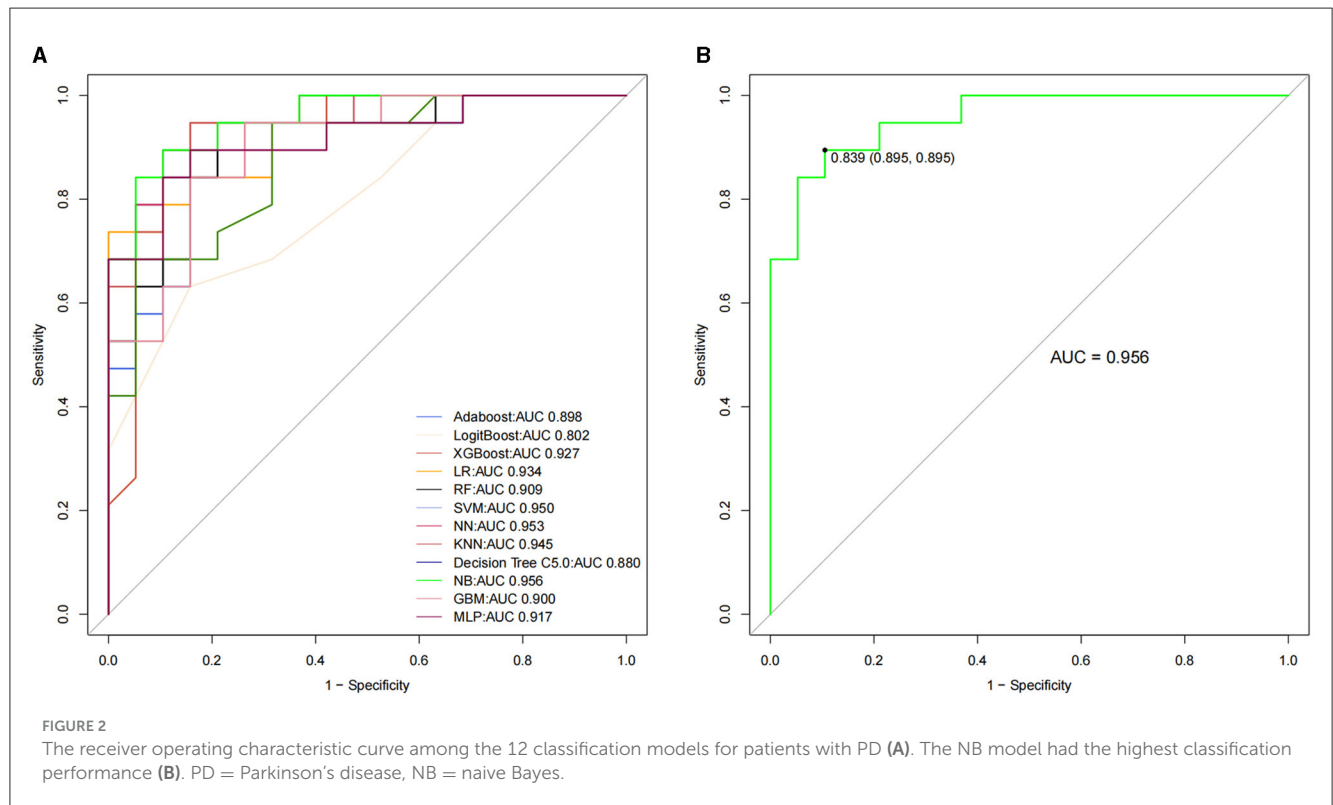
(%stature) [Mean], the Gait: Stride Length R (%stature) [Mean], the Gait: RoM Arm R (degrees) [Mean], the Gait: Peak Shank Velocity R (degrees/s) [Mean], the Gait: Peak Horiz. Trunk Velocity (degrees/s) [Mean], the Gait: Peak Shank Velocity L (degrees/s) [Mean], and the Turn: Peak Velocity (degrees/s) [Mean]. There were positive and negative correlations between the motor features and classification results. In Figure 3B, the colors of points show whether the motor feature was high (in aurantium) or low (in purple) in this study. We found that the increase in the value of all motor features had a positive effect on the "ON" status, driving the prediction of the "ON" status.

The LIME method was also used to explain how eight motor features contributed to the classification results (Figure 4). The sequentially important motor features that contributed to the "OFF" status included the Gait: RoM Shank L (degrees) [Mean] < = 53.4, the Gait: Stride Length L (%stature) [Mean] < = 55.7, the Gait: Stride Length R (%stature) [Mean] < = 55.4, the Turn: Peak Velocity (degrees/s) [Mean] < = 94.8, the Gait: RoM Arm R (degrees) [Mean] < = 7.4, the 7.4 < Gait: RoM Arm R (degrees) [Mean] < = 13.4, the 13.4 < Gait: RoM Arm R (degrees) [Mean] < = 22.9, the Gait: Peak Shank Velocity L (degrees/s) [Mean] < = 269, the Gait: Peak Shank Velocity R (degrees/s) [Mean] < = 271, the Gait: Peak Horiz. Trunk Velocity (degrees/s) [Mean] < = 15.5, and the 15.5 < Gait: Peak Horiz. Trunk Velocity (degrees/s) [Mean] < = 21.8. The sequentially important motor features that contributed to the "ON" status included the 71.0 < Gait: RoM Shank L (degrees) [Mean], the 63.3 < Gait: RoM Shank L (degrees) [Mean] < = 71.0, the 74.6 < Gait: Stride Length L (%stature) [Mean], the 67.5 < Gait: Stride Length L (%stature) [Mean] < = 74.6, the 74.8 < Gait: Stride Length R (%stature) [Mean], the 67.3 < Gait: Stride Length R (%stature) [Mean] < = 74.8, the 144.3 < Turn: Peak Velocity (degrees/s) [Mean], the 22.9 < Gait: RoM Arm R (degrees) [Mean], the 364 < Gait: Peak Shank Velocity L (degrees/s) [Mean], the 318 < Gait: Peak Shank Velocity L (degrees/s) [Mean] < = 364, the 373 < Gait: Peak Shank Velocity R (degrees/s) [Mean], the 320 < Gait: Peak Shank Velocity R (degrees/s) [Mean] < = 373, and the 27.8 < Gait: Peak Horiz. Trunk Velocity (degrees/s) [Mean]. Similar to the SHAP method, we also found that the increase in the value of all motor features has a positive effect on the "ON" status, driving the prediction of the "ON" status.

TABLE 2 Performance of each classification model in the validating dataset.

Model	AUC	Sensitivity	Accuracy	Positive Pred Value	Negative Pred Value	Recall	F1 score
Adaboost	0.898	0.7895 (0.5443, 0.9395)	0.8158 (0.6567, 0.9226)	0.8333 (0.5858, 0.9642)	0.8000 (0.5634, 0.9427)	0.6842	0.7428
LogitBoost	0.802	0.6842 (0.4345, 0.8742)	0.6842 (0.5135, 0.8250)	0.6842 (0.4345, 0.8742)	0.6842 (0.4345, 0.8742)	0.6842	0.6842
XGBoost	0.927	0.7895 (0.5443, 0.9395)	0.8421 (0.6875, 0.9398)	0.8824 (0.6356, 0.9854)	0.8095 (0.5809, 0.9455)	0.7894	0.8333
LR	0.934	0.7895 (0.5443, 0.9395)	0.8421 (0.6875, 0.9398)	0.8824 (0.6356, 0.9854)	0.8095 (0.5809, 0.9455)	0.7894	0.8333
RF	0.909	0.7895 (0.5443, 0.9395)	0.8158 (0.6567, 0.9226)	0.8333 (0.5858, 0.9642)	0.8000 (0.5634, 0.9427)	0.7894	0.8108
SVM	0.95	0.7895 (0.5443, 0.9395)	0.8421 (0.6875, 0.9398)	0.8824 (0.6356, 0.9854)	0.8095 (0.5809, 0.9455)	0.7894	0.8333
NN	0.953	0.7895 (0.5443, 0.9395)	0.8421 (0.6875, 0.9398)	0.8824 (0.6356, 0.9854)	0.8095 (0.5809, 0.9455)	0.7894	0.8333
KNN	0.945	0.8421 (0.6042, 0.9662)	0.8684 (0.7191, 0.9559)	0.8889 (0.6529, 0.9862)	0.8500 (0.6211, 0.9679)	0.8421	0.8648
DT C5.0	0.88	0.7895 (0.5443, 0.9395)	0.7368 (0.5690, 0.8660)	0.7143 (0.4782, 0.8872)	0.7647 (0.5010, 0.9319)	0.7894	0.75
NB	0.956	0.8947 (0.6686, 0.9870)	0.8421 (0.6875, 0.9398)	0.8095 (0.5809, 0.9455)	0.8824 (0.6356, 0.9854)	0.8947	0.85
GBM	0.9	0.7895 (0.5443, 0.9395)	0.8158 (0.6567, 0.9226)	0.8333 (0.5858, 0.9642)	0.8000 (0.5634, 0.9427)	0.7894	0.8108
MLP	0.917	0.7895 (0.5443, 0.9395)	0.8421 (0.6875, 0.9398)	0.8824 (0.6356, 0.9854)	0.8095 (0.5809, 0.9455)	0.7894	0.8333

AUC, area under the curve; LR, logistic regression; RF, random forest; SVM, support vector machine; NN, neural network; KNN, k-nearest neighbors; DT, decision tree; NB, naive Bayes; GBM, gradient boosting machine; MLP, multilayer perceptron.

FIGURE 2
The receiver operating characteristic curve among the 12 classification models for patients with PD **(A)**. The NB model had the highest classification performance **(B)**. PD = Parkinson's disease, NB = naive Bayes.

Differences of eight motor feature values between the "OFF" and "ON" status and correlation analysis

The Wilcoxon signed-rank test was performed to compare the values of eight motor features in the "OFF" and "ON" status, and the results showed that the motor features had significantly higher values in the "ON" status than in the "OFF" status ($p < 0.05$). The details are available in Figure 5

and Supplementary Table S2. The Pearson correlation analysis presented that the improvement rate of each motor feature was significantly and positively correlated with the ADL improvement rate of patients ($p < 0.05$), except for Gait: RoM Arm R (degrees) [Mean] ($p = 0.083$) (Figure 6). This indicated that, compared to the "OFF" status, in the "ON" status, the values of the eight motor features of patients with PD were significantly higher, thus providing patients with a better motor status and quality of life.
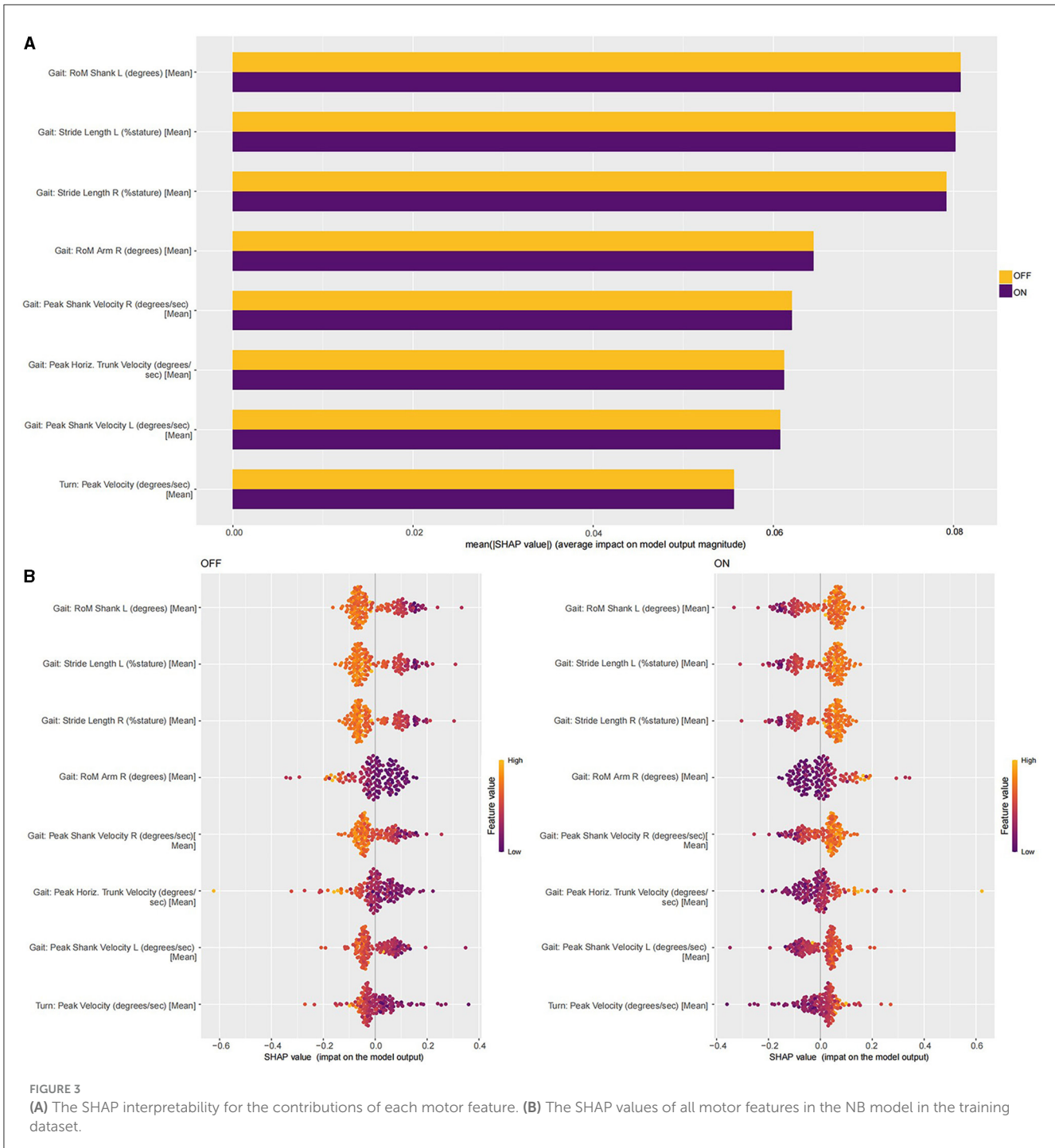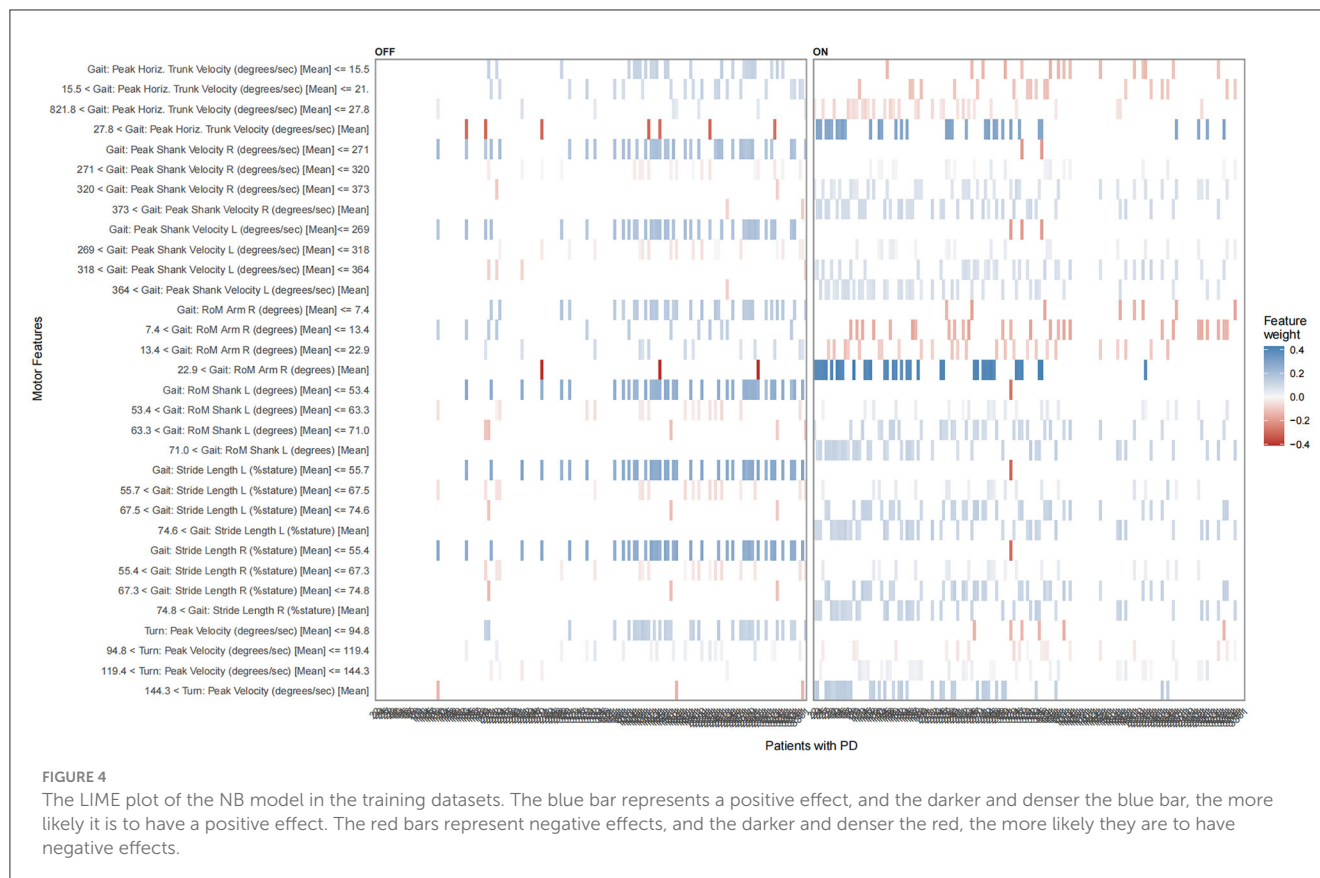
FIGURE 3
(A) The SHAP interpretability for the contributions of each motor feature. **(B)** The SHAP values of all motor features in the NB model in the training dataset.

Discussion

In our study, we aimed to develop an interpretable ML model for classifying the "OFF" and "ON" status of patients with PD, as well as to explore the motor features most associated with changes in clinical symptoms. In total, 12 ML classification models were developed and validated to classify the "OFF" and "ON" status of patients with PD. The classification performance of the XGBoost model outperformed the Adaboost, LogitBoost, LR, RF, SVM, NN, KNN, decision tree C5.0, NB, GBM, and MLP

models. We performed the interpretation for the NB model by the SHAP and LIME methods, which provided a guarantee for its performance and clinical interpretability. In addition, it helped us better understand the classification process of the NB model. Some important motor features and the importance of ranking those features were identified to be associated with identifying the status of patients with PD; this was performed by the SHAP and LIME methods to interpret the NB model. A total of five motor features describing the lower extremities, two motor features describing the trunk of the body, and one motor feature describing the upper

FIGURE 4
The LIME plot of the NB model in the training datasets. The blue bar represents a positive effect, and the darker and denser the blue bar, the more likely it is to have a positive effect. The red bars represent negative effects, and the darker and denser the red, the more likely they are to have negative effects.

limb were included (Figure 3). Then it was confirmed that each of the motor features included in the NB model had significant differences between the "OFF" and "ON" status. The correlation analysis showed the ADL improvement rates were closely related to the improvement rates for each motor feature. The most important finding was that the motor features of the lower limbs might be better for classification performance and could more truly reflect the motor symptoms of patients, mainly including the Gait: RoM Shank L (degrees) [Mean], the Gait: Stride Length L (%stature) [Mean], the Gait: Stride Length R (%stature) [Mean], the Gait: Peak Shank Velocity R (degrees/s) [Mean], and the Gait: Peak Shank Velocity L (degrees/s) [Mean]. The symptoms of PD could be objectively quantified. By utilizing suitable motor features to construct ML models, it became possible to intelligently identify whether patients with PD were in the "ON" or "OFF" status.

The clinical diagnosis of PD is mainly based on motor symptoms, including muscular rigidity, rest tremor, bradykinesia, and postural impairment. As the duration of levodopa treatment in patients with PD increases, the duration and stability of symptom benefits decrease. As a result, the wake time is divided into the time of reduction in PD symptoms and improvement in functional status, that is, the time when levodopa provides favorable benefits (the "ON" status), and the time of PD symptom re-onset and functional status decline, that is, the time when levodopa does not provide good benefits (the "OFF" status) (25, 29). Transitions between the "ON" and "OFF" status are referred to as motor fluctuations. Motor fluctuations may be an important cause of later disability in patients with PD (30). In addition

to the decline in physical function, mood swings and adverse psychological problems in patients with PD are often found in patients with motor fluctuations (31). The MDS-UPDRS-Part III is the main diagnostic tool used in clinical diagnostics and research. However, the results of many studies show the MDS-UPDRS-Part III scores are always not satisfactory in diagnosing PD because they may also lead to a certain degree of misdiagnosis (32). In addition, the evaluation process is time-consuming and heavy, which increases the burden on patients and doctors, and it is reported that many patients with PD have no opportunity to consult with a PD specialist or neurologist for professional clinical assessment, which could lead to rapid progression and disease-related complications (33).

Therefore, doctors and researchers have a strong demand for more objective and continuous evaluation and monitoring methods. With the rise of new sensor-based wearable technology, the traditional diagnosis and treatment model is changing to achieve a more objective diagnostic assessment. The WSD refers to portable and mobile devices worn on the body or embedded in clothing, such as smart glasses, watches, clothes, and pressure shoes, which contain hardware and software technologies and have special functions for collecting spatiotemporal kinematic parameters, data processing, transmission, and storage. At present, the WSD used in the field of PD can realize quantitative evaluation of PD through specific motor tasks and establish data models so that doctors can accurately analyze the movement status of patients (5, 14, 34). Compared with the results of MDS-UPDRS-Part III evaluated by doctors, WSDs were reported as a method with higher objectivity,
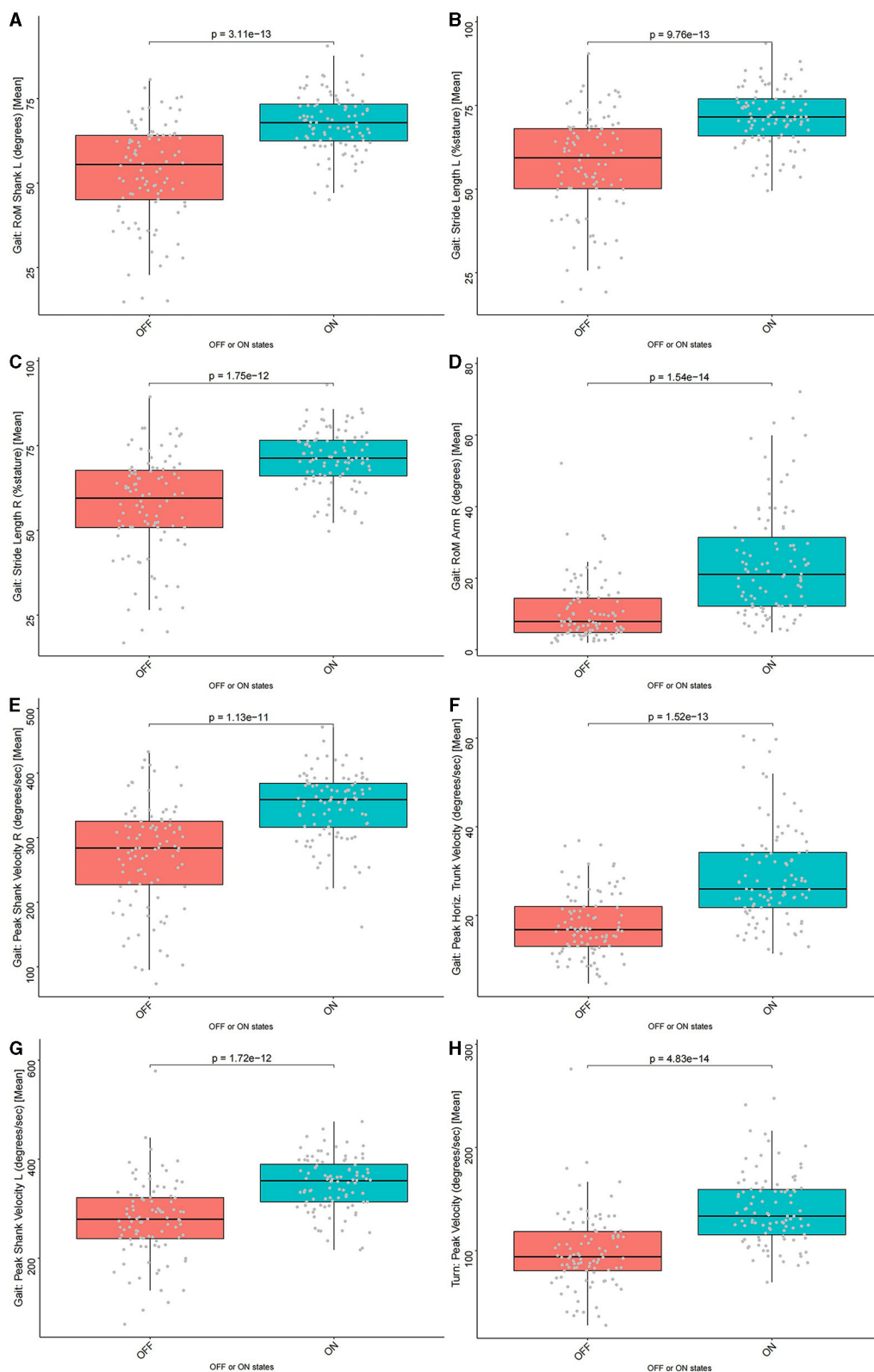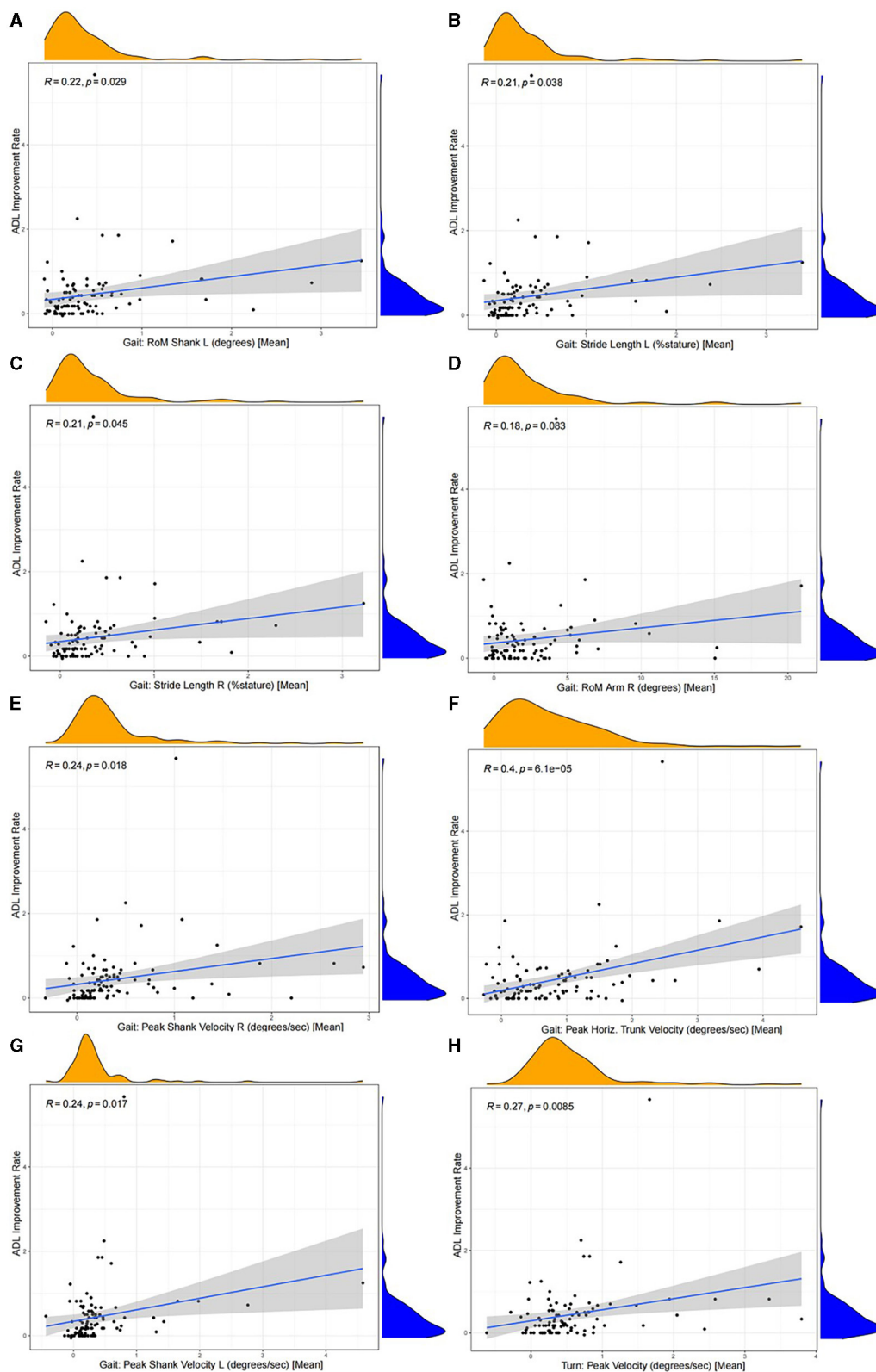
FIGURE 5
The comparison of eight motor feature values between the "OFF" and "ON" status. All motor features had significantly higher values in the "ON" status than in the "OFF" status ($p < 0.05$). **(A)** Gait: RoM Shank L (degrees) [Mean] ($p = 3.11e-13$), **(B)** Gait: Stride Length L (%stature) [Mean] ($p = 9.76e-13$), **(C)** Gait: Stride Length R (%stature) [Mean] ($p = 1.75e-12$), **(D)** Gait: RoM Arm R (degrees) [Mean] ($p = 1.54e-14$), **(E)** Gait: Peak Shank Velocity R (degrees/s) [Mean] ($p = 1.13e-11$), **(F)** Gait: Peak Horiz. Trunk Velocity (degrees/s) [Mean] ($p = 1.52e-13$), **(G)** Gait: Peak Shank Velocity L (degrees/s) [Mean] ($p = 1.72e-12$), and **(H)** Turn: Peak Velocity (degrees/s) [Mean] ($p = 4.83e-14$).

FIGURE 6
The correlation (Pearson correlation analysis) between the improvement rate of each motor feature and the ADL improvement rate. The improvement rate of each motor feature was significantly correlated positively with the ADL improvement rate of patients ($p < 0.05$), except for Gait: RoM Arm R (degrees) [Mean] (R = 0.18, p = 0.083). **(A)** Gait: RoM Shank L (degrees) [Mean] (R = 0.22, p = 0.029), **(B)** Gait: Stride Length L (%stature) [Mean] (R = 0.21, p = 0.038), **(C)** Gait: Stride Length R (%stature) [Mean] (R = 0.21, p = 0.045), **(D)** Gait: RoM Arm R (degrees) [Mean] (R = 0.18, p = 0.083), **(E)** Gait: Peak Shank Velocity R (degrees/s) [Mean] (R = 0.24, p = 0.018), **(F)** Gait: Peak Horiz. Trunk Velocity (degrees/s) [Mean] (R = 0.4, p = 6.1e−05), **(G)** Gait: Peak Shank Velocity L (degrees/s) [Mean] (R = 0.24, p = 0.017), and **(H)** Turn: Peak Velocity (degrees/s) [Mean] (R = 0.27, p = 0.0085).

accuracy, and sensitivity for evaluating the status of patients' motor abilities, such as bradykinesia, dyskinesia, tremor, and freezing of gait (5, 15). Therefore, by using WSDs to collect motor data from patients with PD, the internal relationship between motor data and MDS-UPDRS-Part III scores was studied, and finally, the purpose of quantifying the MDS-UPDRS-Part III scores of patients with PD was achieved. It is very important for the diagnosis and treatment of PD.

In this study, the Gait: RoM Shank L (degrees) [Mean] might be the most important motor feature. It refers to the flexion and extension range of motion of the knee joint (sagittal plane). Both the patients with PD and healthy elderly people could show reduced RoM of the shank. The reduction in RoM of the shank was more pronounced in patients with PD, and it was significantly associated with PD progression (16, 17). It had also been suggested that the RoM of the lower limb joint in patients with PD was severely reduced, and at the same time, the hip and knee significantly moved in the direction of flexion (16). Compared with healthy people individuals, the patients with PD not only had more severely reduced RoM of the lower limb joint but also showed more lower limb flexion movements throughout the gait cycle. Biomechanically, this might be to counteract the enhanced trunk flexion (16). The Gait: Stride Length L (%stature) [Mean], Gait: Stride Length R (%stature) [Mean], Gait: Peak Shank Velocity R (degrees/s) [Mean], and Gait: Peak Shank Velocity L (degrees/s) [Mean] bilaterally described the stride length and peak velocity of the shank. Both stride length and peak velocity of the shank were significantly larger in the "ON" status than in the "OFF" status. Previous studies have argued that variability analysis might be more sensitive in distinguishing gait disorders than other motor features, such as step length and step speed (18, 35, 36). In our study, however, the RoM of the lower limb joint, stride length, and peak velocity of the shank could clearly distinguish the status of patients with PD. However, in our study, the Gait: Peak Horiz. Trunk Velocity (degrees/s) [Mean] and Turn: Peak Velocity (degrees/s) [Mean] were less important than the lower limb motor features, the turn-related motor features were considered the most important factors to distinguish between patients with PD and healthy elderly people (36, 37). Similar to previous studies, patients with PD turned more slowly into the "OFF" status. Turning was a complex act that required everyone to change direction while keeping the trunk stable (37). The patients with PD seemed to have more difficulty turning because they could not precisely control the RoM of the lower limb joint, stride length, and peak velocity of the shank (37). In addition, previous studies suggested that the efficacy of levodopa in improving trunk-related symptoms in patients with PD was limited (38, 39). However, we found that the values of the Gait: Peak Horiz. Trunk Velocity (degrees/s) [Mean] and Turn: Peak Velocity (degrees/s) [Mean] were significantly higher in the "ON" status than in the "OFF" status ($p = 1.52e-13$; $p = 4.83e-14$). Furthermore, the improvement rates of two motor features were significantly and positively correlated with the ADL improvement rate of patients (R = 0.4, $p = 6.1e-05$; R = 0.27, $p = 0.0085$). This might indicate a significant effect of levodopa on improving trunk-related motor features in patients. A reduction in the RoM of arm swing was also an important motor feature for patients with PD (40). Similarly, in our study, the Gait: RoM Arm R (degrees)

[Mean] was significantly smaller in the "OFF" status than in the "ON" status. However, in the correlation analysis, we did not find a significant positive correlation between the ADL improvement rate and the improvement rate of the Gait: RoM Arm R (degrees) [Mean]. It might be because walking was an important factor affecting the quality of life of patients with PD and occupied most of the exercise time in daily life (41, 42).

With their inspiring classification performances, the ML and DL algorithms have been used for constructing models for classifying between patients with PD and healthy elderly people (43–46). However, the interpretability of the model and the importance of the features incorporated into the model were ambiguous. This is the first study to develop an interpretable ML model to classify the "OFF" and "ON" status of patients with PD based on the motor features of different body parts. Our study showed that the use of more sensors in the hospital or laboratory to cover the trunk and limbs of the body more comprehensively reflects the patient's mobility, balance, walking ability, and fall risk. However, this would increase the burden and discomfort of patients and is not conducive to the development of daily monitoring of patients with PD. Therefore, the current research is moving toward a minimum number of sensors worn and daily monitoring (41, 47). In the previous studies, the optimal locations for WSDs, including on the lower limbs, trunk (waist), and upper limbs, have been discussed (41, 48–55). According to published systematic reviews, the lower limbs might be the most common location (41, 48– 51). Some studies also suggested that the waist and upper limbs were the optimal locations for collecting kinematic parameters in patients with PD (41, 52–54). Peraza et al. (55) proposed an automatic gait analysis process based on DL algorithms, with data sourced from triaxial accelerometers placed on the lower limbs, trunk (waist), and upper limbs. The results showed that data from single triaxial accelerometers on the lower limbs and trunk (waist) performed better than those from the upper limbs in assessing gait in patients with PD and healthy elderly people (55). Mikos et al. (56) developed a freezing of gait detection feedback system for patients with PD, which was integrated into a single wearable device sensor attached to the ankle. The system exhibited excellent performance in classification (a sensitivity of 95.6% and a specificity of 90.2%) (56). Recently, Chen et al. (41) proposed a method for patients with PD based on an optimized interpretable DL architecture. They fixed five sensors on the lower limbs, trunk (waist), and upper limbs, to collect the motor data from patients with PD and healthy elderly people during a 10 m walking test (41). After analyzing the raw data and constructing a convolutional neural network classification model, they found that the data collected by the sensor at the waist performed best in classifying patients with PD and healthy elderly people (41). We not only constructed a classification model but also quantified the importance of motion features. In general, the importance of motor features in the lower limbs might be the highest compared to other parts of the body, which provided an evaluation method for the further selection of the best single sensor wearing position. We also found a significant effect of levodopa on improving trunk-related motor features in patients with PD, and this seemed to be contrary to previous research findings. All of these might become new directions for study in the field of Parkinson's disease and WSDs.

Some limitations are found. First, this was a retrospective study, and the absence of some data might have caused some bias in the study results. Second, all data were based on patients with PD from China, and the applicability of the NB model to other ethnic groups is unclear. Third, external verification was lacking. Therefore, the study results should be interpreted cautiously. Finally, our study did not compare patients with PD and healthy elderly people. However, contrasting the "ON" and "OFF" status of patients with PD has already demonstrated the capability of WSDs to objectively quantify the symptoms. Therefore, our study is effective.

Conclusion

The symptoms of PD could be objectively quantified. By utilizing suitable motor features to construct the ML models, it became possible to intelligently identify whether patients with PD were either in the "ON" or "OFF" status. The variations in these motor features were significantly correlated with improvement rates in patients' quality of life. In the future, they might act as objective digital biomarkers to elucidate the changes in symptoms observed in patients with PD and might be used to assist in the diagnosis and treatment of patients with PD.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

XW: Formal analysis, Writing – original draft. LM: Data curation, Formal analysis, Writing – original draft. PW:

Writing – review & editing. YS: Writing – review & editing. PC: Conceptualization, Writing – review & editing. KW: Conceptualization, Methodology, Writing – review & editing. GZ: Writing – review & editing, Conceptualization.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2024.1387477/full#supplementary-material

SUPPLEMENTARY FIGURE S1
A flowchart of the screening process.

References

1. Connolly BS, Lang AE. Pharmacological treatment of Parkinson disease: a review. *JAMA.* (2014) 311:1670–83. doi: 10.1001/jama.2014.3654

2. Dong-Chen X, Yong C, Yang X, Chen-Yu S, Li-Hua P. Signaling pathways in Parkinson's disease: molecular mechanisms and therapeutic interventions. *Signal Transduct Target Ther.* (2023) 8:73. doi: 10.1038/s41392-023-01353-3

3. Kalia LV, Lang AE. Parkinson's disease. *Lancet.* (2015) 386:896–912. doi: 10.1016/S0140-6736(14)61393-3

4. Goetz CG, Tilley BC, Shaftman SR, Stebbins GT, Fahn S, Martinez-Martin P, et al. Movement disorder society-sponsored revision of the unified Parkinson's disease rating scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov Disord.* (2008) 23:2129–70. doi: 10.1002/mds.22340

5. Han Y, Liu X, Zhang N, Zhang X, Zhang B, Wang S, et al. Automatic assessments of parkinsonian gait with wearable sensors for human assistive systems. *Sensors.* (2023) 23:2104. doi: 10.3390/s23042104

6. Parisi F, Ferrari G, Giuberti M, Contin L, Cimolin V, Azzaro C, et al. Inertial BSN-based characterization and automatic UPDRS evaluation of the gait task of Parkinsonians. *IEEE Trans Aff Comput.* (2016) 7:258–71. doi: 10.1109/TAFFC.2016.2549533

7. Post B, Merkus MP, de Bie RM, de Haan RJ, Speelman JD. Unified Parkinson's disease rating scale motor examination: are ratings of nurses, residents in neurology, and movement disorders specialists interchangeable? *Mov Disord.* (2005) 20:1577–84. doi: 10.1002/mds.20640

8. Diao JA, Raza MM, Venkatesh KP, Kvedar JC. Watching Parkinson's disease with wrist-based sensors. *NPJ Digital Med.* (2022) 5:73. doi: 10.1038/s41746-022-00619-4

9. Lu R, Xu Y, Li X, Fan Y, Zeng W, Tan Y, et al. Evaluation of wearable sensor devices in Parkinson's disease: a review of current status and future prospects. *Parkinson's Disease.* (2020) 2020:4693019. doi: 10.1155/2020/4693019

10. Del Din S, Godfrey A, Mazzà C, Lord S, Rochester L. Free-living monitoring of parkinson's disease: lessons from the field. *Mov Disord.* (2016) 31:1293–313. doi: 10.1002/mds.26718

11. Mancini M, Bloem BR, Horak FB, Lewis SJG, Nieuwboer A, Nonnekes J. Clinical and methodological challenges for assessing freezing of gait: future perspectives. *Mov Disord.* (2019) 34:783–90. doi: 10.1002/mds.27709

12. Rovini E, Maremmani C, Cavallo F. How wearable sensors can support parkinson's disease diagnosis and treatment: a systematic review. *Front Neurosci.* (2017) 11:555. doi: 10.3389/fnins.2017.00555

13. Thorp JE, Adamczyk PG, Ploeg HL, Pickett KA. Monitoring motor symptoms during activities of daily living in individuals with Parkinson's disease. *Front Neurol.* (2018) 9:1036. doi: 10.3389/fneur.2018.01036

14. Caballol N, Bayés À, Prats A, Martín-Baranera M, Quispe P. Feasibility of a wearable inertial sensor to assess motor complications and treatment in Parkinson's disease. *PLoS ONE.* (2023) 18:e0279910. doi: 10.1371/journal.pone.0279910

15. Bouça-Machado R, Jalles C, Guerreiro D, Pona-Ferreira F, Branco D, Guerreiro T, et al. Gait kinematic parameters in Parkinson's disease: a systematic review. *J Parkinson's Disease.* (2020) 10:843–53. doi: 10.3233/JPD-201969

16. Pistacchi M, Gioulis M, Sanson F, De Giovannini E, Filippi G, Rossetto F, et al. Gait analysis and clinical correlations in early Parkinson's disease. *Funct Neurol.* (2017) 32:28–34. doi: 10.11138/FNeur/2017.32.1.028

17. Sale P, Pandis MFD, Vimercati SL, Sova I, Foti C, Tenore N, et al. The relation between Parkinson's disease and ageing - Comparison of the gait patterns of young Parkinson's disease subjects with healthy elderly subjects. *Eur J Phys Rehabil Med.* (2013) 4:161–7.

18. Hausdorff JM, Cudkowicz ME, Firtion R, Wei JY, Goldberger AL. Gait variability and basal ganglia disorders: stride-to-stride variations of gait cycle timing in Parkinson's disease and Huntington's disease. *Mov Disord.* (1998) 13:428–37. doi: 10.1002/mds.870130310

19. Mirelman A, Bonato P, Camicioli R, Ellis TD, Giladi N, Hamilton JL, et al. Gait impairments in Parkinson's disease. *Lancet Neurol.* (2019) 18:697–708. doi: 10.1016/S1474-4422(19)30044-4

20. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst.* (2017) 30:4766–75.

21. Ribeiro MT, Singh S, Guestrin C. "Why Should I Trust You?". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* (2016). p. 1135–44.

22. Goetz CG, Poewe W, Rascol O, Sampaio C, Stebbins GT, Counsell C, et al. Movement disorder society task force report on the Hoehn and Yahr staging scale: status and recommendations. *Mov Disord.* (2004) 19:1020–8. doi: 10.1002/mds.20213

23. Mahoney FI, Barthel DW. Functional evaluation: the barthel index. *Md State Med J.* (1965) 14:61–5. doi: 10.1037/t02366-000

24. Rabel C, Le Goff F, Lefaucheur R, Ozel G, Fetter D, Rouille A, et al. Subjective perceived motor improvement after acute levodopa challenge in Parkinson's disease. *J Parkinson's Disease.* (2016) 6:779–85. doi: 10.3233/JPD-160906

25. Jankovic J. Motor fluctuations and dyskinesias in Parkinson's disease: clinical manifestations. *Mov Disord.* (2005) 20 Suppl 11:S11–6. doi: 10.1002/mds.20458

26. Postuma RB, Berg D, Stern M, Poewe W, Olanow CW, Oertel W, et al. MDS clinical diagnostic criteria for Parkinson's disease. *Mov Disord.* (2015) 30:1591–601. doi: 10.1002/mds.26424

27. Dorri F, Azmi P, Dorri F. Missing value imputation in DNA microarrays based on conjugate gradient method. *Comput Biol Med.* (2012) 42:222–7. doi: 10.1016/j.compbiomed.2011.11.011

28. Sanz H, Valim C, Vegas E, Oller JM, Reverter F, SVM-RFE. selection and visualization of the most relevant features through non-linear kernels. *BMC Bioinf.* (2018) 19:432. doi: 10.1186/s12859-018-2451-4

29. Chou KL, Stacy M, Simuni T, Miyasaki J, Oertel WH, Sethi K, et al. The spectrum of "off" in Parkinson's disease: What have we learned over 40 years? *Parkinsonism Relat Disord.* (2018) 51:9–16. doi: 10.1016/j.parkreldis.2018.02.001

30. Chapuis S, Ouchchane L, Metz O, Gerbaud L, Durif F. Impact of the motor complications of Parkinson's disease on the quality of life. *Mov Disord.* (2005) 20:224–30. doi: 10.1002/mds.20279

31. van der Velden RMJ, Broen MPG, Kuijf ML, Leentjens AFG. Frequency of mood and anxiety fluctuations in Parkinson's disease patients with motor fluctuations: a systematic review. *Mov Disord.* (2018) 33:1521–7. doi: 10.1002/mds.27465

32. Rizzo G, Copetti M, Arcuti S, Martino D, Fontana A, Logroscino G. Accuracy of clinical diagnosis of Parkinson disease: a systematic review and meta-analysis. *Neurology.* (2016) 86:566–76. doi: 10.1212/WNL.0000000000002350

33. Dorsey ER, Bloem BR. The Parkinson pandemic-a call to action. *JAMA Neurol.* (2018) 75:9–10. doi: 10.1001/jamaneurol.2017.3299

34. Shokouhi N, Khodakarami H, Fernando C, Osborn S, Horne M. Accuracy of step count estimations in Parkinson's disease can be predicted using ambulatory monitoring. *Front Aging Neurosci.* (2022) 14:904895. doi: 10.3389/fnagi.2022.904895

35. Shah VV, Jagodinsky A, McNames J, Carlson-Kuhta P, Nutt JG, El-Gohary M, et al. Gait and turning characteristics from daily life increase ability to predict future falls in people with Parkinson's disease. *Front Neurol.* (2023) 14:1096401. doi: 10.3389/fneur.2023.1096401

36. Shah VV, McNames J, Mancini M, Carlson-Kuhta P, Nutt JG, El-Gohary M, et al. Digital biomarkers of mobility in Parkinson's disease during daily living. *J Parkinson's Dis.* (2020) 10:1099–111. doi: 10.3233/JPD-201914

37. Netukova S, Horakova L, Szabo Z, Krupicka R. Beyond timing and step counting in 360° turning-in-place assessment: a scoping review. *Biomed Eng.* (2024) 23:208. doi: 10.1186/s12938-024-01208-0

38. Seki M, Takahashi K, Koto A, Mihara B, Morita Y, Isozumi K, et al. Camptocormia in Japanese patients with Parkinson's disease: a multicenter study. *Mov Disord.* (2011) 26:2567–71. doi: 10.1002/mds.23955

39. Margraf NG, Wrede A, Deuschl G, Schulz-Schaeffer WJ. Pathophysiological concepts and treatment of camptocormia. *J Parkinson's Dis.* (2016) 6:485–501. doi: 10.3233/JPD-160836

40. Warmerdam E, Romijnders R, Welzel J, Hansen C, Schmidt G, Maetzler W. Quantification of arm swing during walking in healthy adults and Parkinson's disease patients: wearable sensor-based algorithm development and validation. *Sensors.* (2020) 20:20. doi: 10.3390/s20205963

41. Chen M, Sun Z, Xin T, Chen Y, Su F. An Interpretable deep learning optimized wearable daily detection system for Parkinson's disease. *IEEE Trans Neural Syst Rehab Eng.* (2023) 31:3937–46. doi: 10.1109/TNSRE.2023.3314100

42. Brognara L, Palumbo P, Grimm B, Palmerini L. Assessing gait in Parkinson's disease using wearable motion sensors: a systematic review. *Diseases.* (2019) 7:10018. doi: 10.3390/diseases7010018

43. Zhan A, Mohan S, Tarolli C, Schneider RB, Adams JL, Sharma S, et al. Using smartphones and machine learning to quantify Parkinson disease severity. *JAMA Neurol.* (2018) 75:809. doi: 10.1001/jamaneurol.2018.0809

44. Hssayeni MD, Jimenez-Shahed J, Burack MA, Ghoraani B. Symptom-Based, Dual-Channel LSTM Network for The Estimation of Unified Parkinson's Disease Rating Scale III. In: *2019 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI),* Chicago, IL, USA. (2019), p. 1–4

45. Zia Ur Rehman R, Rochester L, Yarnall AJ, Del Din S. Predicting the Progression of Parkinson's Disease MDS-UPDRS-III Motor Severity Score from Gait Data using Deep Learning. In: *Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual International Conference.* (2021), p. 249–52.

46. Parisi F, Ferrari G, Giuberti M, Contin L, Cimolin V, Azzaro C, et al. Body-sensor-network-based kinematic characterization and comparative outlook of UPDRS scoring in leg agility, sit-to-stand, and gait tasks in Parkinson's disease. *IEEE J Biomed Health Inf.* (2015) 19:1777–93. doi: 10.1109/JBHI.2015.2472640

47. Lonini L, Dai A, Shawen N, Simuni T, Poon C, Shimanovich L, et al. Wearable sensors for Parkinson's disease: which data are worth collecting for training symptom detection models. *NPJ Digital Med.* (2018) 1:64. doi: 10.1038/s41746-018-0071-z

48. Zanardi APJ, da Silva ES, Costa RR, Passos-Monteiro E, Dos Santos IO, Kruel LFM, et al. Gait parameters of Parkinson's disease compared with healthy controls: a systematic review and meta-analysis. *Sci Rep.* (2021) 11:752. doi: 10.1038/s41598-020-80768-2

49. Albán-Cadena AC, Villalba-Meneses F, Pila-Varela KO, Moreno-Calvo A, Villalba-Meneses CP, Almeida-Galárraga DA. Wearable sensors in the diagnosis and study of Parkinson's disease symptoms: a systematic review. *J Med Eng Technol.* (2021) 45:532–45. doi: 10.1080/03091902.2021.1922528

50. Channa A, Popescu N, Ciobanu V. Wearable solutions for patients with parkinson's disease and neurocognitive disorder: a systematic review. *Sensors.* (2020) 20:2713. doi: 10.3390/s20092713

51. Ancona S, Faraci FD, Khatab E, Fiorillo L, Gnarra O, Nef T, et al. Wearables in the home-based assessment of abnormal movements in Parkinson's disease: a systematic review of the literature. *J Neurol.* (2022) 269:100–10. doi: 10.1007/s00415-020-10350-3

52. Juutinen M, Wang C, Zhu J, Haladjian J, Ruokolainen J, Puustinen J, et al. Parkinson's disease detection from 20-step walking tests using inertial sensors of a smartphone: machine learning approach based on an observational case-control study. *PLoS ONE.* (2020) 15:e0236258. doi: 10.1371/journal.pone.0236258

53. Rodriguez-Martin D, Sama A, Perez-Lopez C, Catala A, Moreno Arostegui JM, Cabestany J, et al. Home detection of freezing of gait using support vector machines through a single waist-worn triaxial accelerometer. *PLoS ONE.* (2017) 12:e0171764. doi: 10.1371/journal.pone.0171764

54. Kerr J, Marinac CR, Ellis K, Godbole S, Hipp A, Glanz K, et al. Comparison of accelerometry methods for estimating physical activity. *Med Sci Sports Exerc.* (2017) 49:617–24. doi: 10.1249/MSS.0000000000001124

55. Peraza LR, Kinnunen KM, McNaney R, Craddock IJ, Whone AL, Morgan C, et al. An automatic gait analysis pipeline for wearable sensors: a pilot study in Parkinson's disease. *Sensors.* (2021) 21:286. doi: 10.3390/s21248286

56. Mikos V, Heng CH, Tay A, Yen SC, Chia NSY, Koh KML, et al. A wearable, patient-adaptive freezing of gait detection system for biofeedback cueing in Parkinson's disease. *IEEE Trans Biomed Circuits Syst.* (2019) 13:503–15. doi: 10.1109/TBCAS.2019.2914253