


Biomechanical parameter assessment for classification of Parkinson's disease on clinical scale

International Journal of Distributed
Sensor Networks
2017, Vol. 13(5)
© The Author(s) 2017
DOI: 10.1177/1550147717707417
journals.sagepub.com/home/ijdsn


Abdul Haleem Butt¹, Erika Rovini¹, Dario Esposito¹, Giuseppe Rossi²,
Carlo Maremmani³ and Filippo Cavallo¹

Abstract

The primary goal of this study was to investigate computerized assessment methods to classify motor dysfunctioning of patients with Parkinson's disease on the clinical scale. In this proposed system, machine learning-based computerized assessment methods were introduced to assess the motor performance of patients with Parkinson's disease. Biomechanical parameters were acquired from six exercises through wearable inertial sensors: SensFoot V2 and SensHand V1. All patients were evaluated via neurologist by means of the clinical scale. The average rating was calculated from all exercise ratings given by clinicians to estimate overall rating for each patient. Patients were divided in two groups: slight–mild patients with Parkinson's disease and moderate–severe patients with Parkinson's disease according to average rating (“0: slight and mild” and “1: moderate and severe”). Feature selection methods were used for the selection of significant features. Selected features were trained in support vector machine, logistic regression, and neural network to classify the two groups of patients. The highest classification accuracy obtained by support vector machine classifier was 79.66%, with 0.8790 area under the curve. A 76.2% classification accuracy was obtained with 0.7832 area under the curve through logistic regression. A 83.10% classification accuracy was obtained by neural network classifier, with 0.889 area under the curve. Strong distinguishability of the models between the two groups directs the high possibility of motor impairment classification through biomechanical parameters in patients with Parkinson's disease based on the clinical scale.

Keywords

Slight–mild Parkinson's disease and moderate–severe Parkinson's disease, wearable inertial sensors, biomechanical parameters, Least Absolute Shrinkage and Selection Operator, machine learning, support vector machine, logistic regression, neural network

Date received: 1 July 2016; accepted: 7 April 2017

Academic Editor: Filippo Cavallo

Introduction

The information and communication technology (ICT) system essential for the operation of a complete and suitable care service for patients with Parkinson's disease (PwPD), in terms of early diagnosis and monitoring, is based on the development of modular and wearable technologies and a cloud infrastructure for data management. In this study, the introduced wearable devices, which can measure the PwPD

¹The BioRobotics Institute, Scuola Superiore Sant'Anna, Pontedera, Italy

²Institute of Clinical Physiology, National Research Council, Pisa, Italy

³Area Health Authority District I, Neurology Operative Unit, Carrara, Italy

Corresponding author:

Abdul Haleem Butt, The BioRobotics Institute, Scuola Superiore Sant'Anna, Pontedera 56025, Italy.

Email: abdulhaleem.butt@santannapisa.it



Creative Commons CC-BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License

(<http://www.creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of the work without

further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<http://www.uk.sagepub.com/aboutus/openaccess.htm>).

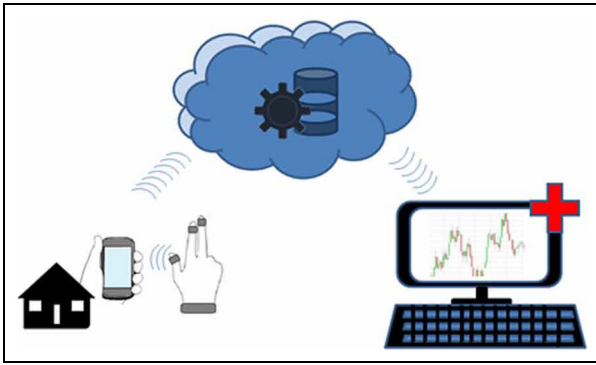


Figure 1. The ICT system.

biomechanical performance finely and objectively, consist of smart inertial units primarily composed of a microcontroller and an inertial measurement unit (IMU). The devices can be used individually or in properly coordinated synchronous networks. In particular, the combination of these units gives origin to modular and wearable sensorized devices adaptable to different parts of the body.

The ICT system shown in Figure 1 would be a future enhancement of this system. The complete system consists of smart devices (e.g. smartphones and smart-TV) which would allow immediate and direct data access at home or on the move, the timely diffusion of useful information (i.e. pharmacological prescriptions and changes in therapy) and healthy guidelines, the counsel of a personal psychological and cognitive diary, and the enjoyment of personalized video-assisted rehabilitative training in a customized care service.¹ The main objective of this healthcare system is to improve the patient's quality of life. At the same time, the system would allow the support of clinical staff for objective assessment and monitoring of a large number of patients. This system would also be helpful to reduce the effect of national healthcare systems. The feasibility of this system has been reported in a previous investigation.²

Current investigation in this article is focused on the machine-learning techniques to classify the PwPD on the clinical scale. Parkinson's disease (PD) is a degenerative disorder of the central nervous system. It is the second most common neurodegenerative disorder after Alzheimer's disease. Most common and early symptoms of PD can include tremors, muscular rigidity, postural instability, bradykinesia, and hypokinesia, caused by a loss of brain dopaminergic neurons. PD affects 1%–2% of people above 50 years of age. According to the PD foundation, 1 million Americans are living with PD, and approximately 60,000 Americans are diagnosed with PD each year. Similarly, 1.2 million Europeans suffer from it, and this number is forecasted to double by 2030.³

Early and accurate diagnosis of PD on the clinical scale is still a challenging task for many reasons. These include the fact that subjects avoid unnecessary examinations; the time, effort, and associated financial costs of therapies; and the possible side effects of treatments. Up-to-date clinical diagnosis is possible when the symptoms are full blown.⁴ Existing methods of PD assessment remain unreliable, which consequently leads to misdiagnosis.⁵ The progressive nature of PD with motor and non-motor problems throughout the disease process complicates the clinical assessment and affects the patient's quality of life and independence. To assess the movement disorder, the neurologist uses visual examination of motor tasks and semi-quantitative rating scales, such as the Hoehn–Yahr Scale and the MDS-UPDRS (Movement Disorder Society–sponsored revision of the Unified Parkinson's Disease Rating Scale).³ These measurements are based on historical progression of the disease and are typically helpful to detect the severity of the disease. These measurements involve repeated clinical visits by the patient. An effective clinical screening process that does not require clinical visits would be beneficial. At the same time, developing a system that could help in diagnosing PD would be useful for clinical professionals.

With PD, the four fundamental motor symptoms consist of tremor, rigidity, bradykinesia, and postural instability. For automatic detection of PwPD symptoms, the common sensors and devices for evaluation are accelerometer, electromyography (EMG), magnetic tracker system, gyroscope, digitizing tablet, video recording, motion detector, and depth sensor.⁶ An accelerometer device is used to measure acceleration forces and capture the movements by converting them into electrical signals that are proportional to the muscular force producing motion. A gyroscope is a sensor device used to measure angular velocity (angular rate); the device senses rotational motion and changes in orientation.⁶ An accelerometer and gyroscope are combined in many motion-sensing instruments. EMG is a technique for evaluating and recording the electrical activity produced by neurologically activated muscles. It records the speed at which nerves can send electrical signals. The accelerometer is the most common sensor used in different studies to assess various symptoms such as tremor, postural instability, bradykinesia, and dyskinesia.⁶ In some studies, motor symptoms of PD are assessed as a single symptom, while in other studies, they are combined with other symptoms. In real time, patients are likely to experience multiple motor symptoms together. It is essential to make a gold standard for clinical ratings to assess the multiple motor symptoms together.

Smart phone,⁷ Microsoft Kinect,⁸ and Leap motion controller⁹ are the latest technologies on the market to assess the PD motor symptoms. These sensors are used

to detect the rigidity and postural instability as a single symptom and also used to detect the bradykinesia and dyskinesia together with the tremor. In addition, wearable sensors such as IMU-based sensors are more feasible for the target population of those suffering from neurological disorders such as PD. Wearable IMU-based sensors can offer low-cost and ubiquitous monitoring solutions for physical activities.¹⁰ Because PD is a progressive disease, it needs continuous monitoring of the symptoms. These IMU-based sensors are available as long as the user is wearing them. These sensors are used to detect the tremor, bradykinesia, and dyskinesia⁹ as single symptoms and used to detect bradykinesia with tremor,¹¹ tremor and postural instability,¹² and tremor with dyskinesia.¹³

Artificial intelligence (AI) techniques such as decision-making, image processing, and classification enable the development of computer systems to perform tasks that typically require human intelligence.⁶ Using meaningful information from the sensor signals, machine learning-based AI techniques show potential for the automatic measurement of Parkinson's symptoms on the clinical scale. The most common techniques such as bagging, boosting, random forest, rotation forest, random subspace, support vector machine (SVM), multilayer perceptron, and decision tree (DT)-based methods are used with minimum-redundancy, maximum-relevance feature-selection algorithms.¹⁴ The wide availability of information on wearable inertial sensors (accelerometers and gyroscopes) in published literature has led to active interest in developing tools for PD assessment on the clinical scale accurately and with less burden on clinicians and patients. All approaches to date have involved classification algorithms that are trained using data from a group of well-characterized patients, and then generalized to individual patients for testing.¹²

In the last decade, machine-learning techniques have been used widely in PD assessment, with the purpose of enhancing the accuracy and effectiveness of PD assessment and minimizing diagnosis error. Many medical decision-making questions can be reduced to binary classification problems, making medical data an ideal domain for several machine-learning techniques.⁵ Machine-learning techniques such as SVM, neural network (NN), and logistic regression (LR) are important to diagnose PD in the early stages. All techniques have the ability not only to diagnose PD in the early stage but also to perform continuous monitoring of PD.¹⁵ SVM is unique among the other machine-learning techniques because of its ability to distribute population in high-dimensional feature space and then categorize the data based on the trained model. Alternatively, LR estimates the probability of PD for a subject based on explanatory variables and is useful to classify the subjects on the clinical scale. SVM and LR show high

performance to classify the PD and normal subjects in the study of subthalamic stimulation in PD on ground reaction force during gait.⁵ In another study by Cancela,¹⁶ which quantifies bradykinesia severity without requiring a standard motor test, five classifiers, SVM, k -nearest neighbors (KNN), NN, and DT, were evaluated. The highest accuracy of 85% was achieved by SVM. In terms of diagnosis, Kupryjanow et al.¹⁷ introduced a new technique to determine the UPDRS sub-score related to motor test (finger tapping and rapid alternating movement of hands). Instead of relying on the subjective assessment from neurologists, four SVM classifiers were trained separately for every hand gesture in another study.¹⁷ Every classifier returns the decision (gesture belongs to the given class or not) and the probability of this decision with high accuracy. Both SVM and LR are increasingly used in neuroimaging studies¹⁵ and also for the classification of PD because they allow characterization at an individual level rather than at the group level, therefore yielding results with a potentially high level of clinical translation.

Classification problems typically involve using high-dimensional features that make the classifier complex and difficult to train. With no feature reduction, both training accuracy and generalization capability will suffer.¹⁸ A straightforward methodology to reduce the complexity of classifiers is to reduce the number of features. Dimensionality reduction is one of the most popular techniques to remove noisy (i.e. irrelevant) and redundant features. Feature selection approaches aim to select a small subset of features that minimizes redundancy and maximizes relevance to the target such as the class labels in classification. Least Absolute Shrinkage and Selection Operator (LASSO) is a powerful feature-selection method proposed by Tibshirani.¹⁹ LASSO penalizes the absolute value of the features' coefficients in a linear regression setting; this leads to some coefficients which are shrunk to zero, which means that features associated with those coefficients are not correlated and are eliminated. Another feature-selection method, Kruskal-Wallis (KW), is a non-parametric one-way analysis of variance (ANOVA) test that has minimal computation, is simple to implement, and is used widely used in clinical datasets.²⁰

To access the feature discriminative ability when differentiating patients on the clinical scale, the linear mixed-effects model (LMM) is a powerful and flexible tool to understand the world. The LMM models are based on a restricted maximum likelihood estimation method and have been used widely in medical diagnostics studies.²¹ To minimize the subjective effect on features, leave-one-subject-out (LOSO) cross-validation is a suitable method for training classifier.²²

In the previous studies, a large number of techniques for automatic detection of PD motor symptoms are

revealed. But most machine-learning algorithms are used to detect single-motor symptoms. PwPD have diverse symptoms. In real time, patients are likely to experience multiple symptoms, thus increasing the chance of false positive and false negative.²² One possible solution to minimize the false-positive and false-negative rate is to detect multiple motor symptoms based on machine-learning algorithms. This study is focused on multiple-motor symptoms to improve accuracy of the machine-learning algorithms.

Materials and methods

In this section, the first part is a description of the instrumentation and protocols. The second part describes the methodology carried out for the classification of PD.

Instrumentation

The system used in this article is composed of two wearable devices that provide an objective and quantitative analysis of the movements of the upper and lower limbs through IMUs which are low cost, low power, non-invasive, small in size, lightweight, wireless, and easy to use.

An IMU integrated in the iNEMO-M1 board based on micro-electro-mechanical systems (MEMS) sensors (three-axis gyroscope L3G4200D and six-axis geomagnetic module LSM303DLHC) and ARM-based 32-bit microcontroller STM32F103RE (STMicroelectronics, Italy) were used to develop the SensFoot V2 device (Figure 2) for lower limb analysis. The system is supplied by a rechargeable LiPo battery and integrated with Bluetooth module (SPBT2632C2A, v3.0, STMicroelectronics) which wirelessly transmits data acquired to a remote PC for offline analysis.²³ The device is placed on the dorsum of the subject's foot with an elastic band to ensure integrity between the foot and sensor.

The SensHand V1 (Figure 2) wearable device for upper limb analysis was developed using the same

inertial sensors integrated into four iNEMO-M1 boards and equipped with dedicated STM32F103RE microcontrollers (ARM 32-bit Cortex™-M3 CPU, STMicroelectronics, Italy). The module placed on the forearm is the coordinator of the system and transmits acquired data toward a generic control station through a wireless communication system based on the ESD 210 (Parani) Bluetooth serial device. The other modules are positioned on the distal phalanges of the thumb, index, and middle fingers. Module coordination and data synchronization are implemented through the CAN-bus standard. A small, rechargeable, and light LiPo battery, integrated into the coordinator module, supplies the system. Both the devices collect data with a sampling frequency of 100 Hz.

Experimental protocol

According to the neurologist and to the tasks required in the motor section of the MDS-UPDRS (MDS-UPDRS III), an experimental protocol composed of six exercises (performed twice, both limbs), such as thumb–forefinger tapping (THFF), hand opening/closing (OPCL), forearm pronation/supination (PSUP), resting tremor (REST), toe tapping with heel pin (TTHP), and heel tapping (HEHE), has been proposed to analyze the motor skills of the upper and lower limbs of the subjects in this study. In addition, every subject attended a short preliminary training to try all the required movements. A neurologist assessed the subjects during the execution of the exercises, assigning them a score according to the tasks in MDS-UPDRS III. MDS-UPDRS III (motor section) tasks are traditionally used for PD assessment and diagnosis (Table 1).

Description of exercises

During the trial session, subjects assumed a comfortable and standardized sitting posture, holding right angles between trunk and thigh (at the hip) and between thigh and shin (at the knee). For each exercise, an initial specific fixed position was established to permit a static acquisition of 3 s to acquire the initial position as reference for each trial. The exercises had to be performed for 10 s, as quickly and widely as possible. The descriptions of the exercises are as follows:

THFF. The subject was directed to keep the hand fixed on the desk, so that the plane between the thumb and the forefinger joined together was parallel to the table. In the starting position, the thumb and the forefinger were in contact, and the subject tapped the forefinger against the thumb (MDS-UPDRS 3.4).

OPCL. The subject was directed to flex the arm that was fixed on the table at the elbow, keeping the



Figure 2. SensHand V1 and SensFoot V2.

Table 1. Biomechanical parameters extracted from SensHand V1 and SensFoot V2.

Exercises	Biomechanical features	Acronym
THFF	Number of taps	H_tapTF
	Opening velocity (deg/s) of the forefinger	H_wotf
	Closing velocity (deg/s) of the forefinger	H_wctf
	Energy expenditure	H_TF_IAV
OPCI	Number of movements	H_numoc
	Opening hand velocity (deg/s)	H_wop
	Closing hand velocity (deg/s)	H_wcl
	Energy expenditure	H_OC_IAV
PSUP	Number of rotation	H_numps
	Amplitude (deg) of the movement	H_excps
	Velocity (deg/s) in supination movements	H_wps
	Velocity (deg/s) in pronation movements	H_wsp
	Energy expenditure	H_PS-IAV
REST	Power in tremor frequency band 3.5–7.5 from accelerometer signal	H_a-pwrpR2
	Power in tremor frequency band 3.5–7.5 from gyroscope signal	H_g-pwrpR2
	Energy expenditure	H_IAV-R
	Number of taps	F_TTHP_taps
TTHP	Toe angle	F_TTHP_T-angle
	Energy expenditure	F_TTHP_IAVxyz
HEHE	Average power of power spectral density	F_HEHE_AVGPWR
	Fundamental frequency	F_HEHE_FREQ
	Maximum peak in signal power	F_HEHE_PEAK
	Energy expenditure	F_HEHE_IAVxyz

The taps represent the number of movements that are performed during an exercise. The velocities are angular velocities (deg/s) measured by the gyroscopes. The amplitude of movement (deg) is calculated by the integration of the angular velocity. The IAV features represent an estimation of the energy expenditure during the execution of the exercises. These features are calculated on the basis of the acceleration values in x , y , and z directions. *IAV* is the integral of magnitude of the total acceleration vector and is calculated as follows

$$IAV = \int \sqrt{a_x^2 + a_y^2 + a_z^2} dt$$

The toe angle in TTHP represents the mean, over the entire exercise, of the maximum amplitude of raising of the toe from the ground at each tap. For REST and HEHE, a frequency analysis was performed. H_a -pwrpR2 and H_g -pwrpR2 represent the ratio between the power of the signal calculated in the frequency band (3.5–7.5 Hz) and the total power of the signal from accelerometer and gyroscope, respectively.

palm of the hand in front of himself or herself. The subject had to alternatively open and close his or her sensorized hand, holding the forearm and the wrist fixed (MDS-UPDRS 3.5).

PSUP. The subject was asked to put the sensorized arm outstretched in front of himself or herself, with the wrist stable and the hand in prone position. The pronation supinations had to be performed in parallel to the floor (MDS-UPDRS 3.6).

REST. The subject was directed to place the sensorized hand on the table in prone position. He remained in rest position for the whole duration of the exercise, keeping the hand fully relaxed (he must not contrast the eventual tremor; MDS-UPDRS 3.17).

TTHP. The subject had to tap his toe on the floor, always keeping the heel in contact with the ground (MDS-UPDRS 3.7).

HEHE. The subject had to tap his heel on the floor, keeping the forefoot always raised from the ground (MDS-UPDRS 3.8).

Participants

A total of 59 PwPD (43 men, 16 women; mean \pm standard deviation: 67.3 \pm 8.8 years old) were involved in this study. All patients were in the on-medication state before and during the experiments. Exclusion criteria were impairments or diseases other than PD (e.g. orthopedic or neurologic) that could affect the performance of daily activities. The exclusion criteria were defined by the neurologist who supported the experimentation. The patients were subjected to neurological examination before involvement in the study to evaluate whether other neurological disorders affected them in addition to PD. The neurologist also asked patients whether they experienced any type of orthopedic impairments (e.g. prosthesis or arthrosis). If such impairments/disorders were revealed, the patients were excluded from the study. All subjects lived independently in the community and gave written informed consent prior to the study. Procedures of the study were approved by the Medical Ethical Committee of

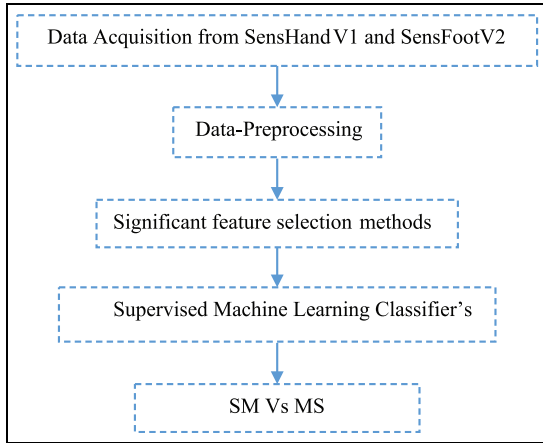


Figure 3. Flowchart purpose classification methodology.

ASL1 (Azienda Sanitaria Locale, Massa and Carrara, Italy; approval no. 1148/12.10.10).

Classification methodology

A flowchart of the methodology carried out in this study is shown in Figure 3. This methodology was established after conducting the literature review, which was discussed in the first section. After data acquisition was accomplished, pre-processing was performed to normalize the values and to remove the high-frequency noise from the dataset. Feature-selection methods were used to obtain the most significant features. At the same time, an LMM was used to determine the random effect on the features. For classification, three different state-of-the-art machine-learning classifiers, that is, SVM, binary LR, and NN, were used. The classification task was composed of three different experiments. In the first experiment, fused significant features were used to train classifiers. In the second composed experiment, SVM was used to measure the contribution of each feature to classify the PwPD. In the third experiment, features were selected based on the measured performance to classify the PwPD (SM and MS).

Data pre-processing. Dedicated algorithms were developed in MATLAB[®] for signal segmentation and event detection to allow the extraction of the features described in the previous paragraph from the motor exercises.

Digital filters, threshold algorithms, and signal integration were applied to conduct the analysis in the spatiotemporal and the frequency domains to obtain the parameters of interest. Specifically, a low-pass fourth-order Butterworth filter with a cut-off frequency of $f_t = 5$ Hz was applied to all the signals (except for REST exercise) to remove the tremor noise. A high-pass fourth-order Butterworth filter ($f_t = 0.5$ Hz) was

implemented for frequency analysis to eliminate high-frequency noise. To classify Parkinsonian patients on the clinical scale, PwPD subjects were divided into two classes. Class 0 belongs to slight and mild (SM) subjects (28 subjects) and Class 1 belongs to moderate and severe (MS) subjects (28 subjects).

Linear scaling was used to scale the features to discrete values between 0 and 1. Linear scaling can be defined as

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (1)$$

where x is the original value of features and x' is the normalized value of features.

Feature selection. LASSO and KW were used to assess the significance of biomechanical features. LASSO is one of the most popular sparse feature-selection methods. It shrinks the regression variables toward zero, keeping all variables in the model to achieve a smooth procedure with less variability. Let us suppose $X = [X_1, \dots, X_N]$, with N representing the number of predictors, and $Y = (y_1, \dots, y_n)$ is the response label.

The LASSO is a penalized least-squares method, imposing a constraint on the L1 norm of the regression coefficients.¹⁸ Thus, the LASSO can be defined as

$$\beta_{Lasso} = \arg \min_{\beta} \|Y - \sum_{i=1}^N X_i \beta_i\|^2 + \lambda \sum_{j=1}^N |\beta_j| \quad (2)$$

where λ is a non-negative tuning parameter controlling the level of sparsity in β . Generically, the usual construct, such as p -values and confidence intervals, does not exist for LASSO estimate. LASSO picks out the top performing features and sparse the other features close to 0.¹⁹ The process of feature selection in LASSO is to select the top performing features; features were ranked using the LASSO fit method (B). A feature's value of 0 or close to 0, according to B , shows high redundancy in the feature subsets. The B -value was used to select the most significant features and to discard the least significant features. Another feature-selection method, KW, was used. This test is based on using the non-parametric ANOVA test to determine the significant features between two or more classes.²⁰ KW can be defined with the given formula

$$H = \frac{12}{N(N+1)} \sum_{i=1}^c \frac{R_i^2}{n_i} - 3(N+1) \quad (3)$$

where c is the number of samples, n_i is the number of observations in the i th sample, N is the sum of all the samples, and R_i is the sum of ranks in the i th sample. For feature selection in KW, the H -value for each feature was compared with the p -value (0.05). If the

H -value was less than the threshold value, then the null-hypothesis could be rejected. KW was also applied on the dataset to obtain significant biomechanical parameters.

LMM. The LMM was performed using SPSS 23 software. This analysis is mostly used to deal with complicated models in which both random and fixed factors are involved. In general, we can define the mixed-effect model as¹¹

$$Y = X\beta + Z\gamma + \varepsilon \quad (4)$$

where Y is $n \times 1$ number of responses, X is an $n \times p$ covariate matrix for fixed-effect β , and Z is an $n \times q$ design matrix of random-effect γ . The $n \times 1$ vector of errors ε is assumed to be multivariate normal with mean 0 and variance matrix random-effect model. To determine the random effect on fixed effect, we add one or more random effects in the mixed model, which gives us the structure of the error term ε .²¹

Classification. The classification task consisted of three separate classification experiments. To train the classifier, LOSO cross-validation was used in SVM and LR. To train the NN classifier, 70% was used for training the model and 15% was used to validate that the network is generalizing and to stop training before overfitting. The last 15% was used as a completely independent test of network generalization. The standard network that is used for function fitting is a two-layer feed-forward network, with a sigmoid transfer function in the hidden layer and a linear transfer function in the output layer. The default number of hidden neurons of 10 was used, which gave the minimum training model error with Neural Network Toolbox in MATLAB 2015b.

To include every feature available might result in overfitting and poor classification performance due to the curse of dimensionality. To overcome this issue, feature-selection methods (KW, Lasso, LMM, and receiver operating characteristic (ROC) curve) were used to select the most significant features. The feature-selection method allows for the selection of the best subset which contains the least number of dimensions that contribute most to accuracy. This is an important stage of pre-processing in machine learning to avoid the curse of dimensionality.

Additionally, ROC provides essential information to measure the individual importance of every input and to discover the variables that produce a statistically significant improvement in the discrimination power of the classification model.²⁴ Another advantage is that it also allows features to be ranked based on area under the curve (AUC).

Classification between SM and MS involved three separate experiments. In the first composed experiment of classification, all the significant features selected from feature-selection methods (KW, Lasso, and LMM) were placed in the classifiers. To measure the classification accuracy of every significant feature, a second experiment was composed in which each significant feature was placed in the SVM classifier with LOSO cross-validation to classify both groups of patients (SM and MS). The area under the ROC curve was used to measure the accuracy of the classification. In the third experiment, features which showed an AUC above the threshold value of 0.5 were fused for classification to improve overall classification accuracy. ROC curve (see Figure 5) is a graphical representation of the fraction of true positives versus the fraction of false positives for a binary classification system. The area under an ROC curve (AUC) is a global measurement of the discrimination performance in a model, and can be used to measure the global accuracy of classification.²⁴ Maximum accuracy corresponds to an AUC value of 1, and minimum discrimination power of 0.5 means a random guess of separation. A minimum discrimination power less than 0.5 is worse than random guessing.

The SVM classifier was implemented using LIBSVM (a Library for Support Vector Machines). The basic idea of SVM in the binary classification task is to separate non-separable input features using high-dimensional separable space based on the selected kernel. A linear kernel function was used for the SM and MS classification. Linear kernel function can be defined as cited by C Hsu et al.²⁵

$$\text{Linear kernel, } K(x_i, x_j) = x_i^T x_j \quad (5)$$

SVMs separate input data into classes using decision boundaries. In the linear kernel function, the mapping rule is linear. For a two-class classification problem, classification decision boundary splits a high-dimensional input space with a hyperplane: all points on one side of the hyperplane are classified as “SM” in our case, while the others are classified as “MS.” It is often used in binary classification problems since it provides representation of the examples as points in space that are mapped, so that the examples of the separate categories are divided by a clear gap that is as wide as possible as shown in Figure 4(a).

LOSO cross-validation strategy was adopted to obtain the unbiased generalization estimates. To avoid biased classification results, samples of one subject were left out at a time, in turn, and a classifier was trained to test the left-out samples. This procedure was repeated for each subject. Average accuracy obtained from paired features is reported in the “Results” section.

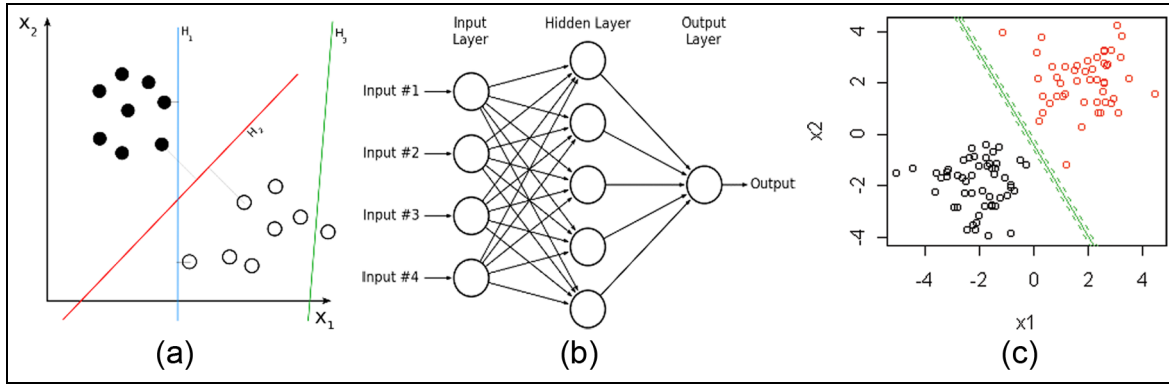


Figure 4. Supervised learning algorithms: (a) support vector machine, (b) neural network, and (c) logistic regression.

Another classifier, the binary LR model, was developed with the Generalized Linear Regression library in MATLAB 2015b. The primary concept of binary LR is to estimate the probability of success in the values given by the explanatory features. The LR classifier is basically used to measure the relationship between dependent variables (target labels) and one or more independent variables (predictors or features) by estimating probabilities using the logistic function or sigmoidal function. It can be defined as²⁶

$$\pi_i = \frac{1}{1 + e^{-\beta x}} \quad (6)$$

In classical statistical techniques, such as linear regression and linear discriminant analysis, this mapping rule is linear, that is, the classification decision boundary or regression curve is a hyperplane (e.g. line or a plane and equivalent geometric object in higher dimensions) shown in Figure 4(c). In contrast, many supervised machine-learning algorithms are in principle capable of discovering any nonlinear relationship, when dealing with biomedical data, so one is often forced to use machine-learning algorithms.²² The classification error was defined as the percentage of patients that were incorrectly classified using leave-one-out cross-validation (LOOCV) in SVM and LR and tenfold cross-validation in NN. The performances of significant features were measured to train the samples of those features across the target value separately in classifiers and confusion matrix, and ROC curves were used to measure the performance of each feature to classify the patients (SM and MS). This method is less computationally expensive and is easy to interpret.²² An NN is a simple artificial neuron with two or more input feature values. Each feature value is multiplied by a weight (w_1, w_2, \dots, w_n), and the result is summed and forms the input to a mathematical function, such as sigmoidal function, which is used to determine the predicted class output of the neuron (0 or 1). A training algorithm

applied to the training data determines the parameters of the algorithm, which are the weight values. Simple artificial neurons such as these are connected, input to output, into networks to make powerful classification algorithms. An example is the “two-layer feed-forward network,” taking four inputs and producing two outputs, which is shown in Figure 4(b).²²

Results

Feature selection

LASSO fit algorithm was used to obtain the significant features. A value of β (mean odd ratio) greater than 1 shows the statistical significance of these features. All other features with β values close to 0 were discarded. This means that LASSO penalized the absolute value of the features and shrunk to 0 because of the huge redundancy in the feature subset values. Another feature-selection method, KW, was used to estimate the significant features. Features with p -values of <0.05 were considered significant. All other features with p -value >0.05 were discarded and were not considered to be significant features. The estimated mean odd ratio and p -value from LASSO and KW for all features are shown in Table 2. Significant features selected from both feature-selection methods are shown in Table 3. It can be observed that five features out of seven are the same features selected from both feature-selection methods. This result endorses the statistical significance of these features for discrimination of PwPD.

LMM

The LMM models were used to assess the discriminative ability of the biomechanical parameters when differentiating between patients on the clinical scale. The LMM models were based on a restricted maximum likelihood estimation method with condition as a fixed effect and subject id as random effect on condition

Table 2. Biomechanical parameter ranking from KW and LASSO.

Features	LASSO (B-value)	Kruskal–Wallis (p-value)
H_tapTF	0.0201	1.45
H_wotf	9.568	0.005213
H_wctf	2.7931	0.002765
H_TF-IAV	0.0054	5.704
H_numoc	0.0206	0.00011
H_wop	5.7019	0.00015
H_wcl	6.6769	2.3019
H_oc-IAV	0.0031	4.03
H_numps	0.0177	0.8119
H_excps	0.0064	6.003
H_wps	0.0015	1.4315
H_wsp	0.0013	1.4568
H_PS-IAV	0.0065	5.532
H_a-pwrpR2	8.2728	0.003814
H_g-pwrpR2	1.662	0.003814
H_IAV-R	0.0024	0.98854
F_HEHE_AVGPWR	0.0074	3.23
F_HEHE_FREQ	0.1296	0.30
F_HEHE_PEAK	0.0033	1.92
F_HEHE_IAVxyz	0.0101	2.33
F_TTHP_taps	0.0188	0.51
F_TTHP_T-angle	0.0043	0.51
F_TTHP_IAVxyz	0.0253	0.38

(target label). Feature is the output. Mixed effect will give p -value for the intercepts b_0 and b_1 , which is for the subjective effect on condition. In this case, our null-hypothesis feature is not affected by the subjective effect, so true groups produce the same feature value. We are looking for p -value >0.05 to reject the null-hypothesis. This means there is insufficient evidence that the feature is affected by random effect. Table 4 shows features that correlated with our hypothesis. The following features were not affected by subjectivity: H_wop ($p = 0.797$), H_wcl ($p = 0.551$), H_excps ($p = 0.067$), H_IAVr ($p = 0.344$), H_gpwrpR2 ($p = 0.133$), F_TTHP-taps ($p = 0.626$), and F_HEHE-fre ($p = 0.946$).

Classification

Classification with significant features. All classification approaches employed the classification error as evaluation criterion. The classification error was defined as the percentage of patients (SM and MS) that were incorrectly classified using LOOCV for SVM and LR. Similarly, in NN classification, error was estimated based on the iterations. To measure the classification performance, the expert diagnosis against the predicted diagnosis from the classifier, all the aware metrics of the diagnostic performance such as sensitivity, specificity, positive and negative predicted values, ROC, and AUC were measured.

Table 3. Selected significant features from KW and LASSO.

Exercise	Features	LASSO (B)	Kruskal–Wallis (p)
THFF	H_wotf	9.568	0.005213
	H_wctf	2.7931	0.002765
OPCL	H_numoc	0.0206	0.00011
	H_wop	5.7019	0.00015
REST	H_wcl	6.6769	2.3019
	H_apwrp-R2	8.2728	0.003814
	H_gpwrp-R2	1.662	0.003814

Table 4. Linear mixed-model estimated p -values for biomechanical parameters.

Exercise	Features	Mixed-effects model (p-value)
THFF	H_tapTF	0.026
	H_woTF	0.000
	H_wcTF	0.000
OPCL	H_TF-IAV	0.000
	H_numOC	0.03
	H_wop	0.797
PSUP	H_wcl	0.551
	H_OC-IAV	0.000
	H_numPS	0.153
	H_excPS	0.067
	H_wps	0.000
	H_wsp	0.000
REST	H_PS-IAV	0.000
	H_a-pwrpR2	0.03
	H_g-pwrpR2	0.133
HEHE	H_IAV-R	0.344
	F_HEHE_AVGPWR	0.000
	F_HEHE_FREQ	0.946
TTHP	F_HEHE_PEAK	0.000
	F_HEHE_IAVxyz	0.000
	F_TTHP_taps	0.626
	F_TTHP_T-angle	0.044
	F_TTHP_IAVxyz	0.000

LOSO cross-validation was used in SVM and LR. The NN 70% dataset was considered as a training set. A level of 15% was used to validate that the network was generalizing and to stop training before overfitting. The last 15% was used as a completely independent test of network generalization with 10 hidden neurons.

In the first classification test, all the significant features were selected based on the feature-selection methods (Lasso, KW, LMM) and entered into the classifiers (SVM, NN, and LR). Table 5 shows a maximum classification accuracy of 76.27% obtained from the SVM classifier with a sensitivity of 92.10% and a specificity of 47.61%. Average AUC from the SVM was 0.9248. The obtained classification accuracy from LR was 66.6%, with a sensitivity of 76.31% and specificity 52.30%. The average AUC from the LR was 0.6078%. NN classifier showed the overall highest classification accuracy of 78.00%, as compared to other classifiers.

Table 5. Classification between MS and SM PwPD with, NN, SVM, and LR with paired significant features.

NN classification			SVM classification			LR classification		
Total instances = 59			Total instances = 59			Total instances = 59		
Correctly classified = 51			Correctly classified = 45			Correctly classified = 40		
Incorrectly classified = 8			Incorrectly classified = 14			Incorrectly classified = 19		
AUC = 0.879			AUC = 0.9248			AUC = 0.6078		
	Class 0	Class 1		Class 0	Class 1		Class 0	Class 1
Class 0	39	4	Class 0	35	3	Class 0	29	9
Class 1	4	12	Class 1	11	10	Class 1	10	11
TPR	89.5%	57.1%	TPR	92.10%	47.61%	TPR	76.31%	52.3%
Accuracy	78.0%		Accuracy	76.27%		Accuracy	66.6%	

MS: moderate and severe; SM: slight and mild; PwPD: patients with Parkinson’s disease; NN: neural network; SVM: support vector machine; LR: logistic regression; AUC: area under the curve; TPR: true positive rate.

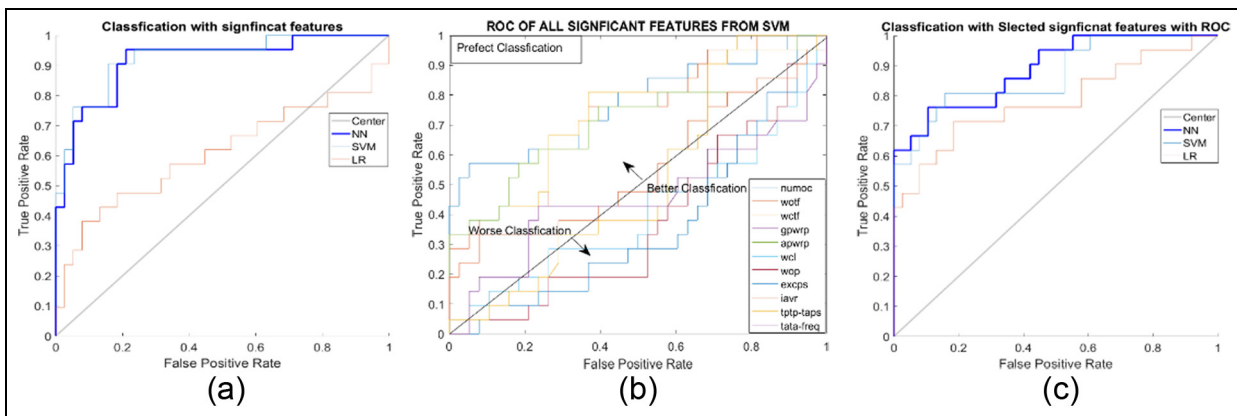


Figure 5. (a) ROC curves of classification test 1, (b) ROC curves of classification test 2, and (c) ROC curves of classification test 3.

Sensitivity was 89.5% and specificity was 57.1% with average AUC 87.90%. ROC curves of all the classifiers are shown in Figure 5.

Features’ classification accuracy measures. In the second composed experiment, classification accuracy of each of the significant features was measured. For this purpose, each of the feature samples was placed in the SVM classifier with linear kernel function in turn. The reason to select the SVM linear kernel is that it allows visualization of the dimensional input space with a hyperplane. In this section, our discussion is focused on the obtained confusion matrix of the significant features. Confusion matrix and ROC curves of the features are shown in Table 6 and Figure 5, respectively.

numoc. The confusion matrix of the hand opening and closing movements showed a classification accuracy of 77.97%, with a sensitivity of 94.73 and specificity of 47.6%. The average AUC was 0.774%, and overall, the feature showed strong potential to classify the PwPD.

Wotf. The confusion matrix for thumb–forefinger tapping opening velocity showed a classification accuracy of 57.63% with a sensitivity of 89.47% and specificity of 0%. The average AUC was 0.7243, and the results revealed that the feature showed limited potential to classify the PwPD.

Wctf. The confusion matrix for thumb–forefinger tapping closing velocity showed a classification accuracy of 57.63% with a sensitivity of 89.47% and specificity of 0%. The average AUC was 0.7243, and overall, the feature showed limited potential to classify the PwPD.

gPwrpR2. The confusion matrix of power in tremor frequency band (3.5–7.5) from accelerometer showed a classification accuracy of 67.80% with a sensitivity of 94.73% and a specificity of 19.07%. The average AUC was 0.7055, and overall, the feature showed limited potential to classify the MS PwPD.

aPwrpR2. The confusion matrix of power in tremor frequency band (3.5–7.5) showed a classification accuracy of 67.80% with a sensitivity of 94.73% and specificity of 19.07%. The average AUC was

Table 6. Classification between MS and MS PwPD for features' classification accuracy measure.

Classification: SVM Feature: numoc AUC = 0.7744			Classification: SVM Feature: wotf AUC = 0.7243			Classification: SVM Feature: wctf AUC = 0.7243			Classification: SVM Feature: gpwrpR2 AUC = 0.7055		
	0	1		0	1		0	1		0	1
Class 0	36	2	Class 0	34	4	Class 0	34	4	Class 0	36	2
Class 1	11	10	Class 1	21	0	Class 1	21	0	Class 1	17	4
TPR	94.73%	47.6%	TPR	89.47%	0%	TPR	89.47%	0%	TPR	94.73%	19.07%
Accuracy	77.97%		Accuracy	57.63%		Accuracy	57.63%		Accuracy	67.80%	
Classification: SVM Feature: apwrpR2 AUC = 0.7055			Classification: SVM Feature: wcl AUC = 0.4085			Classification: SVM Feature: wop AUC = 0.3784			Classification: SVM Feature: excps AUC = 0.3609		
	0	1		0	1		0	1		0	1
Class 0	36	2	Class 0	38	0	Class 0	37	1	Class 0	37	1
Class 1	17	4	Class 1	21	0	Class 1	19	2	Class 1	21	0
TPR	94.73%	19.0%	TPR	100.0%	0%	TPR	97.36%	9.52%	TPR	97.36%	0%
Accuracy	67.80%		Accuracy	64.41%		Accuracy	66.10%		Accuracy	62.71%	
Classification: SVM Feature: IAVR AUC = 0.5539			Classification: SVM Feature: TPTP_Taps AUC = 0.5044			Classification: SVM Feature: TATA_FREQ AUC = 0.4580					
	0	1		0	1		0	1			
Class 0	38	0	Class 0	35	3	Class 0	37	1			
Class 1	21	0	Class 1	21	0	Class 1	19	2			
TPR	100%	0%	TPR	92.10%	0.0%	TPR	97.36%	9.52%			
Accuracy	64.41%		Accuracy	59.32%		Accuracy	66.10%				

MS: moderate and severe; PwPD: patients with Parkinson's disease; SVM: support vector machine; AUC: area under the curve.

0.7055, and overall, the feature showed limited potential to classify the MS PwPD.

wcl. The confusion matrix of hand closing velocity showed a classification accuracy of 64.41% with sensitivity of 100% and specificity of 0%. The average AUC was 0.4085, and overall, the feature failed to classify the MS PwPD.

wop. The confusion matrix of thumb–forefinger tapping closing velocity showed a classification accuracy of 66.10% with sensitivity of 97.36% and specificity of 9.52%. The average AUC was 0.37, and overall, the feature showed limited potential to classify the PwPD.

excPS. The confusion matrix for the amplitude of the movement of pronation and supination showed a classification accuracy of 62.71% with sensitivity of 97.36% and specificity of 0%. The average AUC was 0.3609, and the feature fail to classify the MS PwPD.

iavr. The confusion matrix for rest tremor energy expenditure showed the classification accuracy was 64.14% with sensitivity of 100.0% and specificity of 0%. The average AUC was 0.5539, and the feature fail to classify the MS PwPD.

tthp_taps. The confusion matrix for toe tapping with heel showed the classification accuracy of 59.32%

with sensitivity of 92.10% and specificity of 0%. The average AUC was 0.5044, and the feature fail to classify the MS PwPD.

tata_freq. The confusion matrix of heel tapping frequency showed the classification accuracy of 66.10% with sensitivity of 97.36% and specificity of 9.52%. The average AUC was 0.4580, and very low specificity indicates the limited potential of feature to classify the PwPD in advance stage like MS.

Generally, in all the features, specificity was low as compared to sensitivity. This indicates that the features showed limited potential to classify an advanced stage of PwPD such as MS. Figure 5(b) shows that the AUC in many of the features was significantly lower than 0.5 for the classifier, again suggesting that the classifier as a whole needs to take another fuse of features in to account to classify the PwPD (SM and MS). Only five features showed an AUC above the 0.5.

Classification with selected significant features through AUC. In the third composed experiment, significant features were selected based on the AUC and were paired for the classification. The features which showed an AUC significantly above the threshold value

Table 7. Classification between MS and SM PwPD with NN, SVM, and LR with selected significant features through AUC.

NN classification			SVM classification			LR classification		
Total instances = 59			Total instances = 59			Total instances = 59		
Correctly classified = 49			Correctly classified = 47			Correctly classified = 45		
Incorrectly classified = 10			Incorrectly classified = 12			Incorrectly classified = 14		
AUC = 0.8890			AUC = 0.8709			AUC = 0.7832		
	Class 0	Class 1		Class 0	Class 1		Class 0	Class 1
Class 0	36	8	Class 0	35	3	Class 0	32	6
Class 1	2	13	Class 1	9	12	Class 1	8	13
TPR	94.73%	61.90%	TPR	92.10%	57.14%	TPR	84.21%	61.90%
Accuracy	83.10%		Accuracy	79.66%		Accuracy	76.27%	

MS: moderate and severe; SM: slight and mild; PwPD: patients with Parkinson's disease; NN: neural network; SVM: support vector machine; LR: logistic regression; AUC: area under the curve.

(AUC > 0.5) were fused and entered into the classifiers (NN, SVM, LR). The following features were fused: numoc (number of opening closing hand), wotf (opening velocity of THFF), wctf (closing velocity), gpwrpR2 (power in band (3.5–7.5) from gyroscope), and apwrpR2 (power in band (3.5–7.5) from accelerometer). These fused features were placed in the SVM, LR, and NN classifiers. The classification results are shown in Table 7; the results show an improvement in the classification and endorse the significance of these features to classify the PwPD.

The highest obtained classification accuracy with SVM classifier was 79.66%, with a sensitivity of 92.10% and specificity of 57.14%. The average AUC was 0.8709 as shown in Table 7. Similarly, the highest classification accuracy obtained from LR classifier was 76.7%, with a sensitivity of 84.21% and specificity of 61.90%. The average AUC was 0.7832 as shown in Table 7. NN classifier showed the highest classification accuracy of 83.1% as compared to SVM, and LR with sensitivity of 94.7% and specificity of 61.9%. Average AUC was 0.889 as shown in Table 7. The ROC of each of the classifiers is shown in Figure 5(c). The results endorsed selecting features to train the classifier based on the features' ROC performance to improve the classification accuracy. The maximum classification accuracy is obtained from state-of-the-art classifiers with the following significant features: numoc (number of opening closing hand), wotf (opening velocity of THFF), wctf (closing velocity of THFF), gpwrpR2 (power in band (3.5–7.5) from gyroscope), and apwrpR2 (power in band [3.5–7.5] from accelerometer). Overall, these features showed high potential to classify the advance stage of PwPD.

Discussion

PD assessment on the clinical scale remains a challenging task for clinicians, with visual analysis of motor tasks. In this article, we propose a method for

quantifying PD motor symptoms in both initial and advanced patients experiencing motor fluctuations. The symptoms are quantified by calculating several biomechanical parameters from motor exercises from upper and lower limbs gathered by the sens-Handv1 and sens-Footv2 devices. Based on the biomechanical parameters, three state-of-the-art classifiers, SVM, LR, and NN, were employed to characterize the severity of the motor symptoms and classify them in PD-specific groups. In the first classification experiment, fused significant features selected from feature-selection methods (KW, Lasso, and LMM) were entered in the classification algorithms to obtain the best classification results. However, it is a matter of fact that some features may not monotonically increase or decrease with the UPDRS scale estimated by the clinician. Therefore, they may not represent the actual classification results of medical judgment. However, using all the features may create an overfitting problem due to the curse of dimensionality. In addition to contributing to better diagnosis and monitoring, better characterization might also lead to a better understanding of the disease processes underlying neurodegenerative conditions, which are often poorly understood.²⁷ To obtain the better characteristics for diagnosing the PwPD, further features were selected based on their individual classification performance. The results showed that strategy improved the overall sensitivity and specificity of the classification in all state-of-the-art classifiers' results. The best performing method was NN classifier, which classified SM and MS groups with an accuracy of 83.1%, sensitivity of 94.7%, and specificity of 61.9%, respectively. Overall, the method had good test reliability and provided high discriminating power between both groups. To determine the actual potential of a significant feature to classify the SM and MS patients, SVM was trained from every significant characteristic (feature) of UPDRS separately. LOOCV were used for testing samples (patients) to construct a

hyperplane. The results showed high AUC curve measured from SVM for the following features: number of opening closing hand (numoc), THFF opening velocity (wotf), THFF closing velocity (wctf), and power in the band (3.5–7.5) from gyroscope and accelerometer. In general, the results revealed that all the features have limited potential to classify the PwPD with advanced stage such as MS.

Selected features

The maximum classification accuracy was obtained using five features. Our discussion is focused on these features.

Number of OPCL. The feature showed the highest classification results. Simple, repeated movements such as OPCL make bradykinesia more prominent in PwPD. It is fact that fatigue, hesitation, and freezing in repetitive movements which can be clinically assessed when testing repetitive OPCL. Reduction in the number of the OPCL movements provides essential information to discriminate the PwPD such as SM and MS. Due to the fatigue and hesitation in PwPD with advance stage of the disease, less number of movements should be observed as compared to those PwPD who are in the initial stage of the disease.

THFF opening velocity. The opening velocity of THFF showed a sensitivity, but to classify the advanced stage of PwPD such as MS, it showed limited potential. MS PwPD belongs to a different MDS-UPDRS scale, which complicates to the classification of the PwPD due to high variability in the feature subsets. In general, this feature also showed good a potential to classify the PwPD such as SM. Bradykinesia refers to slowness of movement that is ongoing, akinesia indicates failure of associated movements to occur, and hypokinesia refers to movements that are smaller than desired. These symptoms can be assessed with repetitive movements.²⁸ Repetitive movements in finger tapping result in a progressive reduction in tapping speed and motion amplitude, and increase the use of visual feedback as a compensatory mechanism for a motor system with inherently high variability of motor output.

THFF closing velocity. The closing velocity of THFF also showed a high sensitivity but failed to classify the advanced-stage patients such as MS due to high variability in the feature values. In general, the feature showed a strong contribution in the classification of the PwPD such as SM. Amplitude and speed are the two characteristics mentioned in the MDS-UPDRS that can be more directly related with specific features

from a recorded signal. Bradykinesia is defined as the progressive reduction in speed, amplitude, or both of repetitive actions and is an important diagnostic feature of PD.²⁹

Power in the band (3.5–7.5) from gyroscope and accelerometer. The power from both the gyroscope and accelerometer showed a high sensitivity but low specificity. Overall, the feature had a good potential to classify both groups of PwPD such as SM and MS. One of the most common symptoms of PwPD is REST. Usually, the pathophysiology of rest tremor is largely unknown. It involves unintentional and rhythmic muscle oscillations of an afflicted extremity while the muscles of said extremity are relaxed.²⁷ Rest tremors occur when the body part is in rest position, and it is the most common recognizable symptom of PwPD. Tremor frequency can vary from low (4–5 Hz) to high (8–10 Hz). REST in PwPD is difficult to clinically differentiate from essential tremor. We can differentiate REST objectively because essential tremor has no delay, whereas PD REST re-emerges after a few seconds.

Conclusion

In this article, the objective assessment based on the biomechanical parameters was able to classify the movement disorder in PwPD. The classification results endorse the potential of biomechanical parameters to classify the PwPD on the clinical scale. The mentioned significant features are only valid for this dataset to classify the two groups of patients with the proposed methodology. This is a first step toward the investigation of the gold standard matrix to classify the PwPD on a clinical scale, such as MDS-UPDRS III. To classify the PwPD on a clinical scale, future investigation would be focused on the collection of a big dataset with an equal number of samples from the same MDS-UPDRS III scale. Direct combining of results of suitable biomechanical features based on their performance can improve the accuracy of the machine-learning algorithm. Future investigation also would be focused on the other biomechanical features from SensHand V1 and SensFoot V2, as well as other feature-selection methods to compare the results from different learning algorithms. One limitation of this study was the unavailability of enough samples from each UPDRS scale, which was one reason to distribute samples in the two groups, SM and MS, to maintain the generalization ability of the model. The advanced patients in the MS group affected the TPR due to having the highest variation in the biomechanical parameter-estimated values. Since the number of subjects in the two groups was not equal, this also affected the generalized ability of the model. One possible extension would be to discriminate

the PD subjects on UPDRS with the same proposed methodology. At the same time, the other feature-selection and data-driven machine-learning methods would be investigated to discriminate the PD more efficiently on a clinical scale, especially in the early stages. Due to methodological challenges of assessing the PD in the early stages, the next step is to investigate state-of-the-art data-driven machine-learning methods and new feature extraction from the SensHand V1 and SensFoot V2. Another possible extension would be to accommodate the proposed methodology with other introduced technologies such as leap motion and Kinect sensors to assess the ability of the SensHand V1 device. Leap motion has many limitations. One of them is that leap is very sensitive to motion, which may lead to a great amount of noise. Another one is that there are no guidelines for the users when the hand or finger has crossed the plane. Future leap updates could be helpful in estimating the hand tremor for assessment of PD.³⁰

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was pursued within the project “DAPHNE—innovative and sustainable services for early diagnosis, therapy support and management of Parkinson’s disease by means of mHealth and ICT technologies” supported by a grant from Regione Toscana, Bando FAR-FAS 2014.

References

1. Cavallo F, Esposito D, Rovini E, et al. Preliminary evaluation of SensHand V1 in assessing motor skills performance in Parkinson disease. In: *Proceedings of the 13th international conference on rehabilitation robotics (ICORR)*, Seattle, WA, 24–26 June 2013, pp.1–6. IEEE.
2. Rovini E, Esposito D, Maremmani C, et al. Empowering patients in self-management of Parkinson’s disease through cooperative ICT systems. In: Y Morsi, A Shukla and C Rathore (eds) *Optimizing assistive technologies for aging populations*. Hershey, PA: Medical Information Science Reference, 2016, pp.251–277.
3. Dorsey ER, Constantinescu R, Thompson JP, et al. Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030. *Neurology* 2007; 68: 384–386.
4. Prashanth R, Roy SD and Mandal PK. Automatic classification and prediction models for early Parkinson’s disease diagnosis from SPECT imaging. *Expert Syst Appl* 2014; 41(7): 3333–3342.
5. Mandal I and Sairam N. New machine-learning algorithms for prediction of Parkinson’s disease. *Int J Syst Sci* 2014; 45(3): 647–666.
6. Somayeh A and Westin J. A review of Parkinson’s disease cardinal and dyskinetic motor symptoms assessment methods using sensor systems. In: *3rd EAI international conference on IoT technologies for healthcare*, Västerås, 18–19 October 2016, pp.52–57. Elsevier: IEEE Xplore.
7. Arora S, Venkataraman V, Zhan A, et al. Detecting and monitoring the symptoms of Parkinson’s disease using smartphones: a pilot study. *Parkinsonism Relat D* 2015; 21: 650–653.
8. Yeung L, Cheng KC, Fong CH, et al. Evaluation of the Microsoft Kinect as a clinical assessment tool of body sway. *Gait Posture* 2014; 40(4): 532–538.
9. Hosseinpour N. Implementation of the Kinect sensor and Leap motion for distance study of motor skills in Parkinson’s disease. MSc Thesis, Chalmers University of Technology, Göteborg, 2016.
10. Bennett TR, Wu J, Kehtarnavaz N, et al. Inertial measurement unit-based wearable computers for assisted living applications. *IEEE Signal Proc Mag* 2016; 33(2): 28–35.
11. Patel S, Lorincz K, Hughes R, et al. Monitoring motor fluctuations in patients with Parkinson’s disease using wearable sensors. *IEEE Trans Inf Technol Biomed* 2009; 13: 864–873.
12. Hssayeni M, Burack A, Ghoraani B, et al. Automatic assessment of medication states of patients with Parkinson’s disease using wearable sensors. In: *Proceedings of the 38th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, Orlando, FL, 16–20 August 2016, pp.6082–6085. IEEE.
13. Mellone S, Palmerini L, Cappello A, et al. Hilbert-Huang-based tremor removal to assess postural properties from accelerometers. *IEEE T Bio-Med Eng* 2011; 58: 1752–1761.
14. Tiwari AK. Machine learning based approaches for prediction of Parkinson disease. *Mach Learn Appl* 2016; 3(2): 33–39.
15. Muniz A, Liu H, Lyons K, et al. Comparison among probabilistic neural network, support vector machine and logistic regression for evaluating the effect to subthalamic stimulation in Parkinson disease on ground reaction force during gait a Biomedical Engineering Program. *J Bio-mech* 2010; 43: 720–726.
16. Cancela J, Pansera M, Arredondo MT, et al. A comprehensive motor symptom monitoring and management system: the bradykinesia case. In: *Annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Buenos Aires, Argentina, 31 August–4 September 2010, pp.1008–1011. New York: IEEE.
17. Kupryjanow A, Kunka B and Kostek B. UPDRS tests for diagnosis of Parkinson’s disease employing virtual-touchpad. In: *Proceedings of the workshop on database and expert systems applications (DEXA)*, Bilbao, 30 August–3 September 2010, pp.132–136.
18. Ng WWY, Yeung DS, Firth M, et al. Feature selection using localized generalization error for supervised

- classification problems using RBFNN. *Pattern Recogn* 2008; 41(12): 3706–3719.
19. Tibshirani R. Regression shrinkage and selection via the LASSO. *J R Stat Soc Ser B* 1996; 58: 267–288.
 20. Khan S, Hussain A, Basit A, et al. Kruskal-Wallis-based computationally efficient feature selection for face recognition. *Scientific World J* 2014; 2014: 672630.
 21. Winter B. Linear models and linear mixed effects models in R with linguistic applications, arXiv:1308.5499 [cs], August 2013.
 22. Kubota K, Chen J and Little AM. Machine learning for large-scale wearable sensor data in Parkinson's disease: concepts, promises, pitfalls, and futures. *Mov Disord* 2016; 31: 1314–1326.
 23. Turchetti G, Micera S, Cavallo F, et al. Technology and innovative services. *IEEE Pulse* 2011; 2: 27–35.
 24. Serrano J, Soria E, Mearitn J, et al. Feature selection using roc curves on classification problems. In: *The 2010 international joint conference on neural networks (IJCNN)*, Barcelona, 18–23 July 2010, pp.1–6. New York: IEEE.
 25. Hsu C, Chang C and Lin C. A practical guide to support vector classification. Technical Report, University of Freiburg, Freiburg, 15 July 2003.
 26. Musa A. Comparative study on classification performance between support vector machine and logistic regression. *Int J Mach Learn Cybern* 2013; 4: 13–24.
 27. Lones A, Alty J, Lacy S, et al. Evolving classifiers to inform clinical assessment of Parkinson's disease. In: *IEEE symposium on computational intelligence in health-care and e-health (CICARE)*, Singapore, 15–19 April 2013, pp.76–82. New York: IEEE.
 28. Magrinelli F, Picelli A, Tocco P, et al. Pathophysiology of motor dysfunction in Parkinson's disease as the rationale for drug treatment and rehabilitation. *Parkinsons Dis* 2016; 2016: 9832839.
 29. Marina P, Vincos G, Kern D, et al. Learning more from finger tapping in Parkinson's disease: up and down from dyskinesia to Bradykinesia. *Movement Disord Clin Pract* 2016; 3(2): 184–187.
 30. Khan T, Nyholm D, Westin J, et al. A computer vision framework for finger-tapping evaluation in Parkinson's disease. *Artif Intell Med* 2014; 60: 27–40.