MACHINE LEARNING APPROACHES FOR IDENTIFICATION OF PARKINSON'S DISEASE SEVERITY USING MULTIMODAL FEATURES

Hwayoung Park[1], Changhong Youm[1,2*], Bohyun Kim[1,2], Hyejin Choi[1,2], Juseon Hwang[1,2], and Minsoo Kim[1,2]

[1] Biomechanics laboratory, Dong-A University, Busan, Republic of Korea
[2] Department of Health Sciences, The Graduate School of Dong-A University, Busan, Republic of Korea

This study aimed to create a digital biomarker for assessing the severity of Parkinson's disease (PD) using multimodal features from 50 PD patients and 50 healthy controls. They underwent clinical tests, gait analysis using a motion capture system, postural and functional evaluations, and lifestyle questionnaires. These multimodal features underwent dimensionality reduction techniques such as logistic regression and principal component analysis to identify PD severity using the MDS-UPDRS total score. The results developed six models using machine learning algorithms (Linear Regression and Random Forest), with Model 1 performing the best; spatiotemporal variables from gait analysis were crucial in identifying PD severity. We aim to identify important features correlated with MDS-UPDRS and expect to be applied in clinical settings to monitor the severity of PD.

KEYWORDS: Parkinson's disease, symptom severity, kinematics, multimodal, machine learning

INTRODUCTION: Evaluating the severity of Parkinson's disease (PD) is crucial but challenging due to the subjective nature of current methods reliant on clinical scales such as the Movement Disorder Society-Unified Parkinson's Disease Rating Scale (MDS-UPDRS) (Sotirakis et al., 2023). Limitations include interrater discrepancies hindering disease progression detection and restrictive statistical analyses in clinical research (Almgren et al., 2023). Treatment methods predominantly aim at symptom relief with medications such as levodopa (Albrecht et al., 2022). Leveraging digital biomarkers addresses these limitations by providing personalized therapy and novel intervention targets (Birkenbihl et al., 2023; Deng et al., 2022). The key PD features, such as tremors and gait impairments, can be converted into quantitative surrogates aiding symptom evaluation remotely, minimizing disruptions in daily life (Deng et al., 2022; Sotirakis et al., 2023).

Recent studies explored multimodal datasets encompassing clinical and movement data, voice recordings, magnetic resonance imaging (MRI), and handwriting patterns to identify PD symptom severity across different neurological states (Sotirakis et al., 2023). Among these multimodal features, physical, functional, and kinematic data could be used to identify symptom severity in the early and advanced stages of PD (Deng et al., 2022). Several studies have developed PD severity identification and diagnostic models using machine learning (ML) and deep learning techniques with these digital biomarkers (Albrecht et al., 2022; Almgren et al., 2023; Birkenbihl et al., 2023; Deng et al., 2022; Sotirakis et al., 2023). However, analyzing extensive datasets generated by digital devices necessitates feature-reduction steps to extract clinically meaningful information while avoiding errors (Sotirakis et al., 2023).

In previous research, ML algorithms applied to data from PD clinical questionnaires and wearable devices effectively distinguish between healthy individuals and those with PD or similar disorders. (Almgren et al., 2023; Birkenbihl et al., 2023; Deng et al., 2022; Sotirakis et al., 2023). These algorithms could identify individuals with freezing episodes and falls, as well as detect various symptom signs in individuals with PD. These studies have shown that combining multimodal data collected from individuals with PD and ML algorithms could more effectively track and predict symptom progression than traditional clinical assessment scales (Albrecht et al., 2022; Almgren et al., 2023). However, despite many studies on PD severity identification using digital biomarkers, these have primarily focused on motor assessments such as walking and resting, often with limited datasets.

We previously demonstrated the ability to classify individuals with PD and healthy controls with an accuracy exceeding 90% using ML algorithms based on kinematic features derived from 360° turning gait data collected using a motion capture system (Park et al., 2021). These results for the associations between the clinical and turning characteristics showed that lower turning performance might indicate increased disease severity. Therefore, this study aimed to develop a digital biomarker to identify the severity of PD using multimodal features obtained from PD clinical questionnaires, gait data collected from a motion capture system, postural and physical functional assessments, and lifestyle questionnaires.

METHODS: Fifty individuals with PD and 50 age-matched healthy controls were recruited for a study conducted at Dong-A University Hospital, Busan, Republic of Korea. The study was approved by the experimental protocols by the Institutional Review Board of Dong-A University Medical Center (IRB number: DAUHIRB–22-089). All participants were informed about the study's aims and protocols and signed a written informed consent form to participate in the study. Participants were included if they were diagnosed with PD, received anti-Parkinsonian medication, had no major musculoskeletal problems preventing them from walking or standing, and could walk and stand unassisted during the clinical tests. The control group comprised healthy individuals without cognitive impairment or gait disturbances in the past six months.
All participants were tested using the Mini-Mental State Examination (MMSE) and Montreal Cognitive Assessment (MOCA) when entering the study to ensure that they did not have early dementia (> 24) at the time of providing their consent. All individuals with PD were receiving anti-Parkinsonian medication at the time of their visit to the laboratory for assessment.
The participants were asked to perform four evaluations. 1) Clinical tests: Individuals with PD were evaluated for clinical characteristics using the MDS-UPDRS total score; there are 65 entries on the scale which are divided into four parts (Part I & II: non-motor & motor experiences; Part III: motor examination; Part IV: motor complications); higher scores indicate greater severity in the evaluated aspects, while lower scores indicate less severity, Postural Instability and Gait Disorder (PIGD), modified Hoehn and Yahr scale, New Freezing of Gait Questionnaire (NFOG-Q), Falls Efficacy Scale-Korean (FES-K), Self-Efficacy for Exercise (SEE), and Parkinson's Disease Questionnaire-39 (PDQ-39). 2) Gait analysis: All participants were evaluated for forward and backward straight walking and a 360° turning task, in which they turned left and right three times around a cone at their preferred and faster speed. 3) Postural and functional tests: All participants were evaluated for Short Physical Performance Battery (SPPB), Mini Balance Evaluation Systems Test (Mini-BEST), grip strength, sit-to-stand test five times, and 6-minute walk test. 4) Lifestyle factors: Finally, they were rated using the Nutrition Quotient (NQ) for Korean adults, the 36-item Short Form Health Survey (SF-36), and total Physical Activities (PAs).
For gait analysis, the motion capture system used nine infrared cameras (Vicon MX-T10; Oxford Metrics, UK). An array of 39 reflective markers was used to collect kinematic data, employing a modified version of the Helen Hayes marker set. Markers were placed according to the Plug-in Gait full-body model (Vicon Motion Systems Ltd., Oxford Metrics, UK). The raw data was recorded at a sampling frequency of 100 Hz and was subsequently processed using the Nexus software (version 2.12.0, Vicon, UK) and MATLAB R2018b (MathWorks, Natick, MA). The data were filtered using a fourth-order Butterworth low-pass filter with a 10 Hz cut-off frequency determined through frequency analysis.
Therefore, all multimodal datasets extracted 64 features and were Z-normalized for data preprocessing. Additionally, the left and right limb features were labeled on the more and less affected sides, depending on the individual PD's self-reported side of symptom onset. The more affected side refers to the side that first develops symptoms. The high dimensionality of the kinematic features and collinearities in the dataset necessitate a feature selection step. As the dataset dimensionality reduction methods, we performed logistic regression (LR), principal component analysis (PCA), and multivariate linear regression (MLR). PCA was applied to the entire set of 64 (models 1 and 2) and 34 features (LR) (models 3 and 4), resulting in two sets of the number of principal components that explained two of the eigenvalues of the dataset. In

addition, we estimated PD severity using the MDS-UPDRS total score and MLR based on 34 features derived by LR (models 5 and 6) (Sotirakis et al., 2023).
We employed five-fold cross-validation (CV) to assess the performance of our ML models. The dataset is randomly divided into five subsets. Four subsets (80% of the data) were used as the training set, and the remaining subset (20%) served as the test set. This process was investigated to identify the severity of PD, resulting in six different models using the ML algorithms (Linear Regression and Random Forest). Different feature selection strategies and models were investigated to select the combination that performed best based on the root mean square error (RMSE) derived by five-fold CV analysis (Sotirakis et al., 2023). The RMSE calculation was derived from the predictions generated by our machine learning models compared to the actual values of the MDS-UPDRS total score. Data analysis was performed using SPSS (version 22.0; SPSS Inc., Chicago, IL, USA) and Python 3.8. Figure 1 illustrates the feature selection and model evaluation steps.

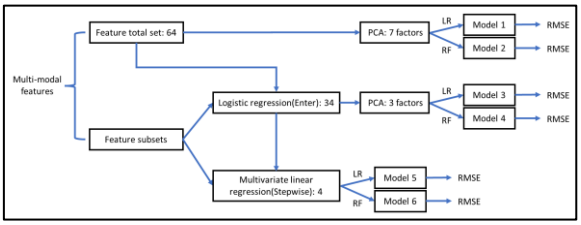

Figure 1: Feature selection and model evaluation pipeline.

RESULTS: Table 1 shows the performance of each model used to estimate the severity of PD, using the MDS-UPDRS total score as the output. The performance metrics were calculated as the average RMSE of the validation set using a 5-fold CV analysis. Model 1 (LR for seven factors using the total feature set) performed the best, with an average RMSE of 15.06. In addition, we have included a comprehensive analysis in Table 2, detailing the feature sets derived from dimensionality reduction methods.

Table 1: Model performance.

Set of features	Dimensionality reduction	Predictor	RMSE (std)
Feature total set (64 features)	**PCA (7 factors)**	**LR (model1)**	**15.06 (3.84)**
		RF (model2)	16.66 (4.42)
Feature subsets (34 features)	PCA (3 factors)	LR (model3)	20.26 (3.79)
		RF (model4)	22.72 (3.62)
	Feature selection (4 features)	**LR (model5)**	**16.01 (3.88)**
		RF (model6)	18.77 (3.42)

Table 2: Multimodal feature set.

Set of features	Features
Feature total set (64 features)	Demographic: gender, age, height, weight, and body mass index; Clinical: symptom & treatment duration, levodopa equivalent dose (LED), MMSE, MOCA, UPDRS total & part Ⅲ, PIGD, H&Y, NFOGQ, FES-K, SEE, and PDQ39; Gait: forward (walking speed, left/right stride length, left/right double support phase) and backward straight walking (walking speed, left/right stride length) and 360° turning task, in which they turned each left and right at prefer and faster speed (walking speed, left/right stride length, left/right double support phase, and left/right contralateral temporal coordination); Postural and functional: SPPB, Mini-BEST, grip strength (kg), sit-to-stand test five times (s), and 6-minute walk (m); Lifestyle: NQ, SF-36(physical, mental, and total score), and PAs.
PCA (7 factors)	PC1: Gait-turning (walking speed and left/right stride length); PC2: Gait-turning (left/right double support phase); PC3: Gait-turning (left/right contralateral temporal coordination); PC4: Clinical (symptom & treatment duration, NFOGQ, LED, and PIGD); PC5: Gait-backward walking (walking speed and left/right stride length); PC6: Clinical (UPDRS total & part Ⅲ, H & Y, and PDQ39); PC7: Gait- forward walking (walking speed, left/right stride length, and left/right double support phase).
Feature subsets (34 features)	Gait: forward (walking speed, left/right stride length, left/right double support phase) and backward straight walking (walking speed, left/right stride length) and 360° turning task, in which they turned each left and right at prefer and faster speed (walking speed, left/right stride length, and left/right double support phase); Postural and functional: SPPB, Mini-BEST, sit-to-stand test five times (s), and 6-minute walk (m); Lifestyle: SF-36(physical) and PAs.
PCA (3 factors)	PC1: Gait-turning (walking speed and left/right stride length); PC2: Gait-turning (left/right double support phase); PC3: Gait- forward walking (walking speed, left/right stride length, and left/right double support phase).
Feature selection (4 features)	Forward and backward walking (right stride length) and turning-right at faster speed (right double support phase and left contralateral temporal coordination).

DISCUSSION: This study demonstrated a quantitative and objective method for identifying the severity of PD using a combination of multimodal data consisting of clinical, physical functional,

and kinematic data and ML algorithms. Dimensionality reduction methods were applied to select various measured multimodal features. Subsequently, we identified a digital biomarker for the severity of PD by verifying the model with the lowest RMSE value (model 1). The results showed the ability to determine the severity of PD based on a set of seven factors obtained through the PCA of the entire set of features. The ability of ML methods to learn patterns from kinematic data and estimate the severity of PD has been demonstrated in previous studies on PD disorders (Sotirakis et al., 2023). They highlighted the significance of individual features, such as the angle of the foot at foot strike and toe-off, as well as stride length, in accurately estimating the MDS UPDRS-III score for identifying the motor symptoms severity of PD. A low foot landing angle has emerged as a distinctive feature and indicator of disease severity in PD (Sotirakis et al., 2023). Our study results in Model 1, including clinical and kinematic variables, performed best; Model 5, also derived from variable selection, showed comparable performance to Model 1. The selected variables primarily consisted of forward/backward straight walking and turning gait variables (Table 2). These results underscore the significance of gait features in assessing PD severity. Specifically, individuals with PD exhibited slower walking speeds and shorter stride lengths during turning, similar to a straight walking trajectory. This characteristic implies an increased risk of falls during turning, suggesting that as the disease progresses, more significant turning impairment may occur owing to increased postural instability (Birkenbihl et al., 2023).

CONCLUSION: The findings of this study, encompassing clinical, physical, functional, and kinematic multimodal features collected through digital devices and analyzed using established ML algorithms, offer valuable insights into assessing PD severity and evaluating treatment efficacy. Furthermore, it may be useful to emphasize the clinical validity and practical application of these findings. We recommend that clinicians prioritize testing parameters, including gait task parameters such as stride length, double support phase, and contralateral temporal coordination, as well as clinical characteristics such as symptom and treatment duration, NFOGQ, LED, and PIGD, which have demonstrated significant potential for distinguishing individuals with PD. By incorporating these multimodal features into clinical practice, healthcare professionals can effectively monitor the severity of PD and tailor treatment strategies to individual patient needs.

REFERENCES

Albrecht, F., Poulakis, K., Freidle, M., Johansson, H., Ekman, U., Volpe, G., ... & Franzén, E. (2022). Unraveling Parkinson's disease heterogeneity using subtypes based on multimodal data. *Parkinsonism & Related Disorders*. 102, 19-29, https://doi.org/10.1016/j.parkreldis.2022.07.014

Almgren, H., Camacho, M., Hanganu, A., Kibreab, M., Camicioli, R., Ismail, Z., ... & Monchi, O. (2023). Machine learning-based prediction of longitudinal cognitive decline in early Parkinson's disease using multimodal features. *Scientific Reports*. 13(1), 13193, https://doi.org/10.1038/s41598-023-37644-6

Birkenbihl, C., Ahmad, A., Massat, N. J., Raschka, T., Avbersek, A., Downey, P., ... & Fröhlich, H. (2023). Artificial intelligence-based clustering and characterization of Parkinson's disease trajectories. *Scientific Reports*. 13(1), 2897, https://doi.org/10.1038/s41598-023-30038-8

Deng, K., Li, Y., Zhang, H., Wang, J., Albin, R. L., & Guan, Y. (2022). Heterogeneous digital biomarker integration out-performs patient self-reports in predicting Parkinson's disease. *Communications biology*, *5*(1), 58, https://doi.org/10.1038/s42003-022-03002-x

Park, H., Shin, S., Youm, C., Cheon, S. M., Lee, M., & Noh, B. (2021). Classification of Parkinson's disease with freezing of gait based on 360 turning analysis using 36 kinematic features. *Journal of NeuroEngineering and Rehabilitation*, *18*, 1-18, https://doi.org/10.1186/s12984-021-00975-4

Sotirakis, C., Su, Z., Brzezicki, M. A., Conway, N., Tarassenko, L., FitzGerald, J. J., & Antoniades, C. A. (2023). Identification of motor progression in Parkinson's disease using wearable sensors and machine learning. *npj Parkinson's Disease*. 9(1), 142, https://doi.org/10.1038/s41531-023-00581-2

ACKNOWLEDGEMENTS: This work was supported by grants from the National Research Foundation of Korea (NRF), funded by the Korea government (MSIT) (No. 2022R1A2C100933711). This research was also supported by the Basic Science Research Program through the NRF, funded by the Ministry of Education (No. 2022R1A6A3A0108756411).