Check for updates

European Journal of Neurology

european journal of neurology
the official journal of the european academy of neurology

ORIGINAL ARTICLE OPEN ACCESS

Machine Learning Predicts Risk of Falls in Parkison's Disease Patients in a Multicenter Observational Study

Maria Chiara Malaguti[1] | Chiara Longo[1] | Monica Moroni[2] | Flavio Ragni[2] | Stefano Bovo[2] | Marco Chierici[2] | Lorenzo Gios[3] | Laura Avanzino[4,5] | Roberta Marchese[5] | Francesca Di Biasio[5] | Matteo Pardini[5,6] | Denise Cerne[6] | Paola Mandich[5,6] | Manuela Marenco[5] | Antonio Uccelli[5] | Bruno Giometto[7] | Giuseppe Jurman[2] | Venet Osmani[2] | the NeuroArtP3 Network†

[1]Azienda Provinciale per i Servizi Sanitari (APSS) di Trento, Trento, Italy | [2]Fondazione Bruno Kessler, Trento, Italy | [3]TrentinoSalute4.0 – Competence Center for Digital Health of the Province of Trento, Trento, Italy | [4]Department of Experimental Medicine, Section of Human Physiology, University of Genoa, Genoa, Italy | [5]IRCCS Ospedale Policlinico San Martino, Genoa, Italy | [6]Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics and Maternal and Child Health, University of Genoa, Genoa, Italy | [7]CISMed, University of Trento, Trento, Italy

Correspondence: Lorenzo Gios (lgios@fbk.eu)

Received: 17 November 2024 | **Revised:** 24 February 2025 | **Accepted:** 27 February 2025

Funding: This work was supported by the Italian Ministry of Health, grant number NET-2018-12366666 NeuroArtP3. This research was partially supported by #NEXTGENERATIONEU (NGEU) and funded by the Ministry of University and Research (MUR), National Recovery and Resilience Plan (NRRP), project MNESYS (PE0000006) – (DN. 1553 11.10.2022).

Keywords: Machine Learning | Multi-Center Validation | Parkinsons Disease

ABSTRACT

Background: Postural instability and gait difficulties are key symptoms of Parkinson's disease (PD), elevating the risk of falls substantially. Falls afflict 35% to 90% of PD patients, representing a major challenge in managing the condition. Accurate prediction of fall risk and identification of contributing factors are essential for timely interventions.

Objectives: Our objective was to develop and validate a machine learning (ML) algorithm across multiple centers in Italy to accurately forecast fall risk and identify related factors using routinely collected clinical data.

Methods: Patient data from two Italian centers ($N=251$) were divided into a training cohort ($N=164$) for ML model development and a validation cohort ($N=87$). External validation was conducted on a subset of PPMI study patients ($N=65$). We compared the performance of logistic regression (LR) and Support Vector Classifier (SVC) models trained on clinical data. The Shapley Additive exPlanations (SHAP) method was employed to examine the predictive power of individual variables.

Results: In the training set, SVC outperformed LR slightly (AUC: LR = 0.779 ± 0.054, SVC = 0.792 ± 0.056). However, LR demonstrated better prediction accuracy in both internal (AUC: LR = 0.753, SVC = 0.733) and external validation cohorts (AUC: LR = 0.714, SVC = 0.676). SHAP analysis on the LR model revealed associations between fall risk and both motor and non-motor variables.

Conclusions: ML-based models effectively estimate fall risk across different clinical centers, enabling tailored interventions to enhance PD patients' quality of life. Challenges persist in predicting falls in US-based patients due to demographic and healthcare system differences.

Maria Chiara Malaguti and Chiara Longo contributed equally to this article.

†NeuroArtP3 Network is details presented in Appendix.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2025 The Author(s). *European Journal of Neurology* published by John Wiley & Sons Ltd on behalf of European Academy of Neurology.

1 | Introduction

Parkinson's disease (PD) is a progressive neurodegenerative disorder characterized by motor impairments and a wide range of non-motor symptoms [1, 2]. Among the various clinical problems associated with PD progression, falls represent a critical and often debilitating issue, resulting in decreased quality of life for patients and increasing the risk of severe injuries with consequent increased healthcare costs [3–6].

Falls affect 50%–60% of PD patients throughout disease progression [2, 7–9], while 40%–70% experiencing falls annually and one-third falling repeatedly [10, 11]. Fall frequency in PD is twice that of age-matched older adults [12], and injurious falls are notably more common [13]. Most falls occur during routine mobility tasks, with over half of patients being repeat fallers [14, 15].

Falls in PD result from a complex interplay of age-related and PD-specific factors [16]. Key contributors include motor symptoms (e.g., axial rigidity, freezing of gait, and postural instability) and non-motor symptoms (e.g., cognitive impairment, depression, and orthostatic hypotension).

Managing falls in PD requires a multidisciplinary approach, where regular assessments of balance, gait, and medication side effects are essential. Personalized interventions, including physical therapy, balance and strength training, cognitive-motor exercises, medication adjustments, and environmental modifications, can help in reducing fall risk [8, 17, 18].

Considering the strong impact of falls on well-being and quality of life of PD patients and socioeconomic cost, developing effective strategies for prediction of falls in PD is a great challenge that will lead to the development of efficacious preventive multidimensional and personalized approaches [19–21].

In this scenario, Machine Learning (ML) can play a significant role in predicting falls in PD patients. ML is a subset of Artificial Intelligence (AI) that focuses on the development of algorithms that allow the learning of data to predict new ones [18]. That is, ML uses input data to create models to predict outcomes.

In PD, an ever-increasing number of articles are studying the disease through ML. A recent systematic review [18] summarized the studies that have applied ML to PD. The authors analyzed 255 studies based on objectives, algorithms used, sample size, source of data, and model performance. Specifically, most studies used voice recording data ($n = 64$), 47 studies used neuroimaging data, 49 used gait data, 25 used EEG data, 13 used genetics data, and only 8 studies used clinical data. Regarding clinical studies, only cross-sectional and classification studies have been performed, and, among these, no studies aimed to predict falls.

To our knowledge, in the literature, only a few studies have investigated the prediction of falls in PD patients using ML. Among the first works on the topic, Gao et al. (2018) [22] used clinical, demographic, and neuroimaging data in different ML algorithms (Logistic Regression, Random Forest, Support Vector Machines, and XGboost) to classify PD patients into fallers or non-fallers. Specifically, using data from 251 patients, they identified that ML algorithms had a classification accuracy of about 70%–80%. Furthermore, the variables that best classify patients who fall or who do not fall are gait speed, Hohen and Yahr (H&Y) stage, postural instability, and gait difficulty-related measurements. A similar study [23] used the XGboost algorithm to classify 305 PDs into fallers or non-fallers using clinical, demographic, pharmacological, balance, and Activities-Specific Balance Confidence Scale (ABC-16) data. The accuracy of the model to classify falls was 72% and the most important features were items 7, 5, 12 in the ABC-16 scale, disease stage, and duration. Another study [24] developed a ML algorithm to classify 109 PDs into fallers and non-fallers using consensus from clinical data from different domains (scale variables, clinical variables, physiological variables, and gait variables). The consensus model reached an accuracy between 75.68% and 77.50% and suggested that different domains contain complementary information and integrating them improves prediction accuracy. Other studies have also used sensor data to predict falls in PD patients. For example, Ullrich et al. [25] used two-week real-word sensor data of 40 PD patients in a Random Forest algorithm. The classification accuracy of the model was 74%.

Previous studies on falls in PD reveal limited analysis using ML and clinical data. Few studies solely utilize clinical data for fall prediction, while non-motor symptoms like orthostatic hypotension or cognitive decline, vital in predicting falls, receive minimal consideration. Classification studies adopt cross-sectional designs, generally lacking the longitudinal aspect crucial for accurate predictions. The present study aims to fill these gaps by developing an AI-driven algorithm using heterogeneous clinical data from PD patients across multiple time points. The objective is to create a predictive tool supporting clinicians' decision-making and tailoring interventions for individual patients [26].

2 | Methods

Detailed description of the project methodological aspects is available in a dedicated paper [26].

2.1 | Ethics Statement and Protocol

The study was approved by the Locals Ethics Committees (Ethics Committee of Azienda Provinciale per i Servizi Sanitari of Trento and Ethics Committee of Liguria Region—Protocol number NET-2018-12366666-3-PD). The protocol for this study was published early in 2024 [26] and refers to the NeuroArtP3 (NET-2018-12366666) project. The NeuroArtP3 project is a four-year multi-center study aimed at promoting the collaboration between clinical and computational centers operating in the field of neurology, co-founded by the Italian Ministry of Health. Written informed consent was obtained from every participant and all assessments were performed according to national and international guidelines.

2.2 | Study Population and Variables of Interest

> A total of 251 patients diagnosed with PD from the NeuroArtP3 project were selected and enrolled in two different clinical centers in North Italy, respectively

14681331, 2025, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ene.70118 by CochraneItalia, Wiley Online Library on [05/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

> 124 at the APSS (Provincial Health Services, Trento, Italy) and 127 at the HSM.
>
> (IRCCS San Martino Hospital, Genoa, Italy)

The entire dataset was divided into training and validation sets. Specifically, APSS patients and a subset of randomly selected HSM patients ($N = 37$) constituted the training cohort ($N = 164$), while the validation cohort comprised the remaining HSM patients (i.e., internal validation cohort; $N = 87$).

Additionally, to further evaluate the performance of the selected prediction models, a third cohort (i.e., external validation cohort; $N = 65$) was selected and comprised a subset of the US-based PD patients enrolled in the PPMI study (www.ppmi-info.org/data; Parkinson Progression Marker Initiative, 2011) (PPMI). The selection of the external validation subjects from the complete PPMI dataset (dataset version: August 2023) was performed considering comparable drug treatments and features consistency.

For each patient, variables collected during the routine clinical practice were investigated, including demographic data, clinical rating scores (e.g., MDS-UPDRS III and Hoehn & Yahr), physical, physiological and cognitive parameters (e.g., presence of ICD, sleep disorders, cognitive status) as well as pharmacological administration. The overall set of variables along with their characteristics is shown in Table 1.

2.3 | Preprocessing and Statistical Analyses

For the analysis, only variables with less than 50% of missing values were considered. Remaining missing values were imputed using the median and the most frequent category, for numerical and categorical features respectively. In order to investigate the between-group differences across the three patient cohorts used to develop and validate the models, statistically significant differences in values distribution were inspected using Kolmogorov–Smirnov and chi-squared tests (Table 2).

Before training the models, numerical predictors values were normalized by scaling them between zero and one, whereas categorical variables were transformed using a one hot encoding strategy.

The outcome variable was encoded by binary labels assigning the positive tag to patients who experienced falls in at least one of the 3 years following the baseline visit, zero otherwise.

2.4 | Machine Learning Model and Validation

The predictive performance of a baseline algorithm, namely Logistic Regression (LR), was compared with that of a Support Vector Classifier (SVC) algorithm.

To prevent overfitting, both models' hyperparameters were optimized on the training set by means of a grid-search, aimed at maximizing the F-1 score metric in a three-fold stratified cross validation setting, repeated two times. The set of parameters defined as a result of this procedure remained fixed for the subsequent evaluation.

Several metrics were considered to evaluate the predictive performance of the models, such as precision, recall, accuracy, balanced accuracy, and the area under the receiver operator characteristic curve (AUC). Moreover, to account for outcome class imbalance, the area under the precision-recall curve (AUPRC), and the Matthews correlation coefficient (MCC) were also computed [27].

2.5 | Predictive Model Interpretability

To inspect the predictive power of individual variables and to increase the interpretability of our machine learning analyses, Shapley Additive exPlanations (SHAP) method was applied to the best performing model. SHAP method deconstructs each prediction into a sum of individual contributions from each variable, emphasizing their influence in the model outcome both at the instance level and throughout the entire population. SHAP values were calculated for each validation cohort and displayed in a Bee swarm plot, to highlight how their values impact the prediction of the model.

2.6 | Data Sharing

The authors takes responsibility for the integrity of the data and the accuracy of the data analysis, while anonymised data could be available upon request.

3 | Results

For this study, patients were recruited from three different cohorts: APSS, HSM, and PPMI. APSS and HSM were used to train ($n = 164$) and test ($n = 87$) a predictive model of PD falls, which was validated also on the PPMI external cohort ($n = 65$).

3.1 | Statistical Differences Between the Cohorts

Before models' training and optimization, dataset characteristics were investigated using univariate statistics.

Overall, as reported in Table 2, there were few statistical differences in the distribution of predictors between the training cohort and the internal validation cohort. Specifically, the only variables that presented strong dissimilarity ($p < 0.001$) between the aforementioned cohorts were years of IMAO at baseline and years of education.

The comparison between the training cohort and the external validation cohort, on the other hand, highlighted a greater difference in variables distribution, revolving around the following features ($p < 0.001$): years of education, LEDD (L-dopa equivalent daily dose), years of Levodopa at baseline, years of dopamine agonists at baseline, years of IMAO at baseline and MDS-UPDRS III (Unified Parkinson's disease rating scale).

14681331, 2025, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ene.70118 by CochraneItalia, Wiley Online Library on [05/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 | Baseline characteristics of the PD patients cohort, divided by outcome and training, internal and external validation cohort.

	Training cohort			Internal validation cohort			External validation cohort		
	Negative	**Positive**	***p***	**Negative**	**Positive**	***p***	**Negative**	**Positive**	***p***
Patients	112	52		61	26		31	34	
Dopamine agonists No (%)	36 (32.1)	18 (34.6)	0.893	14 (23.0)	7 (26.9)	0.902	31 (100.0)	30 (88.2)	0.146
Anxiety No (%)	76 (67.9)	37 (71.2)	0.808	55 (90.2)	20 (76.9)	0.194	27 (87.1)	18 (52.9)	0.007
Falls No (%)	108 (96.4)	34 (65.4)	< 0.001	60 (98.4)	16 (61.5)	< 0.001	28 (90.3)	19 (55.9)	0.005
Depression No (%)	79 (70.5)	31 (59.6)	0.228	48 (78.7)	19 (73.1)	0.771	29 (93.5)	24 (70.6)	0.039
Dysautonomia No (%)	76 (67.9)	26 (50.0)	0.05	44 (72.1)	17 (65.4)	0.09	16 (51.6)	16 (47.1)	0.387
Dysautonomia Orthostatic hypotension (%)	7 (6.2)	8 (15.4)	0.05	0	2 (7.8)	0.09	0	2 (5.9)	0.387
Dysautonomia Altro (%)	29 (25.9)	18 (34.6)	0.05	17 (27.9)	7 (26.9)	0.09	15 (48.4)	16 (47.1)	0.387
Dyskinesia No (%)	94 (83.9)	42 (80.8)	0.781	59 (96.7)	17 (65.4)	< 0.001	31 (100.0)	32 (94.1)	0.514
Sleep–wake disorders No (%)	66 (58.9)	28 (53.8)	0.658	37 (60.7)	14 (53.8)	0.724	22 (71.0)	23 (67.6)	0.983
Phenotype Akinetic (%)	58 (51.8)	32 (61.5)	0.318	13 (21.3)	9 (34.6)	0.3	25 (80.6)	27 (79.4)	1.0
Motor fluctuations No (%)	80 (71.4)	34 (65.4)	0.548	55 (90.2)	17 (65.4)	0.013	31 (100.0)	33 (97.1)	1.0
Freezing No (%)	104 (92.9)	37 (71.2)	0.0	59 (96.7)	17 (65.4)	< 0.001	30 (96.8)	29 (85.3)	0.243
H&Y 1 (%)	34 (30.4)	5 (9.6)	< 0.001	29 (47.5)	10 (38.5)	0.709	3 (9.7)	3 (8.8)	1.0
H&Y 2 (%)	63 (56.2)	26 (50.0)	< 0.001	27 (44.3)	13 (50.0)	0.709	28 (90.3)	31 (91.2)	1.0
H&Y 3 (%)	15 (13.4)	21 (40.4)	< 0.001	5 (8.2)	3 (11.5)	0.709	0	0	1.0
ICD No (%)	101 (90.2)	45 (86.5)	0.67	60 (98.4)	25 (96.2)	1.0	27 (87.1)	26 (76.5)	0.434
IMAO No (%)	56 (50.0)	27 (51.9)	0.951	14 (23.0)	4 (15.4)	0.611	29 (93.5)	31 (91.2)	1.0
Hyposmia No (%)	69 (61.6)	35 (67.3)	0.595	48 (78.7)	22 (84.6)	0.732	15 (48.4)	18 (52.9)	0.906
Levodopa No (%)	34 (30.4)	10 (19.2)	0.191	26 (42.6)	5 (19.2)	0.066	30 (96.8)	32 (94.1)	1.0
Cognitive status Normal (%)	106 (94.6)	41 (78.8)	0.005	60 (98.4)	17 (65.4)	< 0.001	28 (90.3)	32 (94.1)	0.914
Family members with PD No (%)	76 (67.9)	34 (65.4)	0.893	44 (72.1)	19 (73.1)	1.0	21 (67.7)	20 (58.8)	0.626
Sex Male (%)	65 (58.0)	33 (63.5)	0.625	43 (70.5)	11 (42.3)	0.025	21 (67.7)	18 (52.9)	0.335
Years of disease at baseline 0–5 (%)	67 (59.8)	20 (38.5)	< 0.001	45 (73.8)	12 (46.2)	0.006	12 (38.7)	10 (29.4)	0.719

(Continues)

14681331, 2025, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ene.70118 by Cochraneltalia, Wiley Online Library on [05/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 | (Continued)

	Training cohort			Internal validation cohort			External validation cohort		
	Negative	**Positive**	***p***	**Negative**	**Positive**	***p***	**Negative**	**Positive**	***p***
Years of disease at baseline 5–10 (%)	32 (28.6)	13 (25.0)	< 0.001	14 (23.0)	8 (30.8)	0.006	12 (38.7)	13 (38.2)	0.719
Years of disease at baseline > 10 (%)	13 (11.6)	19 (36.5)	< 0.001	2 (3.3)	6 (23.1)	0.006	2 (6.5)	1 (2.9)	0.719
MDS-UPDRS 3 ON, median [Q1, Q3].	18.0 [13.8,25.0]	18.0 [12.0,28.0]	0.701	18.0 [12.0, 26.0]	20.5 [11.8, 27.8]	0.936	22.0 [16.5, 26.5]	24.5 [16.0, 33.8]	0.599
Years of Levodopa at baseline, median [Q1, Q3]	3.0 [1.2, 5.0]	5.5 [3.0, 8.0]	0.017	2.0 [1.0, 4.0]	5.0 [2.0, 9.0]	0.028	2.0 [2.0, 2.0]	0.5 [0.2, 0.8]	0.667
Years of dopamine agonists at baseline, median [Q1, Q3]	3.0 [0.0, 5.0]	4.5 [2.0, 8.0]	0.191	2.0 [1.0, 4.0]	5.0 [1.0, 8.0]	0.013	0	4.0 [1.0, 7.2]	
Years of IMAO at baseline, median [Q1, Q3]	0.0 [0.0, 2.2]	1.0 [0.0, 2.0]	0.815	2.0 [0.0, 4.0]	1.0 [1.0, 5.0]	0.566	2.0 [1.5, 2.5]	8.0 [4.5, 9.5]	0.6
Years between symptoms and diagnosis, median [Q1, Q3]	1.0 [1.0, 3.0]	1.0 [0.0, 2.0]	0.115	1.0 [0.0, 1.0]	1.0 [0.0, 1.0]	1.0	1.0 [0.5, 2.0]	1.0 [0.0, 2.8]	0.934
Age at baseline, median [Q1, Q3]	67.5 [59.8, 73.0]	72.0 [68.0, 78.0]	< 0.001	67.0 [61.0, 72.0]	70.5 [64.8, 74.5]	0.303	63.9 [58.6, 69.8]	67.9 [59.8, 72.1]	0.532
Years of education, median [Q1, Q3]	9.0 [8.0, 13.0]	8.0 [7.2, 13.0]	0.701	13.0 [11.0, 13.0]	10.5 [8.0, 13.0]	0.08	16.0 [15.0, 18.0]	17.0 [14.2, 19.0]	0.715
LEDD, median [Q1, Q3]	300.0 [0.0, 438.0]	399.5 [137.5762.5]	0.008	150.0 [0.0, 300.0]	425.0 [200.0, 600.0]	0.001	0.0 [0.0, 0.0]	0.0 [0.0, 0.0]	1.0

14681331, 2025, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ene.70118 by Cochraneltalia, Wiley Online Library on [05/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 | Between-group differences between the training, and the internal and external validation PD patient cohorts.

	Inter-group differences between training cohort and					
	Internal validation cohort			External validation cohort		
	Statistic	*p*	test	Statistic	*p*	test
MDS-UPDRS 3 ON	0.01	0.99	t-test	**−2.64**	**0.01**	t-test
Years of Levodopa at baseline	8097.00	0.07	Mann–Whitney	**9100.50**	**< 0.001**	Mann–Whitney
Years of dopamine agonists at baseline	6663.50	0.38	Mann–Whitney	**8545.50**	**< 0.001**	Mann–Whitney
Years of IMAO at baseline	**4447.00**	**< 0.001**	Mann–Whitney	**7436.50**	**< 0.001**	Mann–Whitney
Years between symptoms and diagnosis	**8882.50**	**< 0.001**	Mann–Whitney	5657.50	0.45	Mann–Whitney
Age at baseline	0.82	0.41	t-test	**2.46**	**0.01**	t-test
Years of education	**−3.11**	**0.002**	t-test	**1512.50**	**< 0.001**	Mann–Whitney
LEDD	**8359.00**	**0.02**	Mann–Whitney	**9033.50**	**< 0.001**	Mann–Whitney
Dopamine agonists	1.70	0.19	chi2	**66.69**	**< 0.001**	chi2
Anxiety	**8.16**	**0.004**	chi2	0.00	1.00	chi2
Falls	0.00	1.00	chi2	**5.63**	**0.02**	chi2
Depression	2.24	0.13	chi2	**4.07**	**0.04**	chi2
Dysautonomia	4.51	0.11	chi2	**8.60**	**0.01**	chi2
Dyskinesia	0.55	0.46	chi2	**6.83**	**0.01**	chi2
Years of disease at baseline	**8195.50**	**0.03**	Mann–Whitney	**4413.50**	**0.03**	Mann–Whitney
Sleep–wake disorders	0.00	0.95	chi2	2.29	0.13	chi2
Phenotype	**18.96**	**< 0.001**	chi2	**11.43**	**< 0.001**	chi2
Motor fluctuations	**4.53**	**0.03**	chi2	**20.89**	**< 0.001**	chi2
Freezing	0.01	0.91	chi2	0.58	0.45	chi2
H&Y	**14.14**	**< 0.001**	chi2	**28.88**	**< 0.001**	chi2
ICD	**4.71**	**0.03**	chi2	1.68	0.19	chi2
IMAO	**19.94**	**< 0.001**	chi2	**32.76**	**< 0.001**	chi2
Hyposmia	**6.99**	**0.01**	chi2	2.59	0.11	chi2
Levodopa	1.70	0.19	chi2	**85.26**	**< 0.001**	chi2
Cognitive status	0.00	0.95	chi2	0.14	0.71	chi2
Family members with PD	0.53	0.47	chi2	0.18	0.67	chi2
Sex	0.05	0.83	chi2	0.00	1.00	chi2

3.2 | Predictive Performance

The hyperparameters tuning on the training step was finalized to maximize the F1 score in both models, reaching a value of 0.592 ± 0.049 for LR and 0.611 ± 0.095 for SVC. The baseline model had a slightly poorer performance compared to SVC also when considering other metrics, such as AUC (LR: 0.779 ± 0.054, SVC: 0.792 ± 0.056).

Conversely, when tested on the validation sets, never seen during model training, LR achieved a better prediction outcome both within the internal (AUC. LR: 0.753, SVC: 0.733) and the external validation cohort (AUC. LR: 0.714, SVC: 0.676; see Figure 1). Scores for the other considered metrics are reported in Table 3.

3.3 | Model Interpretability

To further explain the predictive performance of the best model (i.e., LR), a saliency analysis using SHAP was performed. Figure 2 depicts the top 9 variables with the highest predictive impact.

The predicted risk of falls within 3 years is associated with both motor and non-motor variables. Regarding the former, the presence of freezing, falls at baseline, low scores on the MDS-UPDRS III test, and high scores (> 2) on the Hoen & Yahr test have a positive influence on the target outcome. For non-motor variables, the presence of dysautonomia, higher age of the patient at baseline visit, a short interval between

14681331, 2025, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ene.70118 by CochraneItalia, Wiley Online Library on [05/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

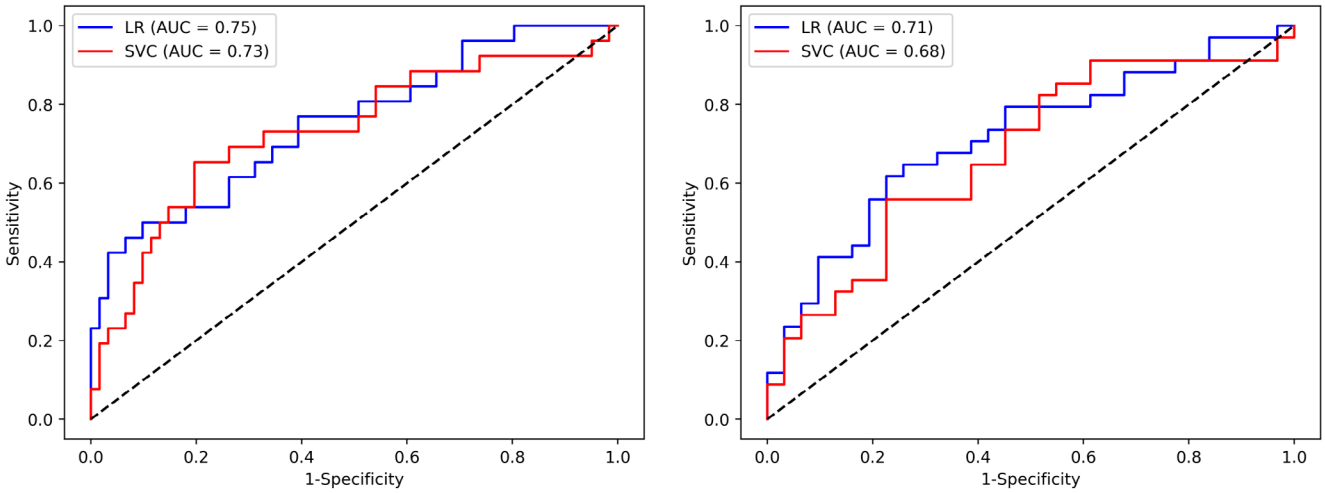

FIGURE 1 | From left to right, AUC performance on the internal and external validation cohort, for the linear regression model (blue line) and support vector classifier (red line).

TABLE 3 | Evaluation metrics for the training, internal and external validation cohort, for both logistic regression and support vector classifier.

	Training cohort		Internal validation cohort		External validation cohort	
	Support vector classifier	Logistic regression	Support vector classifier	Logistic regression	Support vector classifier	Logistic regression
True positive	10.5 (2.665)	10 (2.0)	10	12	10	17
True negative	31.167 (2.483)	31 (3.406)	55	56	27	25
False positive	6.167 (2.639)	6.333 (3.445)	6	5	4	6
False negative	6.833 (2.483)	7.333 (1.966)	16	14	24	17
Correctly classified	41.667 (2.16)	41 (2.191)	65	68	37	42
Incorrectly classified	13 (2.28)	13.667 (2.422)	22	19	28	23
Accuracy	0.762 (0.041)	0.75 (0.043)	0.747	0.782	0.569	0.646
Balanced accuracy	0.72 (0.059)	0.704 (0.036)	0.643	0.69	0.583	0.653
Precision	0.637 (0.082)	0.636 (0.094)	0.625	0.706	0.714	0.739
Recall	0.605 (0.147)	0.577 (0.114)	0.385	0.462	0.294	0.5
F1-score	0.611 (0.095)	0.592 (0.049)	0.476	0.558	0.417	0.596
AUC	0.792 (0.056)	0.779 (0.054)	0.733	0.753	0.676	0.714
Average precision (AUPRC)	0.691 (0.065)	0.688 (0.053)	0.594	0.665	0.709	0.752
MCC	0.449 (0.107)	0.425 (0.068)	0.338	0.438	0.201	0.32

symptoms and diagnosis, the presence of impulse control disorders, and the absence of hyposmia were more closely associated with the presence of falls.

4 | Discussion

Falls in PD lead to significant physical, emotional, and financial burdens, including injuries, reduced mobility, social isolation, and increased medical costs [4, 28]. Early prediction of falls is crucial for proactive risk mitigation, enabling tailored interventions such as medication adjustments, physical therapy, and environmental modifications. These strategies not only reduce fall risk but also improve patients' quality of life. Effective fall prevention requires identifying key risk factors, such as postural impairment, gait instability, and freezing of gait, supported by strong evidence in the scientific literature [29, 30]. Additionally, the presence of cognitive decline, such as dementia, can further

14681331, 2025, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ene.70118 by CochraneItalia, Wiley Online Library on [05/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

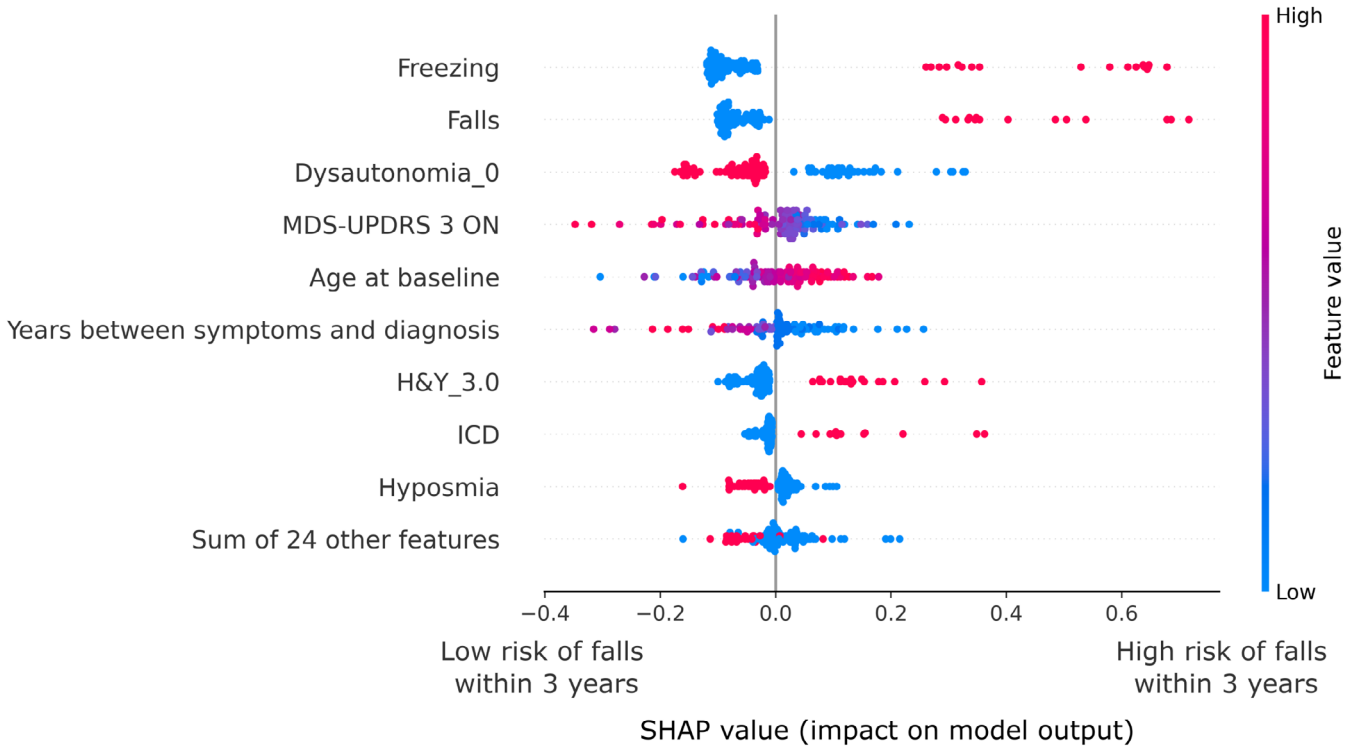

FIGURE 2 | SHAP variable ranking graph, listing the top 9 features impacting model outputs. Blue color represents a low value for a variable, while red the opposite.

exacerbate the risk of falls [31]. Also neurogenic orthostatic hypotension (OH) contributes to increase the risk of falls [30, 32–34]. The pharmacological management of PD can also contribute to falls: some medications, while effective in managing motor symptoms, may lead to orthostatic hypotension, dizziness, and balance issues, increasing the likelihood of falls [28, 35]. Therefore, it emerges that so many different variables can contribute to increased risk of falls. ML allows precisely to integrate different variables by fitting them into a mathematical model to make predictions.

In this study, we developed a reliable algorithm able to predict fall risk and to identify the most important variables in this prediction. The SHAP analysis enabled the identification of variables in our database that increase the likelihood of falling. As anticipated and supported by other studies, the most significant variables are freezing of gait and the presence of falls at baseline, signifying crucial symptoms of disease progression. Additionally, among the most impactful motor variables in predicting falls in the subsequent 3 years, the algorithm emphasizes disease severity, presence of dysautonomia and age [36]. For the latter, there is a growing body of evidence confirming that the higher the age the higher is the risk of falls in both PD patients and general population [36]. Specifically, Logistic Regression identifies patients at a heightened risk of falling with freezing, elevated H&Y scores, and/or delay of diagnosis. Initially, at least, freezing and disease severity are variables that could potentially be modified through external interventions, underscoring the importance of addressing these aspects from both pharmacological and non-pharmacological perspectives, such as physiotherapy. Conversely, the most crucial non-motor variable for the algorithm is dysautonomia, ranking as the third most influential variable in predicting falls, following the presence of falls itself. It is specifically noted that patients without dysautonomia at T0 are less likely to experience falls in the subsequent 3 years. This result is consistent with current evidence in the literature [37]. From a clinical standpoint, this data underscores the importance of both preventing and managing dysautonomia within this patients' group.

However, certain findings from the SHAP analysis pose challenges in interpretation. According to the analysis, a shorter duration between symptom onset and diagnosis correlates with an increased risk of falls, potentially indicating a more rapid and aggressive disease progression [38]. Different hypothesis might partially explain this result. On one side, it is possible that a quicker diagnosis might be linked to more severe or rapidly progressing forms of PD, leading to an increased risk of fall. Alternatively, an early diagnosis may lead to an earlier initiation of treatment, which could involve medications that exacerbate motor symptoms, while patients with early diagnoses might be more aware of their condition, leading to a higher reporting of falls due to heightened monitoring. Further investigation is needed to validate these hypotheses [37].

The role of hyposmia and ICD remains ambiguous; further research is necessary to elucidate their potential impact. Regarding hyposmia, the SHAP analysis suggests a heightened risk of falls in the absence of hyposmia. This contradicts existing evidence that associates hyposmia with disease progression, supported by studies demonstrating a positive correlation between hyposmia severity and disease advancement.

Other studies do not find this correlation. For instance, Boesveldt et al. (2008) examined identification in 404 PD patients using a

14681331, 2025, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ene.70118 by CochraneItalia, Wiley Online Library on [05/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

sophisticated experimental paradigm, discovering that identification remains unaffected by disease severity [39]. However, the dichotomous nature of the hyposmia variable in our dataset (presence/absence) precludes speculation but prompts further investigation through additional studies. For instance, to explore these conflicting findings future research could focus on refining the measurement of hyposmia, to capture varying degrees of severity that might somehow relate to fall risk, while longitudinal studies examining the impact of hyposmia over time on motor and cognitive symptoms could also provide further insight. Additionally, exploring the role of hyposmia in combination with other clinical factors, such as cognitive decline or medication effects, could help clarify its predictive value in fall risk.

Regarding ICD, the SHAP analysis reveals a heightened fall risk associated with its presence. ICDs constitute a category of psychiatric conditions characterized by an inability to resist an urge or temptation to engage in behavior harmful to oneself or others (DSM-5). In PD, these could include gambling, compulsive eating, compulsive buying, compulsive hobbies, and hypersexuality. ICDs can be independent of the disease stage and may also appear early in the disease [40–42]. In fact, it is well-documented in the literature that the presence of ICDs in PD patients correlates with executive dysfunctions such as impulsivity, cognitive inflexibility, and deficits in inhibitory control, with these cognitive abilities involving the prefrontal cortex [43]. Such cognitive dysfunction is also associated with a higher risk of falling [44], hence in our sample, the increased risk of falling in the presence of ICDs could reflect an executive cognitive impairment.

The counterintuitive observation of MDS-UPDRS III (where a lower score corresponds to a higher fall risk) mirrors similar findings observed when evaluating the scale with the PPMI dataset. Additionally, there exists evidence of limited ecological validity [45], making it an unreliable indicator for motor progression, compounded by substantial variance in inter/intra-rater scoring [46]. While widely utilized in clinical settings, UPDRS yields counterintuitive outcomes when scrutinized within AI studies, particularly concerning falls. For instance, a recent study demonstrated significant correlations between motor progression over 1 year and baseline scores, notably Schwab and England ADL progression, and gait parameters, albeit lacking predictive utility [47]. Furthermore, an MDS-commissioned task force assessed clinometric properties of different rating scales, questionaries and timed test that assess disorders of posture, gait and balance (the clinical disorders that are more linked to prediction of falls) [48]. Even if the MDS-UPDRS Postural Instability and Gait Difficulty (PIGD) score was considered "recommended," the authors highlighted that this score lacks sufficient detail for nuanced assessment of gait and balance assessment in PD, particularly with respect to early disease. The limitedness of MDS-UPDRS in assessing the clinical motor aspects mostly predictive of falls in PD may also be an explanation of the poor utility of this scale in the algorithms proposed.

In general, the outcomes of the SHAP analysis validate most established fall risk factors, aligning with existing scientific literature. This concurrence with established findings bolsters the credibility of our results. Nevertheless, the true advantage of AI methodologies lies in their predictive capacity derived from the comprehensive integration of all clinical variables.

From a methodological viewpoint, to improve the models, future research could explore the impact of interaction terms, polynomial features, or domain-specific variables to uncover deeper relationships. Data preprocessing techniques, such as better handling of missing values and normalization, could also enhance performance. Ensemble methods, like Random Forest or Gradient Boosting, may provide a more robust model by combining multiple approaches to reduce overfitting. Further hyperparameter optimization with techniques like Grid Search or Random Search could refine the models. Additionally, more extensive cross-validation methods, such as k-fold or stratified approaches, could improve reliability and generalization.

The distinctiveness of this study lies in its innovative aspects, encompassing both research design and data analysis methodologies. Firstly, a multidisciplinary approach was employed to integrate diverse expertise, spanning IT requirements, clinical aspects, machine learning principles, and privacy and data security compliance. This comprehensive approach is crucial for ensuring a robust triangulation of research questions, adept utilization of advanced AI techniques, and adherence to data protection principles [49, 50].

> A second strand of innovation lies in the multi-site approach adopted. To our knowledge, the present study is among the few studies in this specific field to promote a proper harmonization of clinical variables in the framework of a multi-center study [51, 52].

> A third level of novelty is linked to the use of Parkinson's Progression Markers Initiative (PPMI) as an external evaluation approach. In line with the PPMI principles, a cross-assessment of multiple sources of data (ranging from clinical measures, pharmacological treatment, cognitive assessments) has been used to train the AI algorithms, considering the importance of such variety of information to improve the understanding of diseases progression and related patient's trajectories through this AI models [52, 53]. From this perspective, this study is unique in its reliance on clinical variables to train the AI algorithm, making its outputs broadly applicable to both large research institutions and small clinics, as these variables are commonly available across different settings. Unlike recent studies that use sensor or imaging data to predict falls in PD, this approach emphasizes accessibility and practicality. For instance, wearable sensors have been used to analyze gait variability and postural sway, achieving predictive accuracies of up to 78%, while real-world gait data from foot-worn sensors demonstrated the value of continuous monitoring and highlighting the potential to enhance fall prediction by targeting freezing episodes and gait disturbances [54, 55]. While

14681331, 2025, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ene.70118 by CochraneItalia, Wiley Online Library on [05/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

> these sensor-based methods show promise, they often require specialized equipment or standardized testing environments. In contrast, the clinical variable-based approach in this study is more readily implementable across diverse healthcare settings, offering a practical and scalable solution to guide personalized fall prevention strategies.

In terms of algorithm performance and interpretation, it should be underlined that contextual factors not measured in this project could have an impact on algorithm accuracy. It cannot be excluded that cultural and social variability, as well as socioeconomic and environmental factors could have led to specific biases. Therefore, caution should be taken when generalizing the present results on either datasets or target populations from different countries and contexts [56].

This study highlights the potential of AI to enhance clinical workflows in PD care. By integrating the AI tool into electronic health records or dedicated software, for instance, clinicians could be supported in identifying patients at high risk of falls using key factors available in the clinical records. This could enable P3 approach and related interventions, such as targeted physiotherapy or treatment adjustments, improving patient safety and reducing fall-related burdens, promoting a data-driven decision-support system. Scalable across diverse healthcare settings without the need for specialized tools or infrastructure, the tool could be ideally adopted by a range of institution consider its accessibility and practicality. By advancing a P3 approach, this AI-driven solution aligns with a precision medicine model, supporting better decision-making and improved outcomes for PD patients.

Author Contributions

Maria Chiara Malaguti: conceptualization, investigation, writing – original draft, methodology, validation, writing – review and editing. **Chiara Longo:** conceptualization, investigation, writing – original draft, writing – review and editing, data curation. **Monica Moroni:** writing – original draft, writing – review and editing, formal analysis, data curation. **Flavio Ragni:** writing – original draft, data curation, writing – review and editing, formal analysis. **Stefano Bovo:** formal analysis, data curation, writing – review and editing, writing – original draft. **Marco Chierici:** data curation, formal analysis, writing – review and editing. **Lorenzo Gios:** writing – original draft, writing – review and editing, project administration. **Laura Avanzino:** writing – review and editing. **Roberta Marchese:** writing – review and editing. **Francesca Di Biasio:** writing – review and editing. **Matteo Pardini:** writing – review and editing. **Denise Cerne:** writing – review and editing. **Paola Mandich:** writing – review and editing. **Manuela Marenco:** writing – review and editing, project administration. **Antonio Uccelli:** conceptualization, writing – review and editing, project administration. **Bruno Giometto:** conceptualization, writing – review and editing. **Giuseppe Jurman:** conceptualization, writing – review and editing, supervision, formal analysis. **Venet Osmani:** conceptualization, writing – review and editing, writing – original draft, formal analysis.

Acknowledgments

The authors would like to thank all the patients involved in the NeuroArtP3 NET-2018-12366666 project, as well as the colleagues from the different institutions and departments contributing to this initiative. This research was funded by the Italian Ministry of Health, grant number NET-2018-12366666 NeuroArtP3. This research was partially supported by #NEXTGENERATIONEU (NGEU) and funded by the Ministry of University and Research (MUR), National Recovery and Resilience Plan (NRRP), project MNESYS (PE0000006)—(DN. 1553 11.10.2022). Andrea Falini, Antonella Castellano, Andrea Rossi, Domenico Tortora, Costanza Parodi, Antonio Verrico, Federica Sabatini, Emilio Portaccio, Matteo Betti, Guido Pasquini, Filippo Gerli, Claudia Niccolai, Isabella Cama, Cristina Campi, Alessio Cirone, Sara Garbarino, Michele Piana, Donatella Ottaviani, Raffaella Di Giacopo, Ruggero Bacchin.

Consent

Authors confirm that authorization signed by the patient has been obtained in compliance with the current laws regarding patient authorizations relating to the use or disclosure of protected health information of the jurisdiction(s) to which the patient and the physician are subject. Informed consent has been approved by the local ethics committee. Machine learning predicts risk of falls in Parkison's Disease patients in a multicenter observational study.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

1. R. Balestrino and A. H. V. Schapira, "Parkinson Disease," *European Journal of Neurology* 27, no. 1 (2019): 27–27–42, https://doi.org/10.1111/ene.14108.

2. N. E. Allen, A. K. Schwarzel, and C. G. Canning, "Recurrent Falls in Parkinson's Disease: A Systematic Review," *Parkinson's Disease* 2013 (2013): 2013–906274, https://doi.org/10.1155/2013/906274.

3. S. Ilmaniemi, A. M. Tolppanen, S. K. Herukka, P. Hartikainen, and S. Hartikainen, "Incidence and Outcomes of Head Injuries in People With and Without Parkinson Disease," *European Journal of Neurology* 30, no. 6 (2023): 1648–1657, https://doi.org/10.1111/ene.

4. W. Yang, J. L. Hamilton, C. Kopil, et al., "Current and Projected Future Economic Burden of Parkinson's Disease in the U.S," *NPJ Parkinson's Disease* 6 (2020): 15, https://doi.org/10.1038/s41531-020-0117-1.

5. A. Van Bladel, N. Herssens, K. Bouche, D. Cambier, L. Maes, and N. Lefeber, "Proportion of Falls Reported in Persons With Parkinson's Disease: A Meta-Analysis," *Clinical Rehabilitation* 37, no. 9 (2023): 1260–1277, https://doi.org/10.1177/02692155231158565.

6. S. S. Paul, L. Harvey, C. G. Canning, et al., "Fall-Related Hospitalization in People With Parkinson's Disease," *European Journal of Neurology* 24, no. 3 (2017): 523–529, https://doi.org/10.1111/ene.13238.

7. M. Michałowska, U. Fiszer, A. Krygowska-Wajs, and K. Owczarek, "Falls in Parkinson's Disease. Causes and Impact on Patients' Quality of Life," *Functional Neurology* 20, no. 4 (2005): 163–168.

8. W. Y. Liu, T. H. Tung, C. Zhang, and L. Shi, "Systematic Review for the Prevention and Management of Falls and Fear of Falling in Patients With Parkinson's Disease," *Brain and Behavior: A Cognitive Neuroscience Perspective* 12, no. 8 (2022): e2690, https://doi.org/10.1002/brb3.2690.

9. X. Yang, M. Yao, Z. Guo, X. Shen, and J. Jin, "Development and Validation of Fall Risk Perception Scale for Patients With Parkinson's

14681331, 2025, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ene.70118 by CochraneItalia, Wiley Online Library on [05/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Disease," *Frontiers in Psychology* 15 (2024): 15–1289067, https://doi.org/10.3389/fpsyg.2024.1289067.

10. Y. Balash, C. Peretz, G. Leibovich, T. Herman, J. M. Hausdorff, and N. Giladi, "Falls in Outpatients With Parkinson's Disease: Frequency, Impact and Identifying Factors," *Journal of Neurology* 252, no. 11 (2005): 2005, 13105.

11. B. R. Bloem, V. V. Valkenburg, M. Slabbekoorn, and J. G. van Dijk, "The Multiple Tasks Test. Strategies in Parkinson's Disease," *Experimental Brain Research* 137, no. 3-4 (2001): 478–486, https://doi.org/10.1007/s002210000672.

12. A. L. de Silva Lima, T. Smits, S. K. L. Darweesh, et al., "Home-Based Monitoring of Falls Using Wearable Sensors in Parkinson's Disease," *Movement Disorders: Official Journal of the Movement Disorder Society* 35, no. 1 (2020): 109–115, https://doi.org/10.1002/mds.27830.

13. R. M. Lamont, M. E. Morris, H. B. Menz, J. L. McGinley, and S. G. Brauer, "Falls in People With Parkinson's Disease: A Prospective Comparison of Community and Home-Based Falls," *Gait & Posture* 55 (2017): 62–67, https://doi.org/10.1016/j.gaitpost.2017.04.

14. T. Gazibara, D. K. Tepavcevic, M. Svetel, et al., "Recurrent Falls in Parkinson's Disease After One Year of Follow-Up: A Nested Case-Control Study," *Archives of Gerontology and Geriatrics* 65 (2016): 17–24, https://doi.org/10.1016/j.archger.2016.02.006.

15. L. R. S. Almeida, M. E. P. Piemonte, H. M. Cavalcanti, C. G. Canning, and S. S. Paul, "A Self-Reported Clinical Tool Predicts Falls in People With Parkinson's Disease," *Movement Disorders Clinical Practice* 8, no. 3 (2021): 427–434.

16. A. Fasano, C. G. Canning, J. M. Hausdorff, S. Lord, and L. Rochester, "Falls in Parkinson's Disease: A Complex and Evolving Picture," *Movement Disorders: Official Journal of the Movement Disorder Society* 32, no. 11 (2017): 1524–1536.

17. A. Mirelman, L. Rochester, I. Maidan, et al., "Addition of a Non-Immersive Virtual Reality Component to Treadmill Training to Reduce Fall Risk in Older Adults (V-TIME): A Randomised Controlled Trial," *Lancet (London, England)* 388, no. 10050 (2016): 1170–1182, https://doi.org/10.1016/S0140-6736(16)31325-3.

18. J. Martorell-Marugan, M. Chierici, S. Bandres-Ciga, G. Jurman, and P. Carmona-Sáez, "Machine Learning Applications in the Study of Parkinson's Disease: A Systematic Review," *Current Bioinformatics* 18 (2023): 576–586, https://doi.org/10.2174/15748936186662304.

19. V. V. Shah, A. Jagodinsky, J. McNames, et al., "Gait and Turning Characteristics From Daily Life Increase Ability to Predict Future Falls in People With Parkinson's Disease," *Frontiers in Neurology* 14 (2023): 14–1096401, https://doi.org/10.3389/fneur.2023.1096401.

20. A. Murueta-Goyena, O. Muiño, and J. C. Gómez-Esteban, "Prognostic Factors for Falls in Parkinson's Disease: A Systematic Review," *Acta Neurologica Belgica* 124, no. 2 (2024): 395–406, https://doi.org/10.1007/s13760-023-02428-2.

21. A. S. Monaghan, A. Hooyman, L. E. Dibble, S. H. Mehta, and D. S. Peterson, "Stability Changes in Fall-Prone Individuals With Parkinson Disease Following Reactive Step Training," *Journal of Neurologic Physical Therapy: JNPT* 48, no. 1 (2024): 46–53, https://doi.org/10.1097/NPT.0000000000000442.

22. C. Gao, H. Sun, T. Wang, et al., "Model-Based and Model-Free Machine Learning Techniques for Diagnostic Prediction and Classification of Clinical Outcomes in Parkinson's Disease," *Scientific Reports* 8, no. 1 (2018): 7129, https://doi.org/10.1038/s41598-018-24783-4.

23. P. Panyakaew, N. Pornputtapong, and R. Bhidayasiri, "Using Machine Learning-Based Analytics of Daily Activities to Identify Modifiable Risk Factors for Falling in Parkinson's Disease," *Parkinsonism & Related Disorders* 82 (2021): 77–83.

24. X. Huang, L. Mark, K. Matloob, et al., "A Multi-View Classification Framework for Falls Prediction: Multiple-domain Assessments in Parkinson's Disease," Proceedings of the 54th Hawaii International Conference on System Sciences (2020), http://hdl.handle.net/10125/71029.

25. M. Ullrich, N. Roth, A. Kuderle, et al., "Fall Risk Prediction in Parkinson's Disease Using Real-World Inertial Sensor Gait Data," *IEEE Journal of Biomedical and Health Informatics* 27, no. 1 (2023): 319–328, https://doi.org/10.1109/JBHI.2022.3215921.

26. M. C. Malaguti, L. Gios, B. Giometto, et al., "Artificial Intelligence of Imaging and Clinical Neurological Data for Predictive, Preventive and Personalized (P3) Medicine for Parkinson Disease: The NeuroArtP3 Protocol for a Multi-Center Research Study," *PLoS One* 19, no. 3 (2024): e0300127.

27. D. Chicco, N. Tötsch, and G. Jurman, "The Matthews Correlation Coefficient (MCC) is More Reliable Than Balanced Accuracy, Bookmaker Informedness, and Markedness in Two-Class Confusion Matrix Evaluation," *Biodata Mining* 14, no. 1 (2021): 13.

28. R. Camicioli, M. E. Morris, F. Pieruccini-Faria, et al., "Prevention of Falls in Parkinson's Disease: Guidelines and Gaps," *Movement Disorders Clinical Practice* 10, no. 10 (2023): 1459–1469.

29. K. Marek, S. Chowdhury, A. Siderowf, et al., "Parkinson's Progression Markers Initiative. The Parkinson's Progression Markers Initiative (PPMI) – Establishing a PD Biomarker Cohort," *Annals of Clinical and Translational Neurology* 5, no. 12 (2018): 1460–1477, https://doi.org/10.1002/acn3.644.

30. D. N. Martini, R. Morris, T. M. Madhyastha, et al., "Relationships Between Sensorimotor Inhibition and Mobility in Older Adults With and Without Parkinson's Disease," *Journals of Gerontology. Series A, Biological Sciences and Medical Sciences* 76, no. 4 (2021): 630–637, https://doi.org/10.1093/gerona/glaa300.

31. M. Amboni, P. Barone, and J. M. Hausdorff, "Cognitive Contributions to Gait and Falls: Evidence and Implications," *Movement Disorders: Official Journal of the Movement Disorder Society* 28, no. 11 (2013): 1520–1533, https://doi.org/10.1002/mds.25674.

32. J. L. McKay, K. C. Lang, L. H. Ting, and M. E. Hackney, "Impaired Set Shifting Is Associated With Previous Falls in Individuals With and Without Parkinson's Disease," *Gait & Posture* 62 (2018): 220–226, https://doi.org/10.1016/j.gaitpost.2018.02.027.

33. P. H. S. Pelicioni, J. C. Menant, E. J. Henderson, M. D. Latt, M. A. Brodie, and S. R. Lord, "Mild and Marked Executive Dysfunction and Falls in People With Parkinson's Disease," *Brazilian Journal of Physical Therapy* 25, no. 4 (2021): 437–443, https://doi.org/10.1016/j.bjpt.2020.11.005.

34. S. Heinzel, M. Maechtel, S. E. Hasmann, et al., "Motor Dual-Tasking Deficits Predict Falls in Parkinson's Disease: A Prospective Study," *Parkinsonism & Related Disorders* 26 (2016): 73–77, https://doi.org/10.1016/j.parkreldis.2016.03.007.

35. M. Shadmand, B. Elliott, J. Lautze, and A. Mehdirad, "A Retrospective Analysis of Neurogenic Orthostatic Hypotension in Long-Term Care Facility Residents With Recurrent Falls," *Autonomic Neuroscience: Basic & Clinical* 251 (2024): 103–135, https://doi.org/10.1016/j.autneu.2023.103135.

36. Q. Xu, X. Ou, and J. Li, "The Risk of Falls Among the Aging Population: A Systematic Review and Meta-Analysis," *Frontiers in Public Health* 10 (2022): 902599, https://doi.org/10.3389/fpubh.2022.902599.

37. K. Y. Kwon, S. Park, E. J. Lee, M. Lee, and H. Ju, "Association of Fall Risk Factors and Non-Motor Symptoms in Patients With Early Parkinson's Disease," *Scientific Reports* 11, no. 1 (2021): 5171, https://doi.org/10.1038/s41598-021-84720-w.

38. L. W. Ferguson, M. L. Rajput, N. Muhajarine, S. M. Shah, and A. Rajput, "Clinical Features at First Visit and Rapid Disease Progression in Parkinson's Disease," *Parkinsonism & Related Disorders* 14, no. 5 (2008): 431–435, https://doi.org/10.1016/j.parkreldis.2007.10.012.

14681331, 2025, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ene.70118 by CochraneItalia, Wiley Online Library on [05/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

39. S. Boesveldt, D. Verbaan, D. L. Knol, et al., "A Comparative Study of Odor Identification and Odor Discrimination Deficits in Parkinson's Disease," *Movement Disorders: Official Journal of the Movement Disorder Society* 23, no. 14 (2008): 1984–1990, https://doi.org/10.1002/mds.22155.

40. American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed. (American Psychiatric Association, 2013).

41. G. Santangelo, S. Raimo, M. Cropano, C. Vitale, P. Barone, and L. Trojano, "Neural Bases of Impulse Control Disorders in Parkinson's Disease: A Systematic Review and an ALE Meta-Analysis," *Neuroscience and Biobehavioral Reviews* 107 (2019): 672–685, https://doi.org/10.1016/j.neubiorev.2019.09.041.

42. L. Trojano and C. Papagno, "Cognitive and Behavioral Disorders in Parkinson's Disease: An Update. II: Behavioral Disorders," *Neurological Sciences: Official Journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology* 39, no. 1 (2018): 53–61, https://doi.org/10.1007/s10072-017-3155-7.

43. A. Siquier and P. Andrés, "Cognitive and Behavioral Inhibition Deficits in Parkinson's Disease: The Hayling Test as a Reliable Marker," *Frontiers in Aging Neuroscience* 12 (2020): 621603, https://doi.org/10.3389/fnagi.2020.621603.

44. E. Mak, L. Su, G. B. Williams, et al., "Baseline and Longitudinal Grey Matter Changes in Newly Diagnosed Parkinson's Disease: ICICLE-PD Study," *Brain: A Journal of Neurology* 138, no. Pt 10 (2015): 2974–2986, https://doi.org/10.1093/brain/awv211.

45. M. E. Jenkins, A. M. Johnson, J. D. Holmes, F. F. Stephenson, and S. J. Spaulding, "Predictive Validity of the UPDRS Postural Stability Score and the Functional Reach Test, When Compared With Ecologically Valid Reaching Tasks," *Parkinsonism & Related Disorders* 16, no. 6 (2010): 409–411, https://doi.org/10.1016/j.parkreldis.2010.04.002.

46. L. J. W. Evers, J. H. Krijthe, M. J. Meinders, B. R. Bloem, and T. M. Heskes, "Measuring Parkinson's Disease Over Time: The Real-World Within-Subject Reliability of the MDS-UPDRS," *Movement Disorders: Official Journal of the Movement Disorder Society* 34, no. 10 (2019): 1480–1487, https://doi.org/10.1002/mds.27790.

47. A. Helmy, E. Hamid, M. Salama, A. Gaber, M. El-Belkimy, and A. Shalash, "Baseline Predictors of Progression of Parkinson's Disease in a Sample of Egyptian Patients: Clinical and Biochemical," *Egyptian Journal of Neurology, Psychiatry and Neurosurgery* 58, no. 1 (2022): 9, https://doi.org/10.1186/s41983-022-00445-1.

48. B. R. Bloem, J. Marinus, Q. Almeida, et al., "Measurement Instruments to Assess Posture, Gait, and Balance in Parkinson's Disease: Critique and Recommendations," *Movement Disorders: Official Journal of the Movement Disorder Society* 31, no. 9 (2016): 1342–1355, https://doi.org/10.1002/mds.26572.

49. G. Bincoletto, *Data Protection by Design in the E-Health Care Sector: Theoretical and Applied Perspectives* (Nomos Verlagsgesellschaft, 2021).

50. B. Garcia Santa Cruz, A. Husch, and F. Hertel, "Machine Learning Models for Diagnosis and Prognosis of Parkinson's Disease Using Brain Imaging: General Overview, Main Challenges, and Future Directions," *Frontiers in Aging Neuroscience* 15 (2023): 1216163, https://doi.org/10.3389/fnagi.2023.1216163.

51. K. Bounsall, M. Milne-Ives, A. Hall, C. Carroll, and E. Meinert, "Artificial Intelligence Applications for Assessment, Monitoring, and Management of Parkinson Disease Symptoms: Protocol for a Systematic Review," *JMIR Research Protocols* 12 (2023): e46581, https://doi.org/10.2196/46581.

52. R. T. Gerraty, A. Provost, and L. Li, "Machine Learning Within the Parkinson's Progression Markers Initiative: Review of the Current State of Affairs," *Frontiers in Aging Neuroscience* 15 (2023): 1076657, https://doi.org/10.3389/fnagi.2023.1076657.

53. J. Harvey, R. A. Reijnders, R. Cavill, et al., "Machine Learning-Based Prediction of Cognitive Outcomes in de Novo Parkinson's Disease," *NPJ Parkinson's Disease* 8, no. 1 (2022): 150, https://doi.org/10.1038/s41531-022-00409-5.

54. P. Wu, B. Cao, Z. Liang, and M. Wu, "The Advantages of Artificial Intelligence-Based Gait Assessment in Detecting, Predicting, and Managing Parkinson's Disease," *Frontiers in Aging Neuroscience* 15 (2023): 1191378, https://doi.org/10.3389/fnagi.2023.11913.

55. Z. K. Lim, T. Connie, M. K. O. Goh, and N. B. Saedon, "Fall Risk Prediction Using Temporal Gait Features and Machine Learning Approaches," *Frontiers in Artificial Intelligence* 7 (2024): 1425713, https://doi.org/10.3389/frai.2024.1425713.

56. J. Hastings, "Achieving Inclusivity by Design: Social and Contextual Information in Medical Knowledge," *Yearbook of Medical Informatics* 31, no. 1 (2022): 228–235, https://doi.org/10.1055/s-0042-1742509.

Appendix

NeuroArtP3 Network is composed by: Andrea Falini (Neuroradiology Unit and CERMAC, IRCCS Ospedale San Raffaele, Milano; Università Vita-Salute San Raffaele, Milano), Antonella Castellano (Neuroradiology Unit and CERMAC, IRCCS Ospedale San Raffaele, Milano; Università Vita-Salute San Raffaele, Milano), Andrea Rossi (Neuroradiology Unit, IRCCS Istituto Giannina Gaslini, Genoa, Italy; Department of Health Sciences (DISSAL), University of Genoa, Genoa, Italy), Domenico Tortora (Neuroradiology Unit, IRCCS Istituto Giannina Gaslini, Genoa, Italy), Costanza Parodi (Neuroradiology Unit, IRCCS Istituto Giannina Gaslini, Genoa, Italy), Antonio Verrico (Neuro-Oncology Unit, IRCCS Istituto Giannina Gaslini, Genoa, Italy), Federica Sabatini (UOSD Cell Factory, IRCCS Istituto Giannina Gaslini, Genoa, Italy), Emilio Portaccio (Department of NEUROFARBA, University of Florence, Florence, Italy), Matteo Betti (Department of NEUROFARBA, University of Florence, Florence, Italy), Guido Pasquini (IRCCS Don Carlo Gnocchi Foundation, Florence, Italy), Filippo Gerli (IRCCS Don Carlo Gnocchi Foundation, Florence, Italy), Claudia Niccolai (IRCCS Don Carlo Gnocchi Foundation, Florence, Italy), Isabella Cama (Life Science Computational Laboratory (LISCOMP), IRCCS Ospedale Policlinico San Martino, Genova, Italy; Dipartimento di Matematica (MIDA), Università di Genova, 16132 Genova, Italy), Cristina Campi (Life Science Computational Laboratory (LISCOMP), IRCCS Ospedale Policlinico San Martino, Genova, Italy; Dipartimento di Matematica (MIDA), Università di Genova, 16132 Genova, Italy), Alessio Cirone (IRCCS Ospedale Policlinico San Martino, Genoa, Italy; Life Science Computational Laboratory (LISCOMP), IRCCS Ospedale Policlinico San Martino, Genova, Italy), Sara Garbarino (IRCCS Ospedale Policlinico San Martino, Genoa, Italy; Life Science Computational Laboratory (LISCOMP), IRCCS Ospedale Policlinico San Martino, Genova, Italy), Michele Piana (Life Science Computational Laboratory (LISCOMP), IRCCS Ospedale Policlinico San Martino, Genova, Italy; Dipartimento di Matematica (MIDA), Università di Genova, 16132 Genova, Italy), Donatella Ottaviani (Azienda Provinciale per i Servizi Sanitari (APSS) di Trento, Trento, Italy), Raffaella Di Giacopo (Azienda Provinciale per i Servizi Sanitari (APSS) di Trento, Trento, Italy), Ruggero Bacchin (Azienda Provinciale per i Servizi Sanitari (APSS) di Trento, Trento, Italy).

14681331, 2025, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/ene.70118 by CochraneItalia, Wiley Online Library on [05/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License