The Role of Machine Learning in Cognitive Impairment in Parkinson Disease: Systematic Review and Meta-Analysis

Yanyun Wu, MD; Yangfan Cheng, MD; Yi Xiao, MD; Huifang Shang, MD; Ruwei Ou, MD

Department of Neurology, West China Hospital of Sichuan University, Chengdu, China

Corresponding Author:
Ruwei Ou, MD
Department of Neurology
West China Hospital of Sichuan University
No.37, Guoxue Lane
Chengdu, 610041
China
Phone: 86 18980607525
Fax: 86 028 85423550
Email: ouruwei@aliyun.com

Abstract

Background: Parkinson disease (PD) is a common neurodegenerative disease characterized by both motor and nonmotor symptoms. Cognitive impairment often occurs early in the disease and can persist throughout its progression, severely impacting patients' quality of life. The utilization of machine learning (ML) has recently shown promise in identifying cognitive impairment in patients with PD.

Objective: This study aims to summarize different ML models applied to cognitive impairment in patients with PD and to identify determinants for improving diagnosis and predictive power for early detection of cognitive impairment.

Methods: PubMed, Cochrane, Embase, and Web of Science were searched for relevant articles on March 2, 2024. The risk of bias was assessed using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2). Bivariate meta-analysis was used to estimate pooled sensitivity and specificity results, presented as odds ratio (OR) and 95% CI. A summary receiver operator characteristic (SROC) curve was used.

Results: A total of 38 articles met the criteria, involving 8564 patients with PD and 1134 healthy controls. Overall, 120 models reported sensitivity and specificity, with mean values of 71.07% (SD 13.72%) and 77.01% (SD 14.31%), respectively. Predictors commonly used in ML models included clinical features, neuroimaging features, and other variables. No significant heterogeneity was observed in the bivariate meta-analysis, which included 12 studies. Using sensitivity as the metric, the combined sensitivity and specificity were 0.76 (95% CI 0.67-0.83) and 0.83 (95% CI 0.76-0.88), respectively. When specificity was used, the combined values were 0.77 (95% CI 0.65-0.86) and 0.76 (95% CI 0.63-0.85), respectively. The area under the curves of the SROC were 0.87 (95% CI 0.83-0.89) and 0.83 (95% CI 0.80-0.86) respectively.

Conclusions: Our findings provide a comprehensive summary of various ML models and demonstrate the effectiveness of ML as a tool for diagnosing and predicting cognitive impairment in patients with PD.

Trial Registration: PROSPERO CRD42023480196; https://www.crd.york.ac.uk/PROSPERO/view/CRD42023480196

Introduction

Parkinson disease (PD) is the second most common progressive neurodegenerative disease. It has gained much attention from scientists due to its rising rates of disability and death [1]. The cause of PD involves the clumping of misfolded α-synuclein in Lewy bodies in neurons. It also involves mitochondrial problems, neuroinflammation, and oxidative stress. These processes lead to irreversible cellular damage and neuronal loss. Several factors influence the development of PD, including environmental factors such as exposure to pesticides, lifestyle factors, such as tobacco consumption and sedentary behavior,

genetic mutations, such as *GBA*, *LRRK2*, *PARK*, and *SNCA*, as well as demographic factors like advancing age and male gender [2]. Yet, the cause of PD remains unknown.

The cardinal motor symptoms of PD include bradykinesia, rigidity, rest tremor, and postural instability. In addition, nonmotor symptoms such as cognitive impairment, sleep disorders, and autonomic dysfunction are commonly observed. A recent study has stressed the rising importance of nonmotor symptoms in the later stages of the disease. It has highlighted their key role in PD management [3].

Cognitive impairment in PD is marked by problems in many cognitive domains, including attention, memory, executive function, language, and visuospatial function [4]. The spectrum of cognitive impairment in PD ranges from mild cognitive impairment (MCI) to PD dementia (PD-D) [5]. The definition and progression of PD-MCI and PD-D exhibit significant heterogeneity. PD-MCI is defined as a decline in cognition that deviates from what is expected based on the patient's age and education level but does not meet the criteria for impaired cognitive function [6]. MCI can occur in the early stages of the disease, even before the onset of motor symptoms, and may be overlooked. In contrast, dementia typically appears in the advanced stages of PD and causes impairments across multiple cognitive domains. These deficits are severe enough to disrupt the daily life of motor symptoms. Patients who meet specific criteria [7] can be diagnosed with probable or possible PD-D. Researchers have identified MCI as a significant risk factor for the development of dementia in individuals with PD [8]. A meta-analysis has shown differences in biomarkers and brain imaging between patients with PD-D and those with PD-MCI [9].

Recently, the use of artificial intelligence has grown, specifically machine learning (ML). ML is a subfield of artificial intelligence. It focuses on developing algorithms and statistical models that allow computers to perform tasks without explicit programming. The main goal of ML is to enable computers to learn from data and make predictions or decisions based on that learning [10,11]. ML has 4 types: supervised, unsupervised, semisupervised, and reinforcement learning. ML can be broadly classified into 2 main categories: supervised and unsupervised learning. The former encompassed a range of models, including support vector machine (SVM), random forest, K-nearest neighbors, and naive Bayes, among others. In contrast, unsupervised learning involves clustering, dimensionality reduction, and so on. They differ in their use of labeled or unlabeled data to train models. Semisupervised learning is a hybrid approach. It combines supervised and unsupervised learning. Reinforcement learning does not require data with labels; instead, it learns from experiences by interacting with the environment, observing, and responding to results.

Current evidence has proven that ML is useful in predictive analytics for medicine. It can analyze complex datasets, including clinical data, genetic information, and imaging features to aid in the early detection and diagnosis of diseases. Besides, it can assess the risk of progression and prognosis. For example, in conditions like Alzheimer disease and PD, ML models can find patterns in imaging data for more accurate diagnosis. The

latest research trained an XGBoost (extreme gradient boosting) model to predict PD-associated genes using genomic, transcriptomic, and epigenomic data from brain tissues and dopaminergic neurons [12]. ML has been used as a valuable technique to predict the high risk of disease conversion [13]. Vast health data is available, including clinical, genetic, imaging, and biomarker data. They can be used to diagnose and predict prognosis, rather than being based on symptoms and performed by specialist neurologists.

In this review, our goal is not to give a full literature review of articles applying ML to clinical problems, nor do we aim to delve into the complex mathematical details of numerous ML methods. Instead, we focus on summarizing the existing ML models. They are used to diagnose PD with cognitive impairment and to predict cognitive decline as the disease progresses. We also aim to list predictors that may help in diagnosis, progression, and prognosis in patients with PD with cognitive impairment.

Methods

Search Strategy

Two authors (WYY and XY) independently searched PubMed, Cochrane, Embase, and Web of Science databases for relevant articles using the keywords "Parkinson's disease," "Machine Learning," and "cognitive decline" in different combinations (Table S1 in Multimedia Appendix 1). The included studies were from the start of the database until March 2, 2024. The review was reported according to the PRISMA (Preferred Reporting Items for Systematic Review and Meta-Analyses) statement. It is registered in PROSPERO (International Prospective Register of Systematic Reviews; number CRD42023480196).

Eligibility Criteria and Study Selection

Studies were eligible if they aimed to diagnose or predict the cognitive impairment in PD. The target condition was the normal cognition (PD-NC), PD-MCI, and PD-D. ML models were carried out to predict the target condition. Studies were eligible if these reported data on the following: true positive (TP), true negative (TN), false positive (FP), false negative (FN), sensitivity, specificity, accuracy, positive predictive value, or negative predictive value, and others. In case of not reporting TP, TN, FP, or FN, these were calculated from known variables (sensitivity and specificity).

The articles retrieved from the electronic databases were imported into EndNote (Clarivate) for further analysis. Duplicates, reviews, dissertations, cases, and conference abstracts were systematically eliminated through an automated process. Two authors (WYY and XY) independently screened the titles and abstracts of the rest of the records. Any disagreements were decided by the third author (CYF). Following this step, the full text of the remaining studies was downloaded and carefully reviewed, resulting in the final selection of relevant articles.

The data collected from the chosen studies included publication details such as title, first author, and year of publication, as well as study information such as study design, whether it was

conducted at single or multiple centers, diagnostic criteria for PD and cognitive impairment, ML models employed, predictors used, and value index assessed.

Risk of Bias

There is no widely accepted checklist for assessing the quality of diagnostic ML papers. We chose combined criteria that had been used in a previous report [14]. Two review authors (WYY and XY) independently assessed the risk of bias based on the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) [15], and any disagreements were resolved through the third author (CYF). If one of the questions was scored at a high risk of bias, the domain was scored at a high risk of bias. At least one domain at high risk of bias resulted in an overall score of high risk of bias, and only one domain scored as unclear risk of bias resulting in an overall score of unclear risk of bias for that paper.

Performance Metric and Meta-Analysis

Due to the diversity of models and predictors under study, models were categorized by type (random forest, neural network, etc) irrespective of variable selection procedures in the study. Similarly, predictor variables were grouped for analysis. For example, demographic features (such as age, gender, family history, and education), clinical features (such as motor symptoms, nonmotor symptoms, and various scale tests, such as Mini-Mental Status Examination [MMSE] and Montreal Cognitive Assessment [MoCA]), neuroimaging features (such as magnetic resonance imaging [MRI] and single photon emission computed tomography features), biofluid features (such as -synuclein [-syn], amyloid β [Aβ], and phospho-microtubule-associated protein tau [p-tau] in cerebrospinal fluid and blood plasma), genetic features (such as *GBA*, *LRRK2*, and single-nucleotide polymorphism [SNP] variants), and quantitative electroencephalography.

Given that classification algorithms in ML can indeed be considered as a diagnostic test, and that the results included in the article, ultimately, we used bivariate meta-analysis [16] to estimate the pooled sensitivity and specificity results, along with their corresponding 95% CIs. Estimated sensitivity and specificity were represented as forest plots. A summary receiver operator characteristic (SROC) curve was generated to evaluate the accuracy of ML for the prediction of cognitive impairment in patients with PD [17]. The test of heterogeneity was assessed using the Cochrane Q-test and I^2 test [18]. Heterogeneity was defined as significant if the Q-test results showed $P<.05$ or $I^2>50\%$. Potential publication bias was estimated using Deeks' funnel plot. In the event of bias, it may be evidenced by the visual asymmetry observed in conventional funnel plots. A Fagan nomogram was used to identify the maximum pretest and posttest likelihood. We also conducted sensitivity analysis by removing studies with a high risk of bias to identify potential sources of heterogeneity. Considering not all studies reported the results of both the train and validation sets, we also analyzed the single set. Statistical analyses were performed using GraphPad Prism 9.5 (GraphPad Software, Inc) and Stata (version 17; StataCorp). The level of significance was set at $P<.05$.

Results

Study Selection

Figure 1 shows the literature selection and filtering process following the PRISMA 2020 guidelines [19]. A total of 1674 records were identified from 4 databases. After removing 199 duplicates, 1475 records were screened, with 785 excluded for being unrelated (reviews, editorials, conference abstracts, case reports, etc). Of the 690 records sought for retrieval, 3 were not obtained. Finally, 687 reports were assessed for eligibility, with 137 excluded for not being related to PD and 511 excluded for being irrelevant. Finally, 38 studies [20-57] published between 2013 and 2024 were included in the qualitative analysis, from which 12 studies [20-31] were included in the meta-analysis. A total of 9698 participants were involved in the current study, comprising 8564 patients with PD and 1134 healthy controls in qualitative analysis. Additional information can be found in Table S2 in Multimedia Appendix 1 [20-57].

Figure 1. Flowchart illustrating the process of searching for screening and inclusion. PD: Parkinson disease.
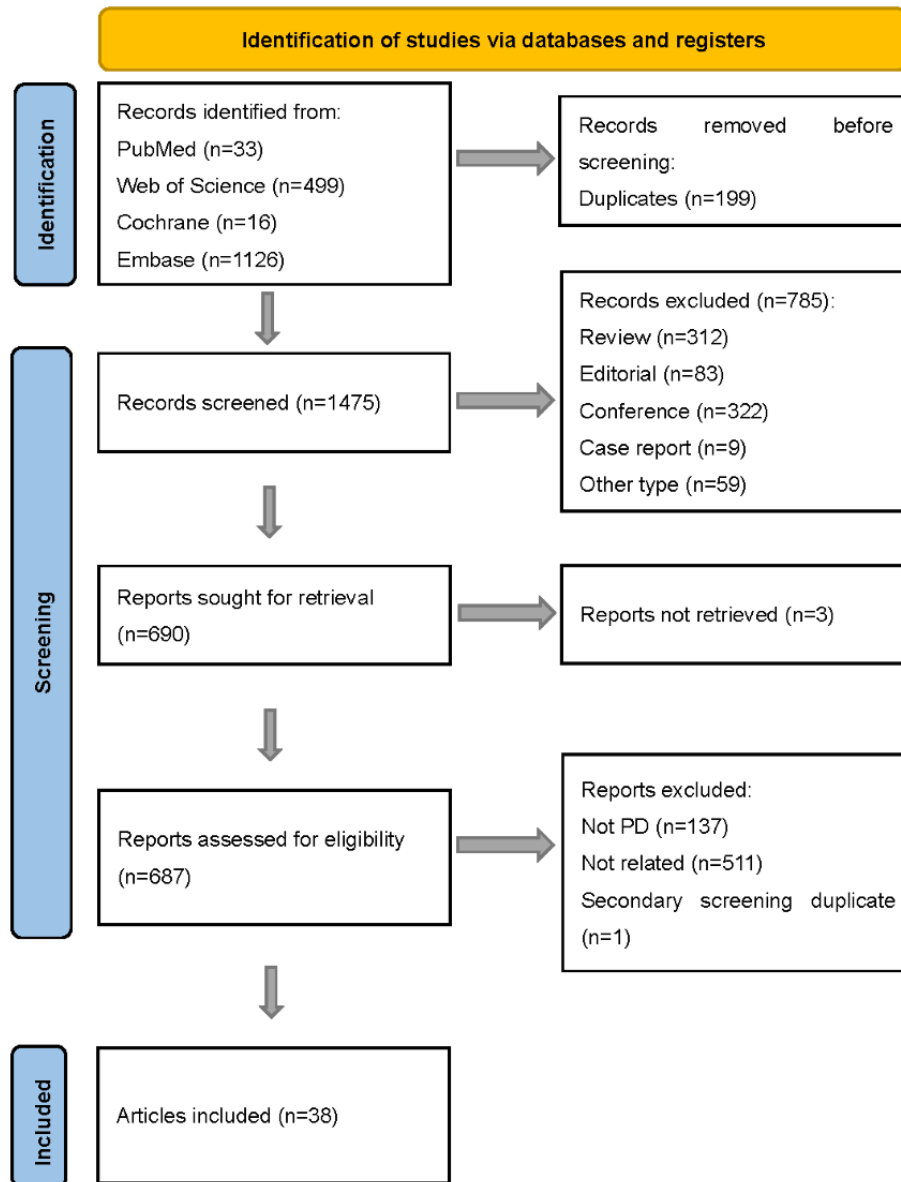
Risk of Bias

Table S3 in Multimedia Appendix 1 shows the risk of bias status of each included article. A total of 18 [24,27,29,32,33,39,42-47,51,52,54-57] out of 38 (47%) studies [20-57] were scored as high risk of bias, 9 [30,31,37,38,40,41,49,50,53] out of 38 (23%) studies were scored as unclear risk of bias, and the rest of the studies were low risk of bias. Studies scored a high risk of bias for failing to describe their study population (patient selection), losing the gold standard for diagnosing cognitive impairment (reference standard).

ML Models and Predictors

In the 38 studies [20-57], 172 models were constructed with different ML models and predictors. However, 142 models were subjected to analysis, with accuracy being the primary outcome. These included 45 SVM models, 29 random forest models, 9 K-nearest neighbors models, 18 naive Bayes models, and other supervised learnings. Also, we included 3 semisupervised

models that used SVM with principal component analysis PCA and 2 neural networks. The overall mean accuracy of 142 models was 74.32% (SD 10.25%). Of the 89 models that reported area under the receiver operating characteristic curve, the mean accuracy was 0.74 (SD 0.12). Also, 120 models reported sensitivity and specificity, with mean values of 71.07% (SD 13.72%) and 77.01% (SD 14.31%), respectively.

This study discovered the predictors used to diagnose and predict cognitive impairment in PD (Figure 2A). Clinical features have the highest percentage, followed by neuroimaging and demographic features. Other predictors such as biofluid markers, genetic variants, electroencephalography, and gait features show a more balanced or lower percentage representation. The top-used predictors were summarized in 142 models (Figure 2B). The demographic and clinical characteristics that emerged as the most significant predictors included age, gender, and education level, as well as the Unified Parkinson's Disease Rating Scale (UPDRS), rapid eye movement sleep behavior disorder, MMSE, and MoCA. Furthermore, additional neuroimaging and biomarkers have been identified as potentially

contributing to the prediction process. Overall, 116 out of 142 models used a single type of feature to diagnose, with a mean accuracy of 73.42% (SD 10.24%). Nearly 18% of models used multiple types of features combined, with a higher accuracy of 78.57% (SD 9.2%). In the single type of model, the predictors with the highest accuracy were neuroimaging features, achieving an accuracy of 78.78% (SD 11.48%). Furthermore, 13 articles used longitudinal data to predict cognitive impairment or conversion, with the majority of predictors being neuroimaging features and biofluid markers.

Figure 2. Predictors used for diagnosing and predicting cognitive impairment in Parkinson disease in included models. EEG: electroencephalography; MMSE: Mini-Mental Status Examination; MoCA: Montreal Cognitive Assessment; MRI: magnetic resonance imaging; p-tau: phospho-microtubule-associated protein tau; RBD: rapid eye movement sleep behavior disorder; SNP: single nucleotide polymorphism; UPDRS: Unified Parkinson's Disease Rating Scale; α-syn: α-synuclein.
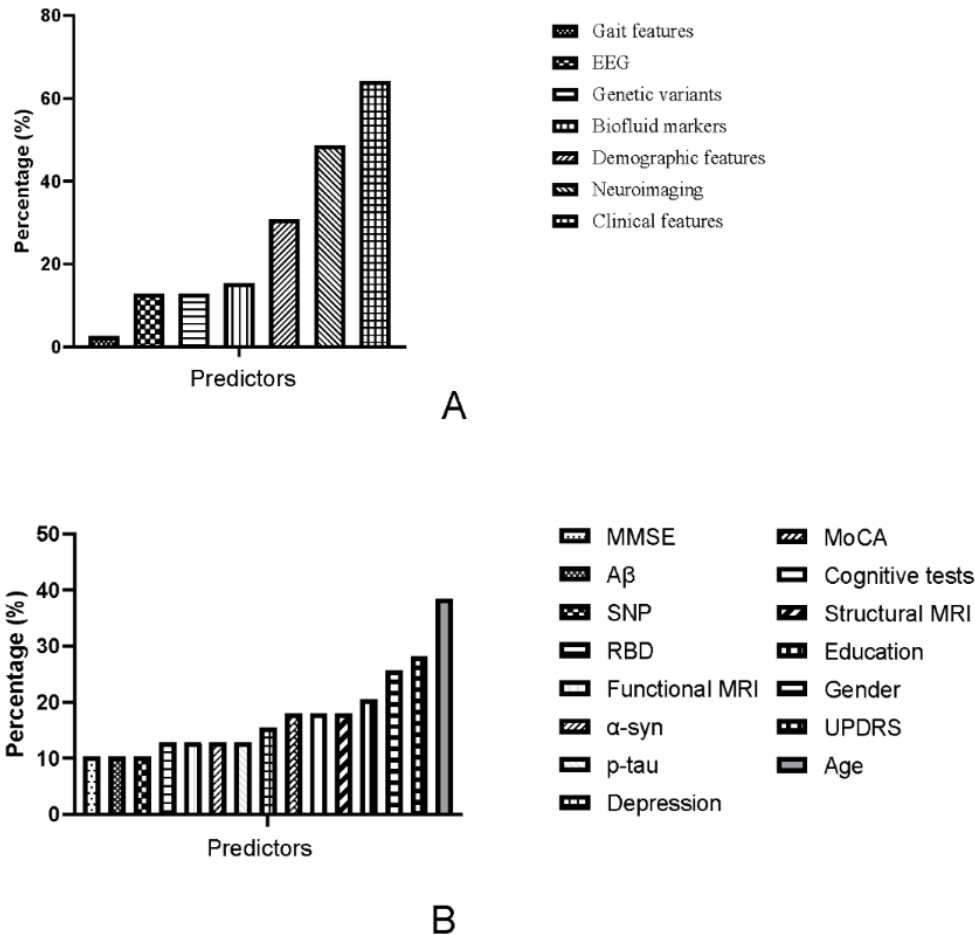
Meta-Analysis

Due to limited reporting of TP, TN, FP, FN, sensitivity, and specificity, only 12 studies were conducted for meta-analysis [20-31]. For studies with multiple models, we prioritized the one with the highest accuracy, reported in 63% (24/38) of the articles, to represent optimal model performance. A total of 7 studies [25-31] reported outcomes in the train set or validation set, while 5 studies [20-24] reported outcomes in both the train set and validation set. Forest plots for sensitivity and specificity are shown in Figure 3 for the train and validation set, respectively, and Figure S1 in Multimedia Appendix 1 for the single set (train or validation set). The training set sensitivity exhibited different results in the Q-test (Q=5.68, P=.22) and I^2

test (I^2=29.53). Other sensitivity and specificity values showed no heterogeneity. Using sensitivity, the combined sensitivity and specificity were 0.76 (95% CI 0.67-0.83) and 0.83 (95% CI 0.76-0.88), respectively. When using specificity, the combined values were 0.77 (95% CI 0.65-0.86) and 0.76 (95% CI 0.63-0.85), respectively. The area under the curves of the SROC were 0.87 (95% CI 0.83-0.89) and 0.83 (95% CI 0.8-0.86), respectively (Figure 4; Figure S2 in Multimedia Appendix 1). Fagan plots were constructed to illustrate the pretest and posttest probability of ML predicting cognitive impairment in patients with PD (Figures S3 and S4 in Multimedia Appendix 1). The Deeks' test showed no significant publication bias for ML (Figures S5 and S6 in Multimedia Appendix 1).

Figure 3. Forest plots for sensitivity and specificity in studies with both train and validation sets.
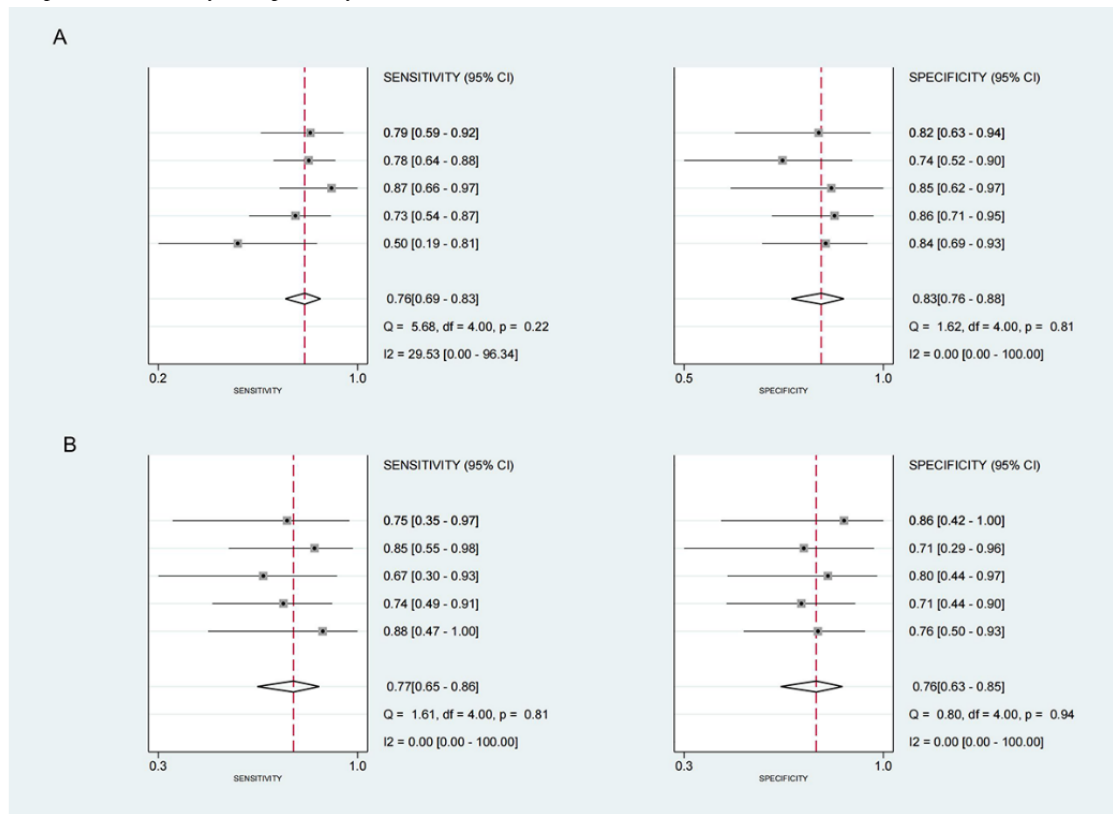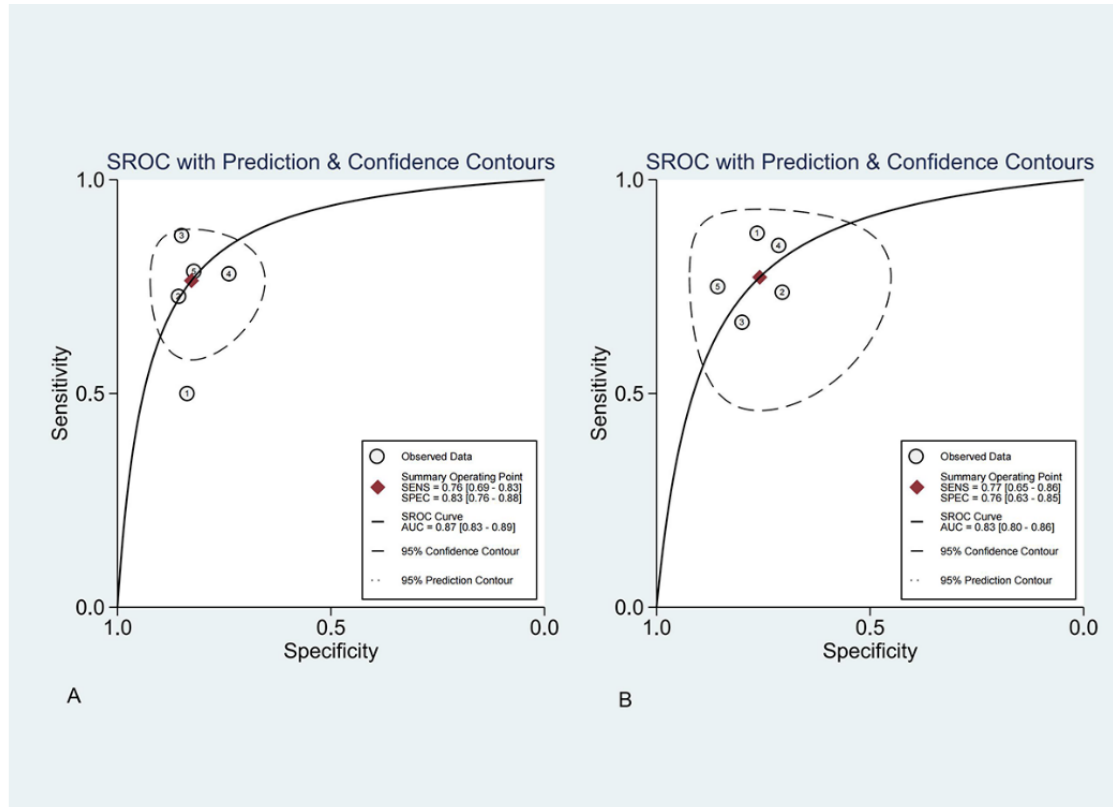

Figure 4. Summary receiver operating characteristic curve for sensitivity and specificity in studies with both train and validation sets. AUC: area under the curve; SROC: summary receiver operating characteristic.

Sensitivity Analysis

According to the risk of bias analysis, 3 articles scored high. After excluding these studies, all results exhibited no statistically significant heterogeneity, for the sensitivity of the training set. The results indicated a moderate degree of heterogeneity between studies (I^2=45.45). However, the results suggested no statistically significant heterogeneity between studies in the

Q-test (Q=5.5, *P*=.14; Figure S7 in Multimedia Appendix 1). The sensitivity, specificity, and SROC were consistent with the original analysis (Figure S9 in Multimedia Appendix 1). In the single set, there exists high heterogeneity both in sensitivity (Q=14.6, *P*=.01; I^2=72.59) and in specificity (Q=96.07, *P*<.01; I^2=95.84; Figure S8 in Multimedia Appendix 1). The results were consistent with the original analysis, with high sensitivity and specificity, and good overall diagnostic accuracy (Figures S9 and S10 in Multimedia Appendix 1). Fagan plots are shown in Figures S11 and S12 in Multimedia Appendix 1. The Deeks' test also found no significant publication bias for ML (Figures S13 and S14 in Multimedia Appendix 1).

Discussion

Principal Findings

This meta-analysis provided new insights into the efficacy of ML in diagnosing cognitive impairment in patients with PD. The average accuracy of 142 models was 74.32%, with 120 models reporting mean sensitivity and specificity values of 71.07% (SD 13.72%) and 77.01% (SD 14.31%), respectively. No significant heterogeneity was found among the sensitivities and specificities reported by the analyzed included in our meta-analysis. Nearly 50% of the included studies had a high risk of bias affecting applicability. A total of 12 studies [22,31,36,37-44,53] explored the cognition changes in a longitudinal cohort of patients with PD. Although complete data were not obtained in this study, we could conclude that the ML can predict potential cognitive impairment at baseline and forecast changes in cognitive state.

According to the QUADAS-2, most of the studies had high or unclear risk in the patient selection and reference standard. Part of the studies lacked detailed patient source descriptions. However, these articles were not excluded due to the use of Movement Disorder Society diagnostic criteria for cognitive impairment in patients with PD [7,58]. Many studies faced similar issues in patient selection, and complete exclusion would significantly reduce the available data. The diverse metrics used to assess ML resulted in incomplete reporting of expected metrics, increasing bias risk. These problems may lead to unstable meta-analysis results and make it difficult to provide meaningful conclusions. Furthermore, the low heterogeneity between studies limits the values of sensitivity analyses.

The meta-analysis of 12 studies [20-31] demonstrated that ML models exhibited greater consistency in terms of sensitivity and specificity in the training and validation sets. In addition, no heterogeneity or publication bias was observed, suggesting that the ML model exhibits superior performance in predicting cognitive impairment in patients with PD. However, further studies are essential to validate its clinical applicability.

Studies with a high risk of bias were excluded, and the results were found to be consistent with those of the original analyses. However, the I^2 in the sensitivity of the train set exhibited a slight increase, while the Q-test suggested no statistically significant heterogeneity. The sensitivity analysis demonstrated the credibility and robustness of the results. The results of the remaining 9 studies [20-23,25,26,28,30,31] remained stable compared to the original analysis and showed high sensitivity and specificity of the ML to diagnose and predict cognitive impairment in patients with PD. While, due to the limited number of studies included in the meta-analysis, further exploration was required.

ML Models and Predictors

A recent review divided ML into traditional ML (including supervised and unsupervised ML), and ML methods based on neural networks (such as convolutional neural networks and recurrent neural networks) [59]. The majority of research still focuses on traditional ML in this review, possibly because supervised learning still has advantages in ML algorithms and processed data. First, supervised learning can be further divided into classification and regression based on their tasks. It can differentiate between patients with cognitive impairment and those with mild cognitive normality and also can identify patients with MCI in PD from those who can be diagnosed with dementia. Unsupervised learning is primarily used for clustering, aiming to discover inherent groupings in the data. It is limited to categorizing patients into distinct groups by reducing the dimensionality of their shared characteristics. It cannot offer a definitive diagnosis due to the absence of labeled features. Second, although unsupervised learning is proving valuable in reducing the cost of labeling data and automatically identifying data structures and patterns, the results still require manual interpretation and analysis for interpretability [60].

Determining the superiority of ML methods is challenging. With the development of imaging technology and the increasing costs of data collection, unsupervised learning, deep learning, and neural networks may become more prevalent in the future. Notably, semisupervised learning also showed great potential [61]. Chen et al [32] combined SVM with principal component analysis to successfully extract 6 features from 32 predictors, improving the performance over the simple SVM model. Other models use unsupervised learning to cluster this unlabeled data and then use supervised learning for longitudinal prediction [33]. Further research and discussion are needed to assess the accuracy and effectiveness of these different models.

We found that cognitive tests such as the MMSE and the MoCA were mostly used as predictors. Meanwhile, the UPDRS score as a specific feature that can reflect the patient's severity of motor and nonmotor symptoms was highly recommended as a potential predictor. In addition, with the advancement of neuroimaging techniques, not only traditional neuroimaging, such as cortical thickness but also functional MRI was widely used to diagnose and predict disease progression, including Fluorine-18 fluorodeoxyglucose positron emission tomography, which can reflect the metabolism in the brain and diffusion tensor imaging which can show the damages to brain connectivity. On the other hand, demographic features such as older age, male sex, and lower education, which were considered as highly risk factors to cause PD, were highly ranked. Biofluid markers, total-tau, p-tau, Aβ, and -syn detected in the cerebrospinal fluid were the most popular approaches and were deemed as the vital factors to show the prognosis. Furthermore, the rise of genomics has become a crucial factor in diagnosis and prediction. The growing popularity of genome-wide

association studies research has made SNP variation a focal point of interest, more and more SNP variants that count were found to cause PD. The *APOE* gene, which is the risk factor in AD, is also a potential predictor to predict cognitive impairment in PD. This review also shows that the performances of ML vary depending on its feature input. Some studies have found inconsistent accuracy when using single or multiple types of features. Some studies have reported greater accuracy in models with a single type of feature model [34], while others have concluded that the models incorporating multiple types of features are more effective [21,24,26]. Despite the remaining discrepancies in the existing research, integrating multiple types of features for predictive models appears to yield better results overall.

This review also summarized the longitudinal studies on predicting cognitive impairment and conversion in PD [22,31,36-39]. We discovered that neuroimaging features and biofluid markers are robust predictors. With the advancement of imaging technology, such as functional MRI and magnetic resonance spectroscopy, and their noninvasive detection capabilities, neuroimaging has become mainstream today. Neural networks and deep learning are especially focused on applications in imaging. Numerous research studies have suggested that biomarkers such as Aβ, -syn, and tau, play an essential role in the pathologic changes in PD. The detection of these biomarkers in fluids such as cerebrospinal fluid and blood becomes a potential approach to improving the accuracy of ML in diagnosing and predicting the status of cognitive impairment and disease progression.

A personal review proposed new diagnostic criteria: a 3-component system (SynNeurGe). It includes the harmful α-synuclein (S) accumulated in tissues or cerebrospinal fluid, evidence of neurodegeneration (N) in imaging, and the disease-causing gene variants (G) for PD. They are associated with clinical symptoms, defined either by a single highly specific clinical feature or by multiple less specific clinical features. A biological classification will aid in both basic and clinical research and bring the field closer to the precision medicine needed to develop disease-modifying therapies [62]. Consequently, a universally applicable ML model that integrates these significant variations can lead to personalized diagnosis and prediction, providing immense value for early detection, intervention, and management.

Strengths and Limitations

A large number of studies have investigated the diagnostic accuracy of ML to diagnose cognitive impairment in patients with PD. With the increasing number of deep learning and detection approaches, a more comprehensive exploration of ML was conducted. We used bivariate analysis to show more robust results. However, there are still several limitations in this review. First, we excluded nonpublished data in the meta-analysis. Due to heterogeneity in intervention methods and outcome measures between studies, it was not possible to synthesize all articles included using meta-analysis. Second, several studies used Parkinson's Disease Progression Markers Initiative datasets, as they focus on the longitudinal cognitive change. However, we did not analyze the predictive power of ML in the follow-up cohort [22,31,36,38-46]. We are also limited to analyzing the optimum models and the predictors. Further investigation is required to determine the most effective model or predictors for identifying patients with PD who have cognitive impairment, predicting the disease progression (including conversion to dementia), and providing personalized treatment options.

Conclusions

ML algorithms have been proven to be highly effective in diagnosing cognitive impairment in patients with PD, especially MCI, compared to normal cognitive function. Furthermore, ML holds significant potential for predicting the cognitive decline in patients with risk factors and the transition from MCI to dementia.

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during this study.

Authors' Contributions

WYY contributed to conceptualization, data curation, investigation, validation and formal analysis, software, visualization, and writing—original draft. ORW assisted with conceptualization, project administration, supervision, and writing—review & editing. SHF handled conceptualization, supervision, and writing—review & editing. XY and CYF contributed to data curation, investigation, and writing—review & editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Additional figures and tables.
[DOCX File , 2001 KB-Multimedia Appendix 1]

Multimedia Appendix 2

PRISMA checklist.

[DOCX File , 31 KB-Multimedia Appendix 2]

References

1. GBD 2021 Nervous System Disorders Collaborators. Global, regional, and national burden of disorders affecting the nervous system, 1990-2021: a systematic analysis for the global burden of disease study 2021. Lancet Neurol. 2024;23(4):344-381. [FREE Full text] [doi: 10.1016/S1474-4422(24)00038-3] [Medline: 38493795]

2. Jankovic J, Tan EK. Parkinson's disease: etiopathogenesis and treatment. J Neurol Neurosurg Psychiatry. 2020;91(8):795-808. [doi: 10.1136/jnnp-2019-322338] [Medline: 32576618]

3. Schapira AHV, Chaudhuri KR, Jenner P. Non-motor features of Parkinson disease. Nat Rev Neurosci. 2017;18(7):435-450. [FREE Full text] [doi: 10.1038/nrn.2017.62] [Medline: 28592904]

4. Hanagasi HA, Tufekcioglu Z, Emre M. Dementia in Parkinson's disease. J Neurol Sci. 2017;374:26-31. [doi: 10.1016/j.jns.2017.01.012] [Medline: 28088312]

5. Aarsland D, Batzu L, Halliday GM, Geurtsen GJ, Ballard C, Ray Chaudhuri K, et al. Parkinson disease-associated cognitive impairment. Nat Rev Dis Primers. 2021;7(1):47. [doi: 10.1038/s41572-021-00280-3] [Medline: 34210995]

6. Litvan I, Goldman JG, Tröster AI, Schmand BA, Weintraub D, Petersen RC, et al. Diagnostic criteria for mild cognitive impairment in Parkinson's disease: movement disorder society task force guidelines. Mov Disord. 2012;27(3):349-356. [FREE Full text] [doi: 10.1002/mds.24893] [Medline: 22275317]

7. Emre M, Aarsland D, Brown R, Burn DJ, Duyckaerts C, Mizuno Y, et al. Clinical diagnostic criteria for dementia associated with Parkinson's disease. Mov Disord. 2007;22(12):1689-707; quiz 1837. [doi: 10.1002/mds.21507] [Medline: 17542011]

8. Saredakis D, Collins-Praino LE, Gutteridge DS, Stephan BCM, Keage HAD. Conversion to MCI and dementia in Parkinson's disease: a systematic review and meta-analysis. Parkinsonism Relat Disord. 2019;65:20-31. [doi: 10.1016/j.parkreldis.2019.04.020] [Medline: 31109727]

9. Delgado-Alvarado M, Gago B, Navalpotro-Gomez I, Jiménez-Urbieta H, Rodriguez-Oroz MC. Biomarkers for dementia and mild cognitive impairment in Parkinson's disease. Mov Disord. 2016;31(6):861-881. [doi: 10.1002/mds.26662] [Medline: 27193487]

10. Deo RC. Machine learning in medicine. Circulation. 2015;132(20):1920-1930. [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.115.001593] [Medline: 26572668]

11. Sarker IH. Machine learning: algorithms, real-world applications and research directions. SN Comput Sci. 2021;2(3):160. [FREE Full text] [doi: 10.1007/s42979-021-00592-x] [Medline: 33778771]

12. Yu E, Larivière R, Thomas RA, Liu L, Senkevich K, Rahayel S, et al. Machine learning nominates the inositol pathway and novel genes in Parkinson's disease. Brain. 2024;147(3):887-899. [FREE Full text] [doi: 10.1093/brain/awad345] [Medline: 37804111]

13. Mallo SC, Valladares-Rodriguez S, Facal D, Lojo-Seoane C, Fernández-Iglesias MJ, Pereiro AX. Neuropsychiatric symptoms as predictors of conversion from MCI to dementia: a machine learning approach. Int Psychogeriatr. 2020;32(3):381-392. [FREE Full text] [doi: 10.1017/S1041610219001030] [Medline: 31455461]

14. Fleuren LM, Klausch TLT, Zwager CL, Schoonmade LJ, Guo T, Roggeveen LF, et al. Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy. Intensive Care Med. 2020;46(3):383-400. [FREE Full text] [doi: 10.1007/s00134-019-05872-y] [Medline: 31965266]

15. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529-536. [FREE Full text] [doi: 10.7326/0003-4819-155-8-201110180-00009] [Medline: 22007046]

16. Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol. 2005;58(10):982-990. [doi: 10.1016/j.jclinepi.2005.02.022] [Medline: 16168343]

17. Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JAC. A unification of models for meta-analysis of diagnostic accuracy studies. Biostatistics. 2007;8(2):239-251. [doi: 10.1093/biostatistics/kxl004] [Medline: 16698768]

18. Dinnes J, Deeks J, Kirby J, Roderick P. A methodological review of how heterogeneity has been examined in systematic reviews of diagnostic test accuracy. Health Technol Assess. 2005;9(12):1-113, iii. [FREE Full text] [doi: 10.3310/hta9120] [Medline: 15774235]

19. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]

20. Abós A, Baggio HC, Segura B, García-Díaz AI, Compta Y, Martí MJ, et al. Discriminating cognitive status in Parkinson's disease through functional connectomics and machine learning. Sci Rep. 2017;7:45347. [FREE Full text] [doi: 10.1038/srep45347] [Medline: 28349948]

21. Amboni M, Ricciardi C, Adamo S, Nicolai E, Volzone A, Erro R, et al. Machine learning can predict mild cognitive impairment in Parkinson's disease. Front Neurol. 2022;13:1010147. [FREE Full text] [doi: 10.3389/fneur.2022.1010147] [Medline: 36468069]

22. Booth S, Park KW, Lee CS, Ko JH. Predicting cognitive decline in Parkinson's disease using FDG-PET-based supervised learning. J Clin Invest. 2022;132(20):e157074. [FREE Full text] [doi: 10.1172/JCI157074] [Medline: 36040832]

23. Zhang J, Li Y, Gao Y, Hu J, Huang B, Rong S, et al. An SBM-based machine learning model for identifying mild cognitive impairment in patients with Parkinson's disease. J Neurol Sci. 2020;418:117077. [doi: 10.1016/j.jns.2020.117077] [Medline: 32798842]

24. Zhang J, Gao Y, He X, Feng S, Hu J, Zhang Q, et al. Identifying Parkinson's disease with mild cognitive impairment by using combined MR imaging and electroencephalogram. Eur Radiol. 2021;31(10):7386-7394. [doi: 10.1007/s00330-020-07575-1] [Medline: 33389038]

25. Cengiz S, Arslan DB, Kicik A, Erdogdu E, Yildirim M, Hatay GH, et al. Identification of metabolic correlates of mild cognitive impairment in Parkinson's disease using magnetic resonance spectroscopic imaging and machine learning. MAGMA. 2022;35(6):997-1008. [doi: 10.1007/s10334-022-01030-6] [Medline: 35867235]

26. Chen B, Xu M, Yu H, He J, Li Y, Song D, et al. Detection of mild cognitive impairment in Parkinson's disease using gradient boosting decision tree models based on multilevel DTI indices. J Transl Med. 2023;21(1):310. [FREE Full text] [doi: 10.1186/s12967-023-04158-8] [Medline: 37158918]

27. García AM, Arias-Vergara T, C Vasquez-Correa J, Nöth E, Schuster M, Welch AE, et al. Cognitive determinants of dysarthria in Parkinson's disease: an automated machine learning approach. Mov Disord. 2021;36(12):2862-2873. [doi: 10.1002/mds.28751] [Medline: 34390508]

28. Lin H, Liu Z, Yan W, Zhang D, Liu J, Xu B, et al. Brain connectivity markers in advanced Parkinson's disease for predicting mild cognitive impairment. Eur Radiol. 2021;31(12):9324-9334. [doi: 10.1007/s00330-021-08086-3] [Medline: 34109485]

29. Morales DA, Vives-Gilabert Y, Gómez-Ansón B, Bengoetxea E, Larrañaga P, Bielza C, et al. Predicting dementia development in Parkinson's disease using bayesian network classifiers. Psychiatry Res. 2013;213(2):92-98. [doi: 10.1016/j.pscychresns.2012.06.001] [Medline: 23149030]

30. Yu Z, Pang H, Yu H, Wu Z, Ding Z, Fan G. Segmental disturbance of white matter microstructure in predicting mild cognitive impairment in idiopathic Parkinson's disease: an individualized study based on automated fiber quantification tractography. Parkinsonism Relat Disord. 2023;115:105802. [doi: 10.1016/j.parkreldis.2023.105802] [Medline: 37734997]

31. Harvey J, Reijnders RA, Cavill R, Duits A, Köhler S, Eijssen L, et al. Machine learning-based prediction of cognitive outcomes in de novo Parkinson's disease. NPJ Parkinsons Dis. 2022;8(1):150. [FREE Full text] [doi: 10.1038/s41531-022-00409-5] [Medline: 36344548]

32. Chen PH, Hou TY, Cheng FY, Shaw JS. Prediction of cognitive degeneration in Parkinson's disease patients using a machine learning method. Brain Sci. 2022;12(8):1048. [FREE Full text] [doi: 10.3390/brainsci12081048] [Medline: 36009111]

33. Kawabata K, Watanabe H, Hara K, Bagarinao E, Yoneyama N, Ogura A, et al. Distinct manifestation of cognitive deficits associate with different resting-state network disruptions in non-demented patients with Parkinson's disease. J Neurol. 2018;265(3):688-700. [doi: 10.1007/s00415-018-8755-5] [Medline: 29392456]

34. Chaturvedi M, Bogaarts JG, Kozak Cozac VV, Hatz F, Gschwandtner U, Meyer A, et al. Phase lag index and spectral power as QEEG features for identification of patients with mild cognitive impairment in Parkinson's disease. Clin Neurophysiol. 2019;130(10):1937-1944. [FREE Full text] [doi: 10.1016/j.clinph.2019.07.017] [Medline: 31445388]

35. Hogue O, Fernandez HH, Floden DP. Predicting early cognitive decline in newly-diagnosed Parkinson's patients: a practical model. Parkinsonism Relat Disord. 2018;56:70-75. [FREE Full text] [doi: 10.1016/j.parkreldis.2018.06.031] [Medline: 29936131]

36. Huang X, He Q, Ruan X, Li Y, Kuang Z, Wang M, et al. Structural connectivity from DTI to predict mild cognitive impairment in de novo Parkinson's disease. Neuroimage Clin. 2024;41:103548. [FREE Full text] [doi: 10.1016/j.nicl.2023.103548] [Medline: 38061176]

37. Liu G, Locascio JJ, Corvol JC, Boot B, Liao Z, Page K, HBS, CamPaIGN, PICNICS, PROPARK, PSG, DIGPD, et al. PDBP. Prediction of cognition in Parkinson's disease with a clinical-genetic score: a longitudinal analysis of nine cohorts. Lancet Neurol. 2017;16(8):620-629. [FREE Full text] [doi: 10.1016/S1474-4422(17)30122-9] [Medline: 28629879]

38. Shin NY, Bang M, Yoo SW, Kim JS, Yun E, Yoon U, et al. Cortical thickness from MRI to predict conversion from mild cognitive impairment to dementia in Parkinson disease: a machine learning-based model. Radiology. 2021;300(2):390-399. [doi: 10.1148/radiol.2021203383] [Medline: 34032515]

39. Tang C, Zhao X, Wu W, Zhong W, Wu X. An individualized prediction of time to cognitive impairment in Parkinson's disease: a combined multi-predictor study. Neurosci Lett. 2021;762:136149. [doi: 10.1016/j.neulet.2021.136149] [Medline: 34352339]

40. Almgren H, Camacho M, Hanganu A, Kibreab M, Camicioli R, Ismail Z, et al. Machine learning-based prediction of longitudinal cognitive decline in early Parkinson's disease using multimodal features. Sci Rep. 2023;13(1):13193. [FREE Full text] [doi: 10.1038/s41598-023-37644-6] [Medline: 37580407]

41. Andersson S, Josefsson M, Stiernman LJ, Rieckmann A. Cognitive decline in Parkinson's disease: a subgroup of extreme decliners revealed by a data-driven analysis of longitudinal progression. Front Psychol. 2021;12:729755. [FREE Full text] [doi: 10.3389/fpsyg.2021.729755] [Medline: 34566817]

42. Nguyen AA, Maia PD, Gao X, F Damasceno P, Raj A. Dynamical role of pivotal brain regions in Parkinson symptomatology uncovered with deep learning. Brain Sci. 2020;10(2):73. [FREE Full text] [doi: 10.3390/brainsci10020073] [Medline: 32019067]

43. Ostertag C, Visani M, Urruty T, Beurton-Aimar M. Long-term cognitive decline prediction based on multi-modal data using multimodal3DSiameseNet: transfer learning from alzheimer's disease to Parkinson's disease. Int J Comput Assist Radiol Surg. 2023;18(5):809-818. [FREE Full text] [doi: 10.1007/s11548-023-02866-6] [Medline: 36964477]

44. Salmanpour MR, Shamsaei M, Saberi A, Setayeshi S, Klyuzhin IS, Sossi V, et al. Optimized machine learning methods for prediction of cognitive outcome in Parkinson's disease. Comput Biol Med. 2019;111:103347. [doi: 10.1016/j.compbiomed.2019.103347] [Medline: 31284154]

45. Ortelli P, Ferrazzoli D, Versace V, Cian V, Zarucchi M, Gusmeroli A, et al. Optimization of cognitive assessment in parkinsonisms by applying artificial intelligence to a comprehensive screening test. NPJ Parkinsons Dis. 2022;8(1):42. [FREE Full text] [doi: 10.1038/s41531-022-00304-z] [Medline: 35410449]

46. Betrouni N, Delval A, Chaton L, Defebvre L, Duits A, Moonen A, et al. Electroencephalography-based machine learning for cognitive profiling in Parkinson's disease: Preliminary results. Mov Disord. Feb 2019;34(2):210-217. [doi: 10.1002/mds.27528] [Medline: 30345602]

47. Byeon H, Jin H, Cho S. Development of Parkinson's disease dementia prediction model based on verbal memory, visuospatial memory, and executive function. J Med Imaging Health Inform. 2017;7(7):1517-1521. [FREE Full text]

48. Byeon H. Application of machine learning technique to distinguish Parkinson's disease dementia and Alzheimer's dementia: predictive power of Parkinson's disease-related non-motor symptoms and neuropsychological profile. J Pers Med. Apr 28, 2020;10(2):31. [FREE Full text] [doi: 10.3390/jpm10020031] [Medline: 32354187]

49. Byeon H. Best early-onset Parkinson dementia predictor using ensemble learning among Parkinson's symptoms, rapid eye movement sleep disorder, and neuropsychological profile. World J Psychiatry. Nov 19, 2020;10(11):245-259. [FREE Full text] [doi: 10.5498/wjp.v10.i11.245] [Medline: 33269221]

50. Byeon H. Is the random forest algorithm suitable for predicting Parkinson's disease with mild cognitive impairment out of Parkinson's disease with normal cognition? Int J Environ Res Public Health. Apr 10, 2020;17(7):2594. [FREE Full text] [doi: 10.3390/ijerph17072594] [Medline: 32290134]

51. Byeon H. Exploring Parkinson's disease predictors based on basic intelligence quotient and executive intelligence quotient. Int J Adv Comput Sci Appl. 2021;12(4):106-111.

52. Byeon H. Predicting the severity of Parkinson's disease dementia by assessing the neuropsychiatric symptoms with an SVM regression model. Int J Environ Res Public Health. Mar 04, 2021;18(5):2551. [FREE Full text] [doi: 10.3390/ijerph18052551] [Medline: 33806474]

53. Chen F, Li Y, Ye G, Zhou L, Bian X, Liu J. Development and validation of a prognostic model for cognitive impairment in Parkinson's disease with REM sleep behavior disorder. Front Aging Neurosci. 2021;13:703158. [FREE Full text] [doi: 10.3389/fnagi.2021.703158] [Medline: 34322014]

54. Geraedts VJ, Koch M, Contarino MF, Middelkoop HAM, Wang H, van Hilten JJ, et al. Machine learning for automated EEG-based biomarkers of cognitive impairment during Deep Brain Stimulation screening in patients with Parkinson's Disease. Clin Neurophysiol. May 2021;132(5):1041-1048. [FREE Full text] [doi: 10.1016/j.clinph.2021.01.021] [Medline: 33743299]

55. Jeon J, Kim K, Baek K, Chung SJ, Yoon J, Kim YJ. Accuracy of machine learning using the Montreal Cognitive Assessment for the diagnosis of cognitive impairment in Parkinson's disease. J Mov Disord. May 2022;15(2):132-139. [FREE Full text] [doi: 10.14802/jmd.22012] [Medline: 35670022]

56. Kibtia H, Abdullah S, Bustamam A. Comparison of random forest and support vector machine for prediction of cognitive impairment in Parkinson's disease. AIP Conf Proc. 2020;2296(1). [FREE Full text]

57. Ramezani M, Mouches P, Yoon E, Rajashekar D, Ruskey JA, Leveille E, et al. Investigating the relationship between the SNCA gene and cognitive abilities in idiopathic Parkinson's disease using machine learning. Sci Rep. Mar 01, 2021;11(1):4917. [FREE Full text] [doi: 10.1038/s41598-021-84316-4] [Medline: 33649398]

58. Litvan I, Goldman JG, Tröster AI, Schmand BA, Weintraub D, Petersen RC, et al. Diagnostic criteria for mild cognitive impairment in parkinson's disease: movement disorder society task force guidelines. Mov Disord. 2012;27(3):349-356. [FREE Full text] [doi: 10.1002/mds.24893] [Medline: 22275317]

59. Greener JG, Kandathil SM, Moffat L, Jones DT. A guide to machine learning for biologists. Nat Rev Mol Cell Biol. 2022;23(1):40-55. [doi: 10.1038/s41580-021-00407-0] [Medline: 34518686]

60. Barredo Arrieta A, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, et al. Explainable atificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion. 2020;58:82-115. [FREE Full text] [doi: 10.1016/j.inffus.2019.12.012]

61.   Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Red HookNY. Curran Associates Inc; 2017. Presented at: NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems; 2017 December 4 - 9:4768-4777; Long Beach California USA.

62.   Höglinger GU, Adler CH, Berg D, Klein C, Outeiro TF, Poewe W, et al. A biological classification of parkinson's disease: the synNeurGe research diagnostic criteria. Lancet Neurol. 2024;23(2):191-204. [doi: 10.1016/S1474-4422(23)00404-0] [Medline: 38267191]

Abbreviations

-syn: -synuclein
Aβ: amyloid β
FN: false negative
FP: false positive
MCI: mild cognitive impairment
ML: machine learning
MMSE: Mini-Mental Status Examination
MoCA: Montreal Cognitive Assessment
MRI: magnetic resonance imaging
PD: Parkinson disease
PD-D: Parkinson disease dementia
PD-MCI: Parkinson disease mild cognitive impairment
PD-NC: Parkinson disease normal cognition
PRISMA: Preferred Reporting Items for Systematic Review and Meta-Analyses
PROSPERO: International Prospective Register of Systematic Reviews
p-tau: phospho-microtubule-associated protein tau
QUADAS-2: Quality Assessment of Diagnostic Accuracy Studies-2
SNP: single-nucleotide polymorphism
SROC: summary receiver operator characteristic
SVM: support vector machine
TN: true negative
TP: true positive
UPDRS: Unified Parkinson's Disease Rating Scale
XGBoost: extreme gradient boosting