The Role of Machine Learning in Cognitive Impairment in Parkinson's Disease: A Systematic Review and meta-analysis

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

Submitted to: Journal of Medical Internet Research
on: April 20, 2024

Table of Contents

Original Manuscript **4**
Supplementary Files **26**
Figures **27**
Figure 1 ..... **28**
Figure 2 ..... **29**
Figure 3 ..... **30**
Figure 4 ..... **31**
Figure 5 ..... **32**

The Role of Machine Learning in Cognitive Impairment in Parkinson's Disease: A Systematic Review and meta-analysis

Yanyun Wu[1]; Yangfan Cheng[1]; Yi Xiao[1]; Huifang Shang[1]; Ruwei Ou[1]

[1]Department of Neurology, West China Hospital, Sichuan University Chengdu, Sichuan CN

Corresponding Author:
Ruwei Ou
Department of Neurology, West China Hospital, Sichuan University
Department of Neurology, West China Hospital, Sichuan University, No.37, Guoxue Lane
Chengdu, Sichuan
CN

Abstract

Parkinson's disease (PD) is a common neurodegenerative disease characterized by both motor and non-motor 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 PD patients. This study aims to summarize different ML models applied to cognitive impairment in PD patients and to identify determinants for improving the diagnosis and predictive power to find cognitive impairment at an early stage. PubMed, Cochrane, Embase, and Web of Science for relevant articles were conducted on March 2, 2024. A total of 43 articles met the criteria, involving 9,139 PD patients and 1,353 healthy controls. A total of 151 models were analyzed, with an accuracy ranging from 60% to 90%. Predictors commonly used in ML models included clinical features, neuroimaging features, and other variables. In the bivariate meta-analysis, including only 12 studies, no significant heterogeneity was observed. 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.

(JMIR Preprints 20/04/2024:59649)
DOI: https://doi.org/10.2196/preprints.59649

Original Manuscript

The Role of Machine Learning in Cognitive Impairment in Parkinson's Disease: A Systematic Review and meta-analysis

Yanyun Wu[1], Yangfan Cheng[1], Yi Xiao[1], Huifang Shang[1], Ruwei Ou[1*]

[1] Department of Neurology, West China Hospital, Sichuan University, Chengdu, Sichuan, China

* Correspondence to Ruwei Ou, Email: ouruwei@aliyun.com; Telephone: 0086-18980607525; FAX:0086-028-85423550; Postal address: Department of Neurology, West China Hospital, Sichuan University, No.37, Guoxue Lane, 610041, Chengdu, Sichuan, China.

Abstract

Parkinson’s disease (PD) is a common neurodegenerative disease characterized by both motor and non-motor 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 PD patients. This study aims to summarize different ML models applied to cognitive impairment in PD patients and to identify determinants for improving the diagnosis and predictive power to find cognitive impairment at an early stage. PubMed, Cochrane, Embase, and Web of Science for relevant articles were conducted on March 2, 2024. A total of 43 articles met the criteria, involving 9,139 PD patients and 1,353 healthy controls. A total of 151 models were analyzed, with an accuracy ranging from 60% to 90%. Predictors commonly used in ML models included clinical features, neuroimaging features, and other variables. In the bivariate meta-analysis, including only 12 studies, no significant heterogeneity was observed. 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.

Keywords: Parkinson’s disease; cognitive impairment; machine learning; systematic review; meta-analysis

1 Introduction

Parkinson's disease (PD) is the second most common progressive brain 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 use and sedentary behavior, genetic mutations (e.g. *GBA*, *LRRK2*, *PARK*, and *SNCA*), 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. Additionally, non-motor symptoms such as cognitive impairment, sleep disorders, and autonomic dysfunction are commonly observed. A recent study has stressed the rising importance of non-motor 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 (AI) has grown, specifically machine learning (ML). ML is a subfield of AI. 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 four types: supervised, unsupervised, semi-supervised, and reinforcement learning. Supervised and unsupervised learning are the two main categories of ML. They differ in their use of labeled or unlabeled data to train models. Semi-supervised 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 diseases like Alzheimer's and Parkinson's, ML models can find patterns in imaging data for more accurate diagnosis. The latest research trained an XGBoost 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 PD patients with cognitive impairment.

2 Method

1.1 Search Strategy

We 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). The study included was from the start of the database until March 2, 2024. The review was reported according to the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) statement. It is registered in PROSPERO, the International Prospective Register of Systematic Reviews (No. CRD42023480196).

1.2 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 (Sen), specificity (Spe), accuracy, positive predictive value (PPV), or negative predictive value (NPV) and others. In case of not reporting TP, TN, FP, or FN, these were calculated from known variables (Sen and Spe).

The articles retrieved from the electronic databases were imported into Endnote for further analysis. Duplicates, reviews, dissertations, cases, and conference abstracts were systematically eliminated through automated process. To avoid any irrelevant studies related to PD, a thorough screening of titles and abstracts was carried out. 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 utilized, and value index assessed.

1.3 Risk of Bias

There is no widely accepted checklist for assessing the quality of diagnostic ML papers. We chose a combined criteria that have been used in a previous report [14]. Two review authors 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 authors.

1.4 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 (age; gender; familial history; education); clinical features (motor symptoms, non-motor symptoms, and various scale tests); neuroimaging features (MRI, SPECT features); biofluid features (a-synuclein, Abeta in cerebrospinal fluid and blood plasma); genetic features (*GBA*, *LRRK2*, and SNPs variants). Data are expressed as mean ± standard deviation (SD).

The bivariate meta-analysis [16] was used to estimate the pooled results for sensitivity and specificity, with their corresponding 95% confidence intervals (CIs). As various definitions were used in the studies selected, we selected MCI as the presence of disease and constructed 2 × 2 tables. A summary receiver operator characteristic (SROC) curve was generated to evaluate the accuracy of ML for the prediction of cognitive impairment in PD patients [17]. The statistical heterogeneity

from non-threshold was assessed using the Q value and I^2 statistic [18]. Potential publication bias was estimated using Deeks' funnel plot. A Fagan nomogram was used to identify the maximum pre-test and post-test likelihood. Analyses were carried out with SAS.

3 Results

1.5 Study Selection

Figure 1 shows the literature selection and filtering process following the PRISMA2020 guidelines [19]. Finally, 43 studies published between 2013 and 2024 were included in the qualitative analysis, from which 12 were included in the meta-analysis. A total of 10,492 participants was involved in the study, comprising 9,139 patients with PD and 1,353 healthy controls in qualitative analysis. Additional information can be found in Table S2.

1.6 Risk of Bias

In general, the articles were deemed to have a high risk of bias due to the absence of a gold standard for diagnosing and predicting cognitive impairment, and the failure to provide the outcomes in the majority of included articles. The quality assessment is presented in Figure 2.

1.7 Machine Learning Models and Predictors

In the 43 studies, 181 models were constructed with different ML models and predictors. However, 151 models were subjected to analysis, with accuracy being the primary outcome. These included 42 SVM models, 32 RF models, 12 KNN models, 10 NB models, and other supervised learnings. Also, we included 3 semi-supervised models that used SVM with PCA, and two neural network. The overall mean accuracy of 151 models was 74.54% (10.16%). Of the 99 models that reported AUCROC, the mean accuracy was 0.75 (0.12). Also, 128 models reported sensitivity

and specificity, with mean values of 71.38% (13.65%) and 77.18% (13.95%), respectively.

The study discovered that the predictors used to diagnose and predict cognitive impairment in PD included demographic, clinical, neuroimaging, QEEG features, biofluid biomarkers, and genetic features (Figure 3A). The top-used predictors were summarized in 151 models (Figure 3B). 125 out of 151 models used a single type of feature to diagnose, with a mean accuracy of 73.71% (10.15%). Nearly 17% of models used multiple types of features combined, with a higher accuracy of 78.57% (9.38%). In the single type of model, the predictors with the highest accuracy were neuroimaging features, achieving an accuracy of 78.86% (11.70%). Furthermore, 10 articles utilized longitudinal data to predict cognitive impairment or convertion, with the majority of predictors being neuroimaging features and biofluid markers.

1.8 Meta-analysis

Because only a few articles reported TP, TN, FP, FN, sensitivity, and specificity, 12 studies were included in the meta-analysis [20–31].In the event of multiple models from a single study, we prioritized the selection of the highest accuracy which was reported in 63% of the articles, to express optimal model performance. A total of 7 studies has reported outcomes in train set or validation set, and 5 studies have reported outcomes in both train set and validation set. Sensitivity and specificity forest plots for the train set and validation set respectively are presented in Figure 4, and for the single set in Figure S1. Figure 5 and Figure S2 display the SROC curve illustrating the summary point. No heterogeneity was found in the sensitivity and specificity values. Fagan plots were constructed to illustrate the pre-test and post-test probability of ML predicting cognitive impairment in PD patients (Figure S3 and S4). The Deeks' test found no significant publication bias for ML (Figure S5 and S6).

4 Discussion

1.9 Principal findings

This meta-analysis provided new results regarding the value of ML in the diagnosis of cognitive impairment in PD patients. The average accuracy of 151 models was 74.54%, and 128 models reported sensitivity and specificity, with mean values of 71.38% and 77.18%, respectively. There was no significant heterogeneity among the sensitivities and specificities reported by the studies included. Moreover, many of the studies had major methodological weaknesses. More than 50% of the included studies in this systematic review and meta-analysis had a high risk of bias for patient selection, leading to applicability concerns.

1.10 Machine learning models and Predictors

A recent review divided ML into traditional ML (including supervised ML and unsupervised ML), and ML methods based on neural networks (such as convolutional neural networks and recurrent neural networks) [32]. 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. Firstly, 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. Secondly, 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 [33].

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, semi-supervised learning also showed great potential [34]. Che et al. combined SVM with PCA to successfully extract 6 features from 32 predictors, improving the performance over the simple SVM model [35]. Other models use unsupervised learning to cluster this unlabelled data and then employ supervised learning for longitudinal prediction [36]. Further research and discussion are needed to assess the accuracy and effectiveness of these different models.

We found that cognitive tests such as MoCA and MMSE 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 functional MRI (fMRI) was widely used to diagnose and predict disease progression, including FDP-PET which can reflect the metabolism in the brain and DTI 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 α-synuclein detected in the CSF 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 (GWAS) 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 [37], while others have

concluded that the models incorporating multiple types of features are more effective [38–40]. 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 study on predicting cognitive impairment and conversion in PD [22,41–46]. We discover that neuroimaging features and biofluid markers are robust predictors. With the advancement of imaging technology, such as fMRI and Magnetic Resonance Spectroscopy (MRS), and their non-invasive detection capabilities, neuroimaging has become mainstream today. Neural networks and deep learning are especially focused on applications in imaging. Numerous research has suggested that biomarkers such as Aβ, α-synuclein, 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 three-component system (SynNeurGe). It includes the harmful α-synuclein (S) accumulated in tissues or cerebrospinal fluid (CSF), 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 [47]. 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.

1.11 Strength and Limitations

A large number of studies have investigated the diagnostic accuracy of ML to diagnose cognitive impairment in PD patients. With the increasing number of deep

learning and detection approaches, a more comprehensive exploration of ML was conducted. We employed bivariate analysis to show more robust results. However, there are still several limitations in this review. Primarily, we excluded non-published data in the analysis. Due to heterogeneity in intervention methods and outcome measures between studies, it was not possible to synthesize all articles included using meta-analysis. Secondly, several studies used Parkinson's Disease Progression Markers Initiative (PPMI) datasets, and we lack all the results. Consequently, it weas not possible to assess the heterogeneity exactly. Further investigation is required to determine the most effective model or predictors for identifying PD patients who suffer from cognitive impairment, predicting the disease progression (including conversion to dementia), and providing personalized treatment options.

5 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.

6 Acknowledgement

The authors thank all subjects for their participation in the study.

7 Conflict of Interest

No potential conflict of interest was reported by the authors.

Reference

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*, 344–381, doi:10.1016/S1474-4422(24)00038-3.
2. Jankovic, J.; Tan, E.K. Parkinson's Disease: Etiopathogenesis and Treatment. *J. Neurol. Neurosurg. Psychiatry* **2020**, *91*, 795–808, doi:10.1136/jnnp-2019-322338.
3. Schapira, A.H.V.; Chaudhuri, K.R.; Jenner, P. Non-Motor Features of Parkinson Disease. *Nat. Rev. Neurosci.* **2017**, *18*, 435–450, doi:10.1038/nrn.2017.62.
4. Hanagasi, H.A.; Tufekcioglu, Z.; Emre, M. Dementia in Parkinson's Disease. *J. Neurol. Sci.* **2017**, *374*, 26–31, doi:10.1016/j.jns.2017.01.012.
5. Aarsland, D.; Batzu, L.; Halliday, G.M.; Geurtsen, G.J.; Ballard, C.; Ray Chaudhuri, K.; Weintraub, D. Parkinson Disease-Associated Cognitive Impairment. *Nat. Rev. Dis. Primer* **2021**, *7*, 47, doi:10.1038/s41572-021-00280-3.
6. Litvan, I.; Goldman, J.G.; Tröster, A.I.; Schmand, B.A.; Weintraub, D.; Petersen, R.C.; Mollenhauer, B.; Adler, C.H.; Marder, K.; Williams-Gray, C.H.; et al. Diagnostic Criteria for Mild Cognitive Impairment in Parkinson's Disease: *Movement* Disorder Society Task Force Guidelines. *Mov. Disord.* **2012**, *27*, 349–356, doi:10.1002/mds.24893.
7. Emre, M.; Aarsland, D.; Brown, R.; Burn, D.J.; Duyckaerts, C.; Mizuno, Y.; Broe, G.A.; Cummings, J.; Dickson, D.W.; Gauthier, S.; et al. Clinical Diagnostic Criteria for Dementia Associated with Parkinson's Disease. *Mov. Disord.* **2007**, *22*, 1689–1707, doi:10.1002/mds.21507.
8. Saredakis, D.; Collins-Praino, L.E.; Gutteridge, D.S.; Stephan, B.C.M.; Keage, H.A.D. 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.
9. Delgado-Alvarado, M.; Gago, B.; Navalpotro-Gomez, I.; Jiménez-Urbieta, H.; Rodriguez-Oroz, M.C. Biomarkers for Dementia and Mild Cognitive Impairment in Parkinson's Disease. *Mov. Disord.* **2016**, *31*, 861–881, doi:10.1002/mds.26662.
10. Deo, R.C. Machine Learning in Medicine. *Circulation* **2015**, *132*, 1920–1930, doi:10.1161/CIRCULATIONAHA.115.001593.
11. Sarker, I.H. Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Comput. Sci.* **2021**, *2*, 160, doi:10.1007/s42979-021-

00592-x.

12. Yu, E.; Larivière, R.; Thomas, R.A.; Liu, L.; Senkevich, K.; Rahayel, S.; Trempe, J.-F.; Fon, E.A.; Gan-Or, Z. Machine Learning Nominates the Inositol Pathway and Novel Genes in Parkinson's Disease. *Brain* **2024**, *147*, 887–899, doi:10.1093/brain/awad345.

13. Mallo, S.C.; Valladares-Rodriguez, S.; Facal, D.; Lojo-Seoane, C.; Fernández-Iglesias, M.J.; Pereiro, A.X. Neuropsychiatric Symptoms as Predictors of Conversion from MCI to Dementia: A Machine Learning Approach. *Int. Psychogeriatr.* **2020**, *32*, 381–392, doi:10.1017/S1041610219001030.

14. Fleuren, L.M.; Klausch, T.L.T.; Zwager, C.L.; Schoonmade, L.J.; Guo, T.; Roggeveen, L.F.; Swart, E.L.; Girbes, A.R.J.; Thoral, P.; Ercole, A.; et al. Machine Learning for the Prediction of Sepsis: A Systematic Review and Meta-Analysis of Diagnostic Test Accuracy. *Intensive Care Med.* **2020**, *46*, 383–400, doi:10.1007/s00134-019-05872-y.

15. Whiting, P.F.; Rutjes, A.W.S.; Westwood, M.E.; Mallett, S.; Deeks, J.J.; Reitsma, J.B.; Leeflang, M.M.G.; Sterne, J.A.C.; Bossuyt, P.M.M.; QUADAS-2 Group QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. *Ann. Intern. Med.* **2011**, *155*, 529–536, doi:10.7326/0003-4819-155-8-201110180-00009.

16. Reitsma, J.B.; Glas, A.S.; Rutjes, A.W.S.; Scholten, R.J.P.M.; Bossuyt, P.M.; Zwinderman, A.H. Bivariate Analysis of Sensitivity and Specificity Produces Informative Summary Measures in Diagnostic Reviews. *J. Clin. Epidemiol.* **2005**, *58*, 982–990, doi:10.1016/j.jclinepi.2005.02.022.

17. Harbord, R.M.; Deeks, J.J.; Egger, M.; Whiting, P.; Sterne, J.A.C. A Unification of Models for Meta-Analysis of Diagnostic Accuracy Studies. *Biostatistics* **2007**, *8*, 239–251, doi:10.1093/biostatistics/kxl004.

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*, doi:10.3310/hta9120.

19. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. *BMJ* **2021**, n71, doi:10.1136/bmj.n71.

20. Abos, A.; Baggio, H.C.; Segura, B.; Garcia-Diaz, A.I.; Compta, Y.; Marti, M.J.; Valldeoriola, F.; Junque, C. Discriminating Cognitive Status in Parkinson's Disease through Functional Connectomics and Machine Learning. *Sci Rep* **2017**, *7*, doi:10.1038/srep45347.

21. Amboni, M.; Ricciardi, C.; Adamo, S.; Nicolai, E.; Volzone, A.; Erro, R.; Cuoco,

S.; Cesarelli, G.; Basso, L.; D'Addio, G.; et al. Machine Learning Can Predict Mild Cognitive Impairment in Parkinson's Disease. *Front. Neurol.* **2022**, *13*, doi:10.3389/fneur.2022.1010147.

22. Booth, S.; Park, K.W.; Lee, C.S.; Ko, J.H. Predicting Cognitive Decline in Parkinson's Disease Using FDG-PET-Based Supervised Learning. *J. Clin. Invest.* **2022**, *132*, e157074, doi:10.1172/JCI157074.

23. Zhang, J.; Li, Y.; Gao, Y.; Hu, J.; Huang, B.; Rong, S.; Chen, J.; Zhang, Y.; Wang, L.; Feng, 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*, doi:10.1016/j.jns.2020.117077.

24. Zhang, J.; Gao, Y.; He, X.; Feng, S.; Hu, J.; Zhang, Q.; Zhao, J.; Huang, Z.; Wang, L.; Ma, G.; et al. Identifying Parkinson's Disease with Mild Cognitive Impairment by Using Combined MR Imaging and Electroencephalogram. *Eur Radiol* **2021**, *31*, 7386–7394, doi:10.1007/s00330-020-07575-1.

25. Cengiz, S.; Arslan, D.B.; Kicik, A.; Erdogdu, E.; Yildirim, M.; Hatay, G.H.; Tufekcioglu, Z.; Ulug, A.M.; Bilgic, B.; Hanagasi, H.; et al. Identification of Metabolic Correlates of Mild Cognitive Impairment in Parkinson's Disease Using Magnetic Resonance Spectroscopic Imaging and Machine Learning. *Magn. Reson. Mater. Phys. Biol. Med.* **2022**, *35*, 997–1008, doi:10.1007/s10334-022-01030-6.

26. Chen, B.; Xu, M.; Yu, H.; He, J.; Li, Y.; Song, D.; Fan, G.G. 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*, doi:10.1186/s12967-023-04158-8.

27. García, A.M.; Arias-Vergara, T.; C. Vasquez-Correa, J.; Nöth, E.; Schuster, M.; Welch, A.E.; Bocanegra, Y.; Baena, A.; Orozco-Arroyave, J.R. Cognitive Determinants of Dysarthria in Parkinson's Disease: An Automated Machine Learning Approach. *Mov. Disord.* **2021**, *36*, 2862–2873, doi:10.1002/mds.28751.

28. Lin, H.; Liu, Z.; Yan, W.; Zhang, D.; Liu, J.; Xu, B.; Li, W.; Zhang, Q.; Cai, X. Brain Connectivity Markers in Advanced Parkinson's Disease for Predicting Mild Cognitive Impairment. *Eur. Radiol.* **2021**, *31*, 9324–9334, doi:10.1007/s00330-021-08086-3.

29. Morales, D.A.; Vives-Gilabert, Y.; Gomez-Anson, B.; Bengoetxea, E.; Larranaga, P.; Bielza, C.; Pagonabarraga, J.; Kulisevsky, J.; Corcuera-Solano, I.; Delfino, M. Predicting Dementia Development in Parkinson's Disease Using Bayesian Network Classifiers. *Psychiatry Res.-Neuroimaging* **2013**, *213*, 92–98, doi:10.1016/j.pscychresns.2012.06.001.

30. Russo, M.; Amboni, M.; Barone, P.; Pellecchia, M.T.; Romano, M.; Ricciardi,

C.; Amato, F. Identification of a Gait Pattern for Detecting Mild Cognitive Impairment in Parkinson's Disease. *Sensors* **2023**, *23*, doi:10.3390/s23041985.

31. 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. *Park. Relat Disord* **2023**, *115*, doi:10.1016/j.parkreldis.2023.105802.
32. Greener, J.G.; Kandathil, S.M.; Moffat, L.; Jones, D.T. A Guide to Machine Learning for Biologists. *Nat. Rev. Mol. Cell Biol.* **2022**, *23*, 40–55, doi:10.1038/s41580-021-00407-0.
33. Barredo Arrieta, A.; Díaz-Rodríguez, N.; Del Ser, J.; Bennetot, A.; Tabik, S.; Barbado, A.; Garcia, S.; Gil-Lopez, S.; Molina, D.; Benjamins, R.; et al. Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI. *Inf. Fusion* **2020**, *58*, 82–115, doi:10.1016/j.inffus.2019.12.012.
34. Lundberg, S.M.; Lee, S.-I. A Unified Approach to Interpreting Model Predictions. In Proceedings of the Proceedings of the 31st International Conference on Neural Information Processing Systems; Curran Associates Inc.: Red Hook, NY, USA, December 4 2017; pp. 4768–4777.
35. Chen, P.-H.; Hou, T.-Y.; Cheng, F.-Y.; Shaw, J.-S. Prediction of Cognitive Degeneration in Parkinson's Disease Patients Using a Machine Learning Method. *Brain Sci.* **2022**, *12*, 1048, doi:10.3390/brainsci12081048.
36. Kawabata, K.; Watanabe, H.; Hara, K.; Bagarinao, E.; Yoneyama, N.; Ogura, A.; Imai, K.; Masuda, M.; Yokoi, T.; Ohdake, R.; 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*, 688–700, doi:10.1007/s00415-018-8755-5.
37. Chaturvedi, M.; Bogaarts, J.G.; Kozak (Cozac), V.V.; Hatz, F.; Gschwandtner, U.; Meyer, A.; Fuhr, P.; Roth, V. 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*, 1937–1944, doi:10.1016/j.clinph.2019.07.017.
38. Amboni, M.; Ricciardi, C.; Adamo, S.; Nicolai, E.; Volzone, A.; Erro, R.; Cuoco, S.; Cesarelli, G.; Basso, L.; D'Addio, G.; et al. Machine Learning Can Predict Mild Cognitive Impairment in Parkinson's Disease. *Front. Neurol.* **2022**, *13*, 1010147, doi:10.3389/fneur.2022.1010147.
39. Chen, B.; Xu, M.; Yu, H.; He, J.; Li, Y.; Song, D.; Fan, G.G. 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*, 310, doi:10.1186/s12967-023-04158-8.

40. Zhang, J.; Gao, Y.; He, X.; Feng, S.; Hu, J.; Zhang, Q.; Zhao, J.; Huang, Z.; Wang, L.; Ma, G.; et al. Identifying Parkinson's Disease with Mild Cognitive Impairment by Using Combined MR Imaging and Electroencephalogram. *Eur. Radiol.* **2021**, *31*, 7386–7394, doi:10.1007/s00330-020-07575-1.

41. Harvey, J.; Reijnders, R.A.; Cavill, R.; Duits, A.; Kohler, S.; Eijssen, L.; Rutten, B.P.F.; Shireby, G.; Torkamani, A.; Creese, B.; et al. Machine Learning-Based Prediction of Cognitive Outcomes in de Novo Parkinson's Disease. *Npj Park. Dis.* **2022**, *8*, doi:10.1038/s41531-022-00409-5.

42. Hogue, O.; Fernandez, H.H.; Floden, D.P. Predicting Early Cognitive Decline in Newly-Diagnosed Parkinson's Patients: A Practical Model. *Parkinsonism Relat. Disord.* **2018**, *56*, 70–75, doi:10.1016/j.parkreldis.2018.06.031.

43. Huang, X.; He, Q.; Ruan, X.; Li, Y.; Kuang, Z.; Wang, M.; Guo, R.; Bu, S.; Wang, Z.; Yu, S.; et al. Structural Connectivity from DTI to Predict Mild Cognitive Impairment in de Novo Parkinson's Disease. *Neuroimage-Clin.* **2024**, *41*, doi:10.1016/j.nicl.2023.103548.

44. Liu, G.; Locascio, J.J.; Corvol, J.-C.; Boot, B.; Liao, Z.; Page, K.; Franco, D.; Burke, K.; Jansen, I.E.; Trisini-Lipsanopoulos, A.; et al. Prediction of Cognition in Parkinson's Disease with a Clinical-Genetic Score: A Longitudinal Analysis of Nine Cohorts. *Lancet Neurol.* **2017**, *16*, 620–629, doi:10.1016/S1474-4422(17)30122-9.

45. Shin, N.-Y.; Bang, M.; Yoo, S.-W.; Kim, J.-S.; Yun, E.; Yoon, U.; Han, K.; Ahn, K.J.; Lee, S.-K. Cortical Thickness from MRI to Predict Conversion from Mild Cognitive Impairment to Dementia in Parkinson Disease: A Machine Learning-Based Model. *Radiology* **2021**, *300*, 390–399, doi:10.1148/radiol.2021203383.

46. 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.

47. Höglinger, G.U.; Adler, C.H.; Berg, D.; Klein, C.; Outeiro, T.F.; Poewe, W.; Postuma, R.; Stoessl, A.J.; Lang, A.E. A Biological Classification of Parkinson's Disease: The SynNeurGe Research Diagnostic Criteria. *Lancet Neurol.* **2024**, *23*, 191–204, doi:10.1016/S1474-4422(23)00404-0.

Figure Legend

Figure 1. Flowchart illustrating the process of searching for, screening, and inclusion.

Figure 2. Quality assessment results of the selected studies.
Figure 3. Predictors used in ML models.
Figure 4. Forest plot of sensitivity and specificity.
Figure 5. Summary Receiver-Operating Characteristic curve of the sensitivity and specificity.

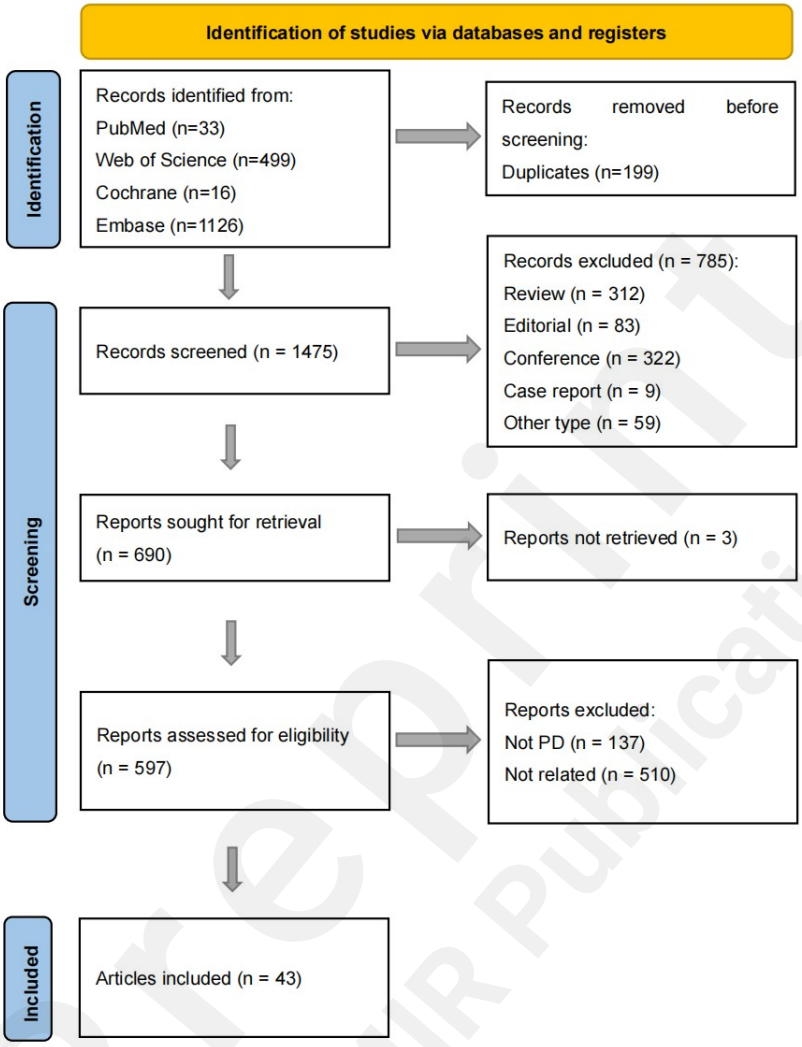


Figure 1. Flowchart illustrating the process of searching for screening and inclusion.

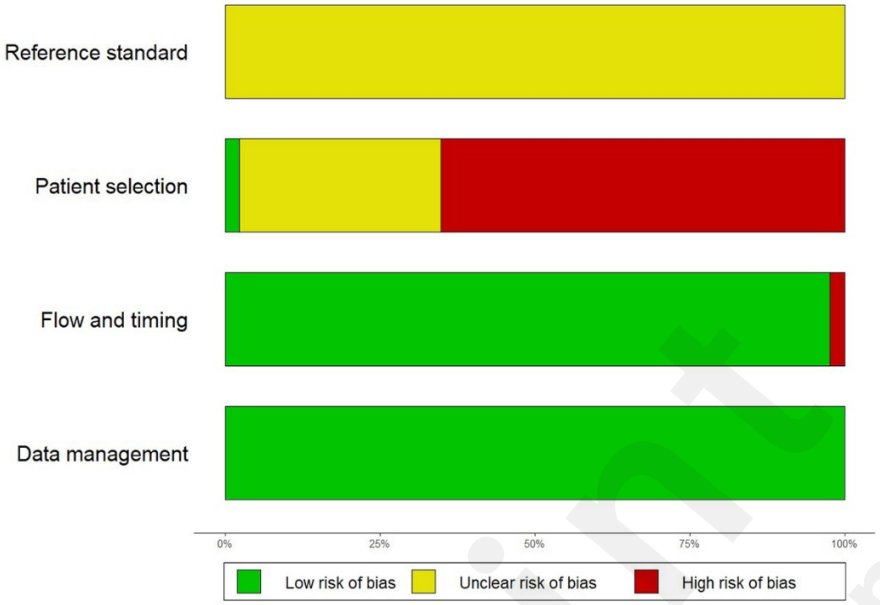


Figure 2. Quality assessment results of the selected studies.

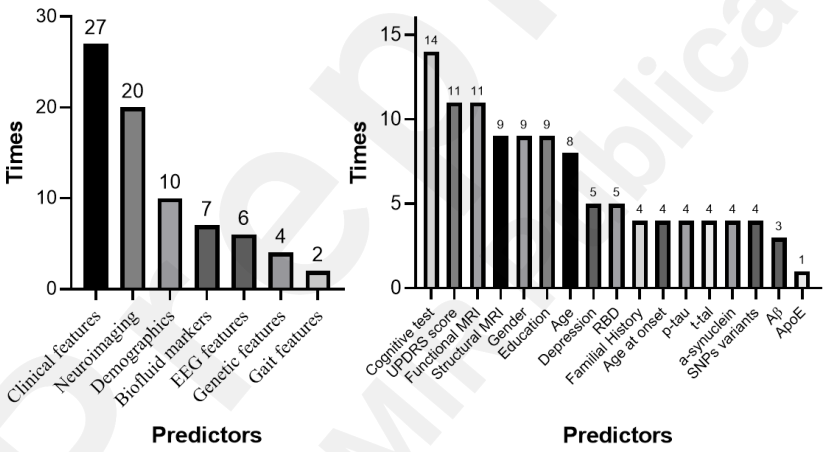


Figure 3. Predictors used in ML models. (A) shows each type of predictor put into models; (B) shows the detailed predictors which are classified by type.

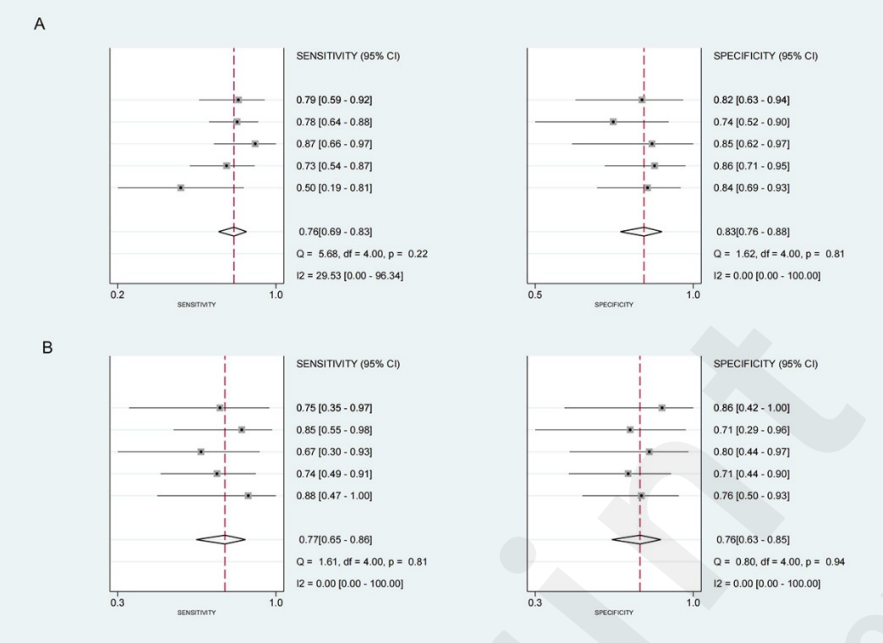


Figure 4. Forest plot of sensitivity and specificity. (A) Forest plot of sensitivity and specificity for train set; (B) Forest plot of sensitivity and specificity for validation set.

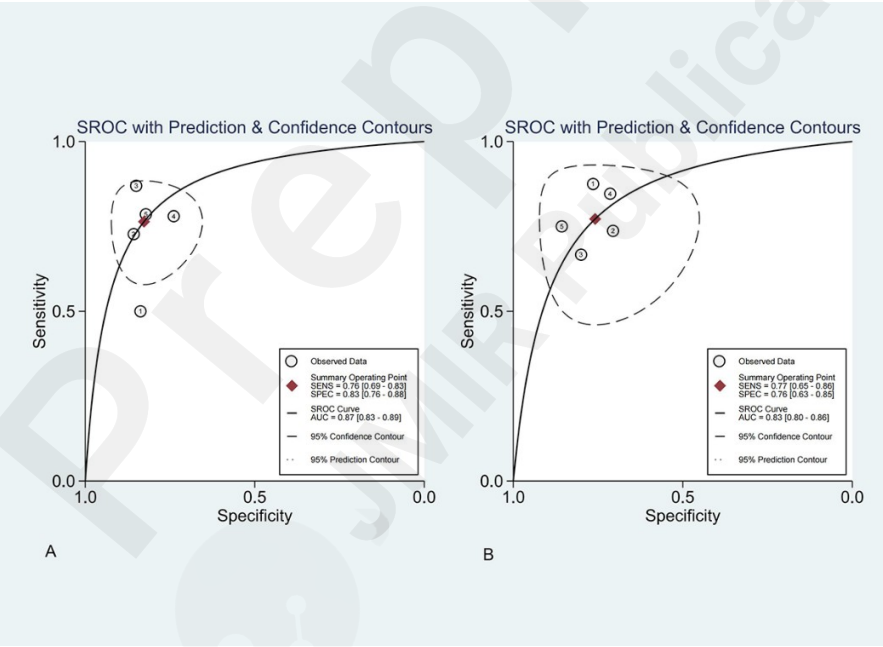
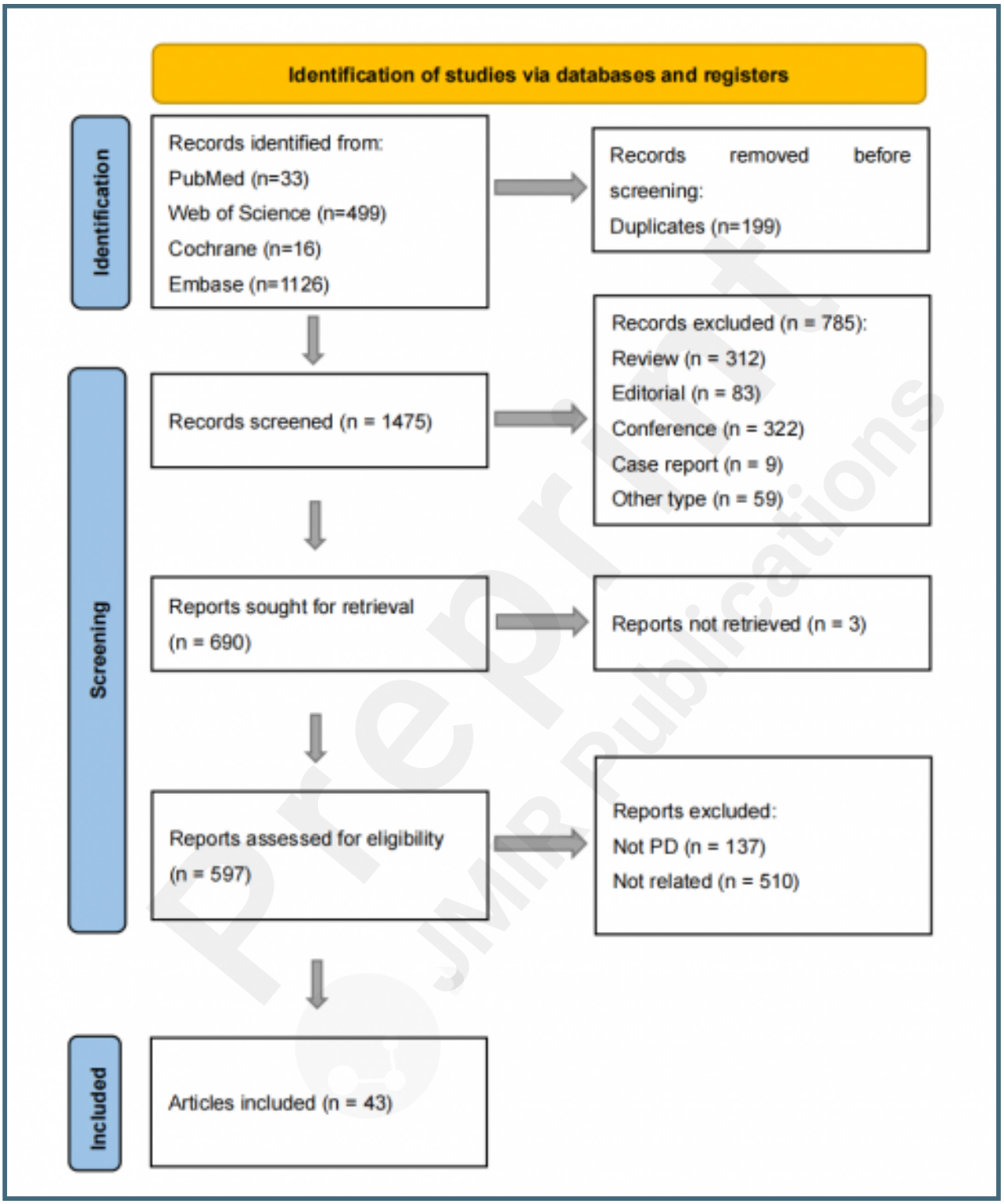


Figure 5. Summary Receiver-Operating Characteristic curve of the sensitivity and specificity. (A) For the train set; (B) For the validation set.

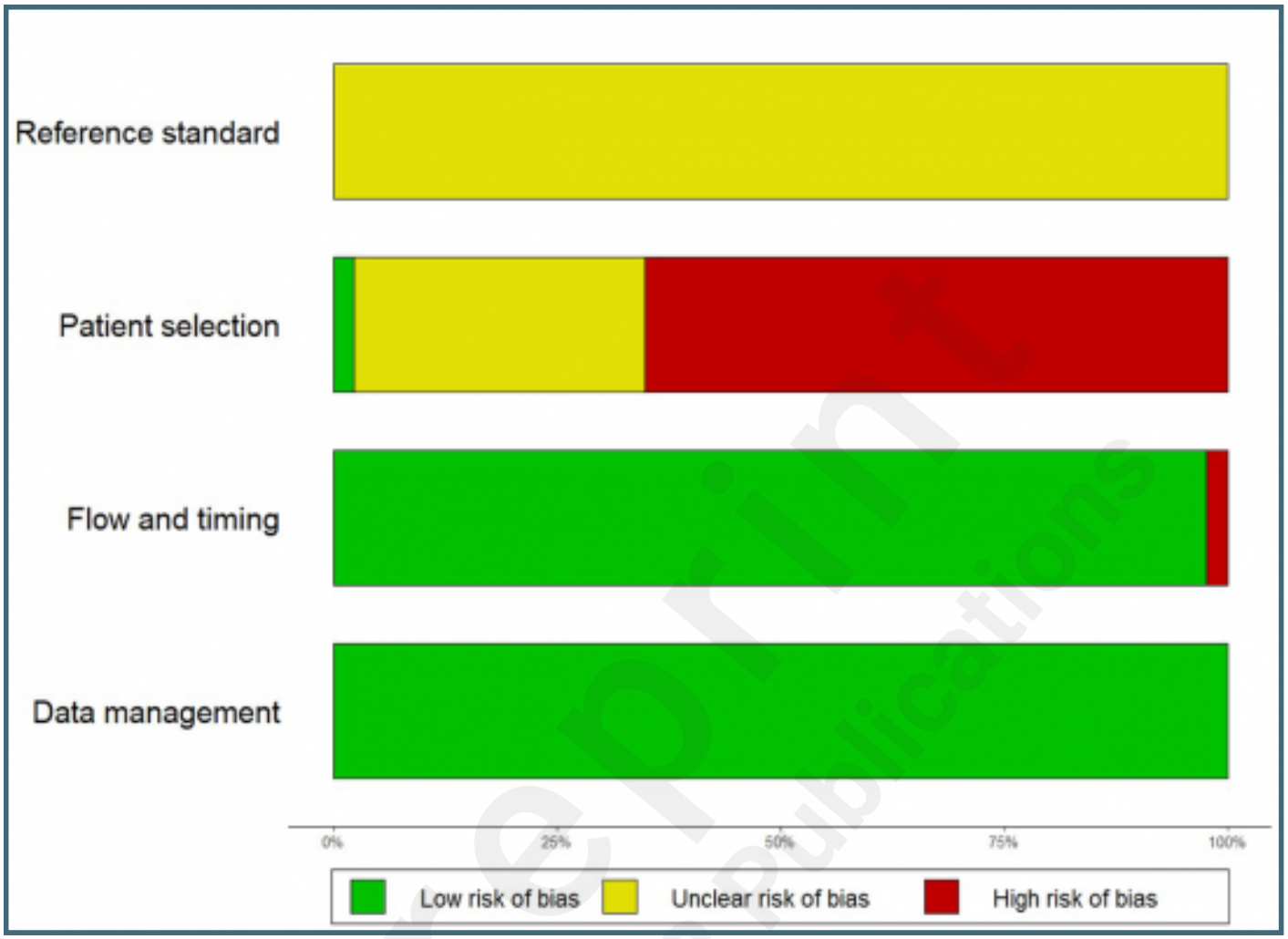
Supplementary Files

Figures

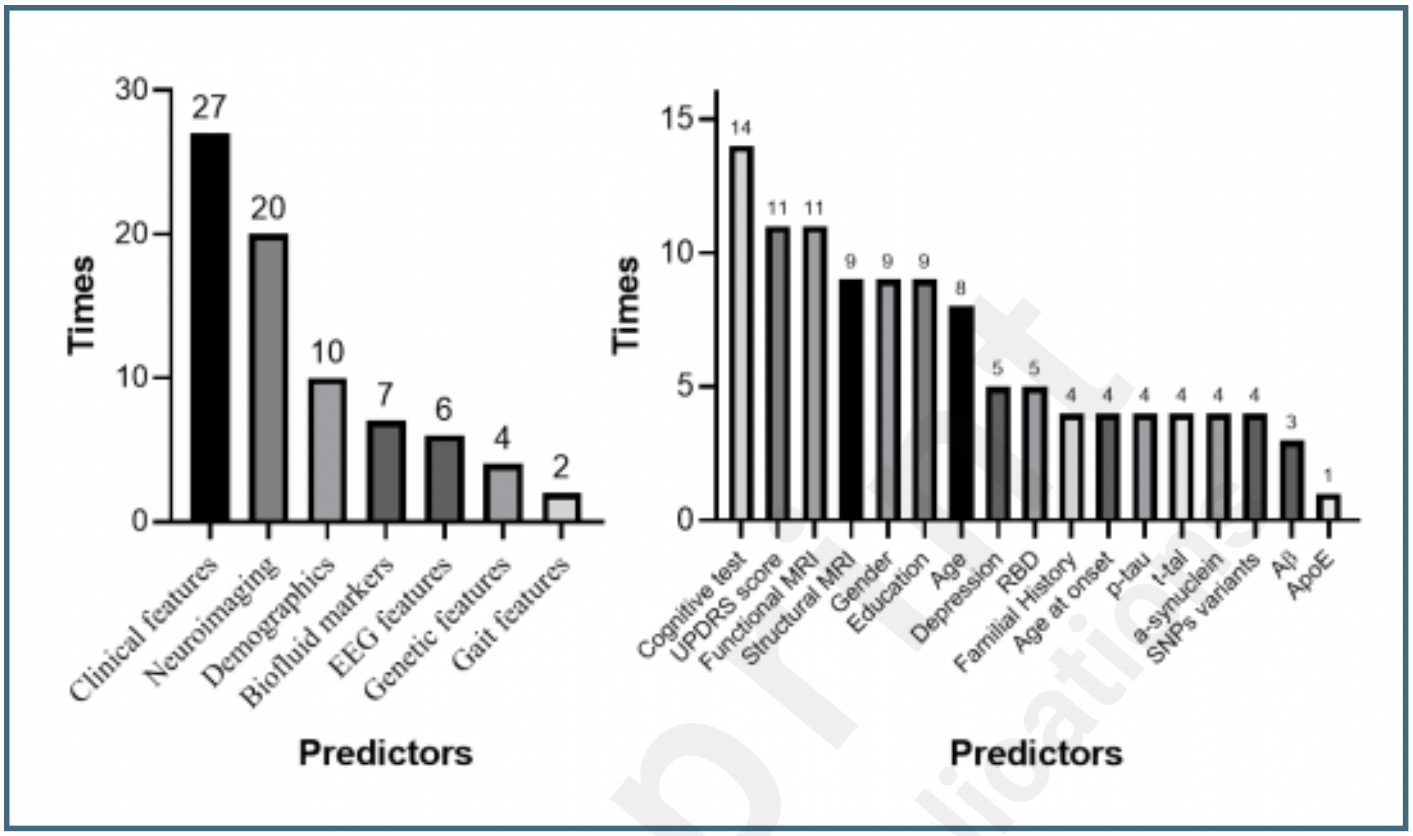
Flowchart illustrating the process of searching for, screening, and inclusion.



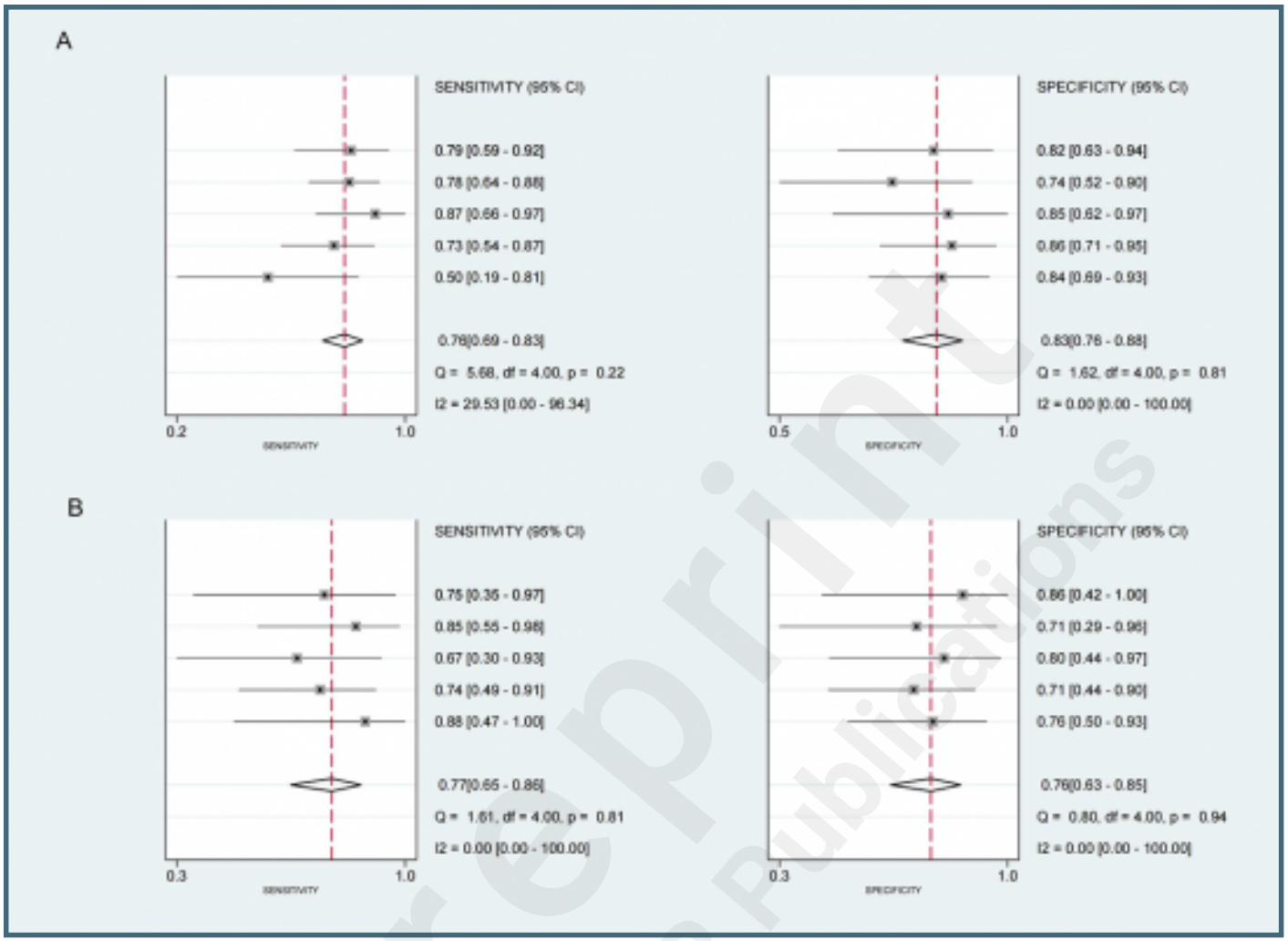
Quality assessment results of the selected studies.



Predictors used in ML models.



Forest plot of sensitivity and specificity.



Summary Receiver-Operating Characteristic curve of the sensitivity and specificity.