BMC Medical Imaging

RESEARCH Open Access

Machine learning-based brain magnetic resonance imaging radiomics for identifying rapid eye movement sleep behavior disorder in Parkinson's disease patients

Yandong lian[1†], Yibin Xu[2†], Linlin Hu[2], Yuguo Wei[3] and Zhaoge Wang[4*]

Abstract

Background Traditional clinical diagnostic methods of rapid eye movement sleep behavior disorder (RBD) have certain limitations, especially in the early stages. This study aims to develop and validate an magnetic resonance imaging (MRI) radiomics-based machine learning classifier to accurately detect RBD patients with Parkinson's disease (PD).

Methods Data from 183 subjects, including 63 PD patients with RBD, sourced from the PPMI database were utilized in this study. The data were randomly divided into training (70%) and testing (30%) sets. Quantitative radiomic features of white matter, gray matter, and cerebrospinal fluid were extracted from whole-brain structural MRI images. Feature reduction was performed on the training set data to construct radiomics signatures. Additionally, multi-factor logistic regression analysis identified clinical predictors associated with PD-RBD, and these clinical features were integrated with the radiomics signatures to develop predictive models using various machine learning algorithms. The model exhibiting the best performance was selected, and receiver operating characteristic (ROC) curves were used to evaluate its performance in both the training and testing sets. Furthermore, based on the optimal cut-off value of the model, subjects were categorized into low- and high-risk groups, and differences in the actual number of RBD patients between the two sets were compared to assess the clinical effectiveness of the model.

Results The radiomics signatures achieved areas under the curve (AUC) of 0.754 and 0.707 in the training and testing sets, respectively. Multi-factor logistic regression analysis revealed that postural instability was an independent predictor of PD-RBD. The random forest model, which integrated radiomics signatures with postural instability, demonstrated superior performance in predicting PD-RBD. Specifically, its AUCs in the training and testing sets were 0.917 and 0.882, with sensitivities of 0.933 and 0.889, and specificities of 0.786 and 0.722, respectively. Based on the optimal cut-off value of 0.3772, significant differences in the actual number of PD-RBD patients were observed between low-risk and high-risk groups in both the training and testing sets ($P < 0.05$).

†Yandong lian and Yibin Xu authors contributed equally to this work.

*Correspondence:
Zhaoge Wang
11035070@qq.com

Full list of author information is available at the end of the article

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Conclusion MRI-based radiomic signatures have the potential to serve as biomarkers for PD-RBD. The random forest model, which integrates radiomic signatures with postural instability, and shows improved performance in identifying PD-RBD. This approach offers valuable insights for prognostic evaluation and preventive treatment strategies.

Keywords Magnetic resonance imaging, Radiomics, Rapid eye movement sleep behavior disorder, Machine learning, Biomarkers

Introduction

Parkinson's disease (PD) is one of the most common movement disorders, primarily affecting motor control and presenting with symptoms such as tremors, rigidity, and bradykinesia [1]. As the disease progresses, non-motor symptoms become increasingly prominent, with Rapid Eye Movement Sleep Behavior Disorder (RBD) being one of the most common and troublesome non-motor symptoms for both patients and clinicians [2]. RBD is a significant sleep abnormality in PD patients, characterized by the absence of normal muscle paralysis during rapid eye movement (REM) sleep [3]. RBD not only impacts the patient's quality of life but is also considered an early indicator of PD, often emerging before the onset of motor symptoms [4]. Therefore, early recognition and diagnosis of RBD are critical for the clinical management of PD.

Currently, the diagnosis of RBD primarily relies on clinical symptoms, patient self-reporting, and polysomnography (PSG) [5]. However, traditional clinical diagnostic methods have certain limitations, especially in the early stages, when patients may not consciously notice changes in sleep behavior, or clinicians may fail to detect subtle changes in sleep behavior abnormalities [6]. While PSG provides more accurate diagnostic information, its application is restricted by various factors, including expensive equipment, complex procedures, and patient discomfort [7]. Thus, the development of a simpler, non-invasive, and efficient diagnostic approach has become a research focus in the field of neuroscience. With the advancement of medical imaging technology, Magnetic Resonance Imaging (MRI) has become an indispensable tool in the diagnosis of brain disorders [8]. Particularly, the rise of radiomics has enabled the extraction of a large number of high-dimensional features from MRI images, allowing for the creation of mathematical models that can accurately classify and predict diseases, offering clinicians a more objective and quantitative diagnostic basis [9]. Recent progress in radiomics has shown significant advancements in various neurodegenerative diseases, such as Alzheimer's Disease (AD) and stroke [10, 11]. However, radiomics research on Parkinson's disease and its complications, such as RBD, remains relatively sparse, and its application in this field is still in the exploratory stage.

Machine Learning (ML), as a vital branch of artificial intelligence, has been widely applied in medical imaging analysis [12]. By training algorithms to learn underlying patterns from large datasets, machine learning can extract valuable information from complex imaging data and classify, predict, and evaluate diseases based on this information [13]. In recent years, research combining radiomics and machine learning has become a hot topic in the field of medical imaging, especially in PD-related studies, such as the work by Panahi et al, which demonstrated the utility of multi-modality radiomics (T1-weighted and DTI) for distinguishing PD motor subtypes in early-stage patients [14]. Notably, recent advances have explored ML-based detection of RBD using presynaptic dopaminergic imaging [15]. However, such approaches face significant limitations, including high costs, reliance on radioactive tracers, and limited accessibility in resource-constrained settings, which hinder their utility for population-level screening. In contrast, structural MRI-based radiomics offers a cost-effective, non-invasive alternative, leveraging widely available imaging protocols to capture neurodegenerative patterns. Our previous studies have demonstrated that a machine learning classifier built from whole-brain MRI radiomics can accurately identify cognitive decline [16]. Despite the growing use of radiomics in neurodegenerative research, there is a notable gap in studies specifically targeting the detection of RBD in PD patients using MRI-based radiomic features integrated with machine learning. Therefore, this study aims to fill this gap by developing a hybrid model, and we hypothesize that a machine learning classifier based on MRI radiomics can also achieve precise detection of RBD in PD patients.

This study aims to develop a machine learning classifier based on MRI radiomics for the high-accuracy detection of RBD in PD patients. By conducting this research, we hope to provide a more convenient and efficient method for detecting RBD in PD patients and offer a powerful decision-support tool for clinicians.

Materials and methods

Patient information

This study was conducted in accordance with the guidelines of the Declaration of Helsinki. This study collected patient information from the Parkinson's Progression Markers Initiative (PPMI) database (https://www.ppmi-info.org/), a global, large-scale, multi-center clinical research platform designed to systematically gather multi-dimensional data from PD patients and control

groups, including clinical features, imaging data, genomic information, laboratory results, and biomarkers. In this study, we focused on PD patients with RBD from the PPMI database. Patient selection was based on the following inclusion and exclusion criteria. Inclusion criteria: all patients met the clinical diagnostic criteria for Parkinson's disease and had clear diagnostic records in the PPMI database; the presence or absence of RBD was confirmed using validated clinical assessment tools, including the Rapid Eye Movement Sleep Behavior Disorder Screening Questionnaire (RBDSQ) and polysomnography (PSG). Diagnosis of RBD adhered to the International Classification of Sleep Disorders; patients had completed MRI scans and provided complete clinical data and imaging information. Exclusion criteria: patients with other neurodegenerative diseases or known brain lesions, such as Alzheimer's disease, multiple system atrophy, or Huntington's disease; patients whose MRI data could not be extracted due to technical issues preventing the extraction of radiomic features. Based on these criteria, 183 PD patients were ultimately included, of whom 63 had RBD, and remaining patients are defined as PD negative control(–). The cases were randomly divided into training and testing sets in a 7:3 ratio, with the training set used for model development and the testing set used to validate the model's reliability.

Image data preprocessing

All experimental data were acquired using a 1.5 T or 3.0 T MRI system. To minimize the impact of scanning parameters on feature extraction, structural T1-weighted images (T1WI) were preprocessed prior to analysis. First, the image resolution was adjusted to 1×1x1 mm^3, and the grayscale intensity was normalized to 32 levels[17]. The preprocessed images were then imported into the SPM12 software (version V2.5.5) on the MATLAB platform for automated segmentation, where regions of the brain, including white matter, gray matter, and cerebrospinal fluid (CSF), were extracted [18]. Next, the segmented results were further refined manually using ITK-SNAP software (version 4.2.2, http://www.itksnap.org/pmwiki/pmwiki.php) [19]. The manual corrections primarily included: 1) removal of non-brain tissues, brainstem, and cerebellum regions; and 2) correction of any inaccuracies in the automated segmentation. All manual corrections were performed by two experienced neuroradiologists independently, who remained blinded to the clinical information of the patients during the process. After completing the manual adjustments, the final segmented regions (white matter, gray matter, and CSF were used as masks and imported into the Pyradiomics software(version V3.0.1) for feature extraction [20]. This refined process ensures the accuracy and consistency of the image segmentation, thus enhancing the quality of the features extracted for subsequent analysis.

Radiomic feature extraction

In this study, radiomic features were extracted from the regions of interest (ROI), including white matter, gray matter, and CSF, using the PyRadiomics platform [21]. The extracted features were classified into two major categories: original image features and filter-based features. Original image features included 14 shape features, 18 first-order statistical features, and various texture features, such as 22 from the gray-level co-occurrence matrix (GLCM), 16 from the gray-level run-length matrix (GLRLM), 16 from the gray-level size zone matrix (GLSZM), and 14 from the gray-level dependence matrix (GLDM). Filter-based features included wavelet transforms and Laplacian transforms, and wavelet decompositions were performed using 3D Haar transforms with one decomposition level, generating eight frequency sub-bands. A Laplacian filter with a 3×3x3 kernel and standardized coefficients (central: – 6, adjacent: 1) was applied to enhance edge-related features. For each brain region (white matter, gray matter, and CSF), a total of 1,132 features were extracted, yielding a final set of 3,396 radiomic features, with detailed feature information provided in Table S1.

Feature dimensionality reduction

To prevent potential loss of interactive patterns that may arise from processing features of individual regions in isolation, this study employed a pre-fusion strategy, integrating radiomic features from white matter, gray matter, and cerebrospinal fluid prior to dimensionality reduction, thereby comprehensively assessing cross-regional interactions critical for PD-RBD detection.To ensure the robustness of the features, the spearman correlation coefficients were computed for all radiomic features across 183 subjects between the two neuroradiologists segmentations. Features with coefficients greater than 0.8 were retained, and their average values were used for model building. This process eliminated features sensitive to manual refinement variability, ensuring data reliability and included in subsequent analyses [22]. Given the high dimensionality of the data, which may lead to model overfitting, feature selection and dimensionality reduction were performed on the feature sets extracted from the training set. Specifically, we applied two methods for dimensionality reduction: Maximum Relevance Minimum Redundancy (mRMR) and Least Absolute Shrinkage and Selection Operator (LASSO). The mRMR method selects the most representative features by maximizing their relevance to the target variable while minimizing redundancy between features. In this study, the mRMR algorithm was applied to select 30 top features

(k = 30) based on mutual information maximization, with parameter tuning performed via 5-fold cross-validation. This approach ensured a balance between relevance to PD-RBD and minimal inter-feature redundancy. LASSO regression, which incorporates an L1 regularization term, effectively eliminates features with minimal contribution to classification performance. The LASSO regularization parameter (λ) was optimized using 5-fold cross-validation on the training set. The detailed steps for dimensionality reduction are provided in the supplementary materials.

Construction of the joint model

After dimensionality reduction, radiomic features were used to construct radiomic signatures through logistic regression, and a score for each case, termed Rad-score, was calculated to reflect the probability of RBD occurrence. The classification accuracy of the Rad-score was evaluated through the area under the receiver operating characteristic (ROC) curve (AUC) in both the training and testing sets, and calibration curves were used to check for overfitting. To further build the combined prediction model, multivariate logistic regression was used in the training set to select independent predictive factors based on clinical features and Rad-score. Subsequently, five machine learning algorithms such as Support Vector Machine (SVM), Bayesian Classifier (Bayes), k-Nearest Neighbors (KNN), Decision Tree, and Random Forest were applied to construct prediction models based on independent predictive factor. The model training followed a two-layer nested cross-validation process. In the outer loop, the training set was randomly and stratified divided into training and testing subsets, repeated 200 times to assess the model's classification performance; in the inner loop, 5-fold cross-validation was used, and GridSearchCV was employed to automatically optimize hyperparameters [23]. Each split generated one model, and ultimately, 200 models were constructed. The AUC of each model was calculated using the testing subset, and the ratio of the mean to the standard deviation of the AUC values from the 200 iterations was used as an indicator of the model's performance stability. A higher ratio indicates more stable performance. Detailed methods are provided in the supplementary materials.

Validation of the joint model

To validate the improvement in performance after integrating radiomic signatures, we used AUC to evaluate the classification performance of each independent predictive factor and conducted statistical comparisons using DeLong test evaluated the joint model against two benchmarks: the radiomics-only model and the clinical-only model [24]. In this study, the precision-recall curve complements ROC analysis by illustrating the diagnostic trade-offs between precision and recall in identifying PD-RBD patients. The precision-recall curve, a tool for evaluating classification models, and this metric is particularly informative for imbalanced datasets, as it emphasizes classifier performance on the minority class. In addition, the goodness of fit of the joint model was assessed using the Hosmer-Lemeshow test [25], and the clinical net benefit of the model was evaluated using DCA curve. To further assess the generalizability of the joint model, the optimal classification threshold was determined in the training cohort by maximizing the Youden index. both the training and testing sets were classified into low- and high-risk groups based on the model's optimal cutoff value. Confusion matrix analysis was performed to evaluate the model's classification accuracy [26], including sensitivity, specificity, positive predictive value (PPV), and F1-score. Furthermore, to elucidate the contribution of features in constructing the random forest model, we calculated SHapley Additive exPlanations(SHAP) values using SHAP analysis, which quantifies the individual impact of each predictor on the model output. The study workflow is shown in Fig. 1.

Statistical analysis

All statistical analyses were performed using SPSS software version 17.0 (IBM, Armonk, NY) and R software version 3.4.1 (http://www.r-project.org). The normality of quantitative data was assessed using the Kolmogorov-Smirnov test [27]. For normally distributed data, group differences were compared using the independent sample t-test, and results are presented as Mean ± SD [28]. For non-normally distributed data, the Mann-Whitney U test was used, and results are presented as median (interquartile range, M[IQR]) [29]. Categorical data are presented as frequency and percentage (n, %), and group comparisons were conducted using the chi-square test (χ^2 test) [30]. A two-tailed p-value < 0.05 was considered statistically significant.

Results

Comparison of clinical factors

The average age of the participants included in this study was 63.82 years, with 118 male patients. In terms of clinical data, there were no statistically significant differences between the training and testing groups for all clinical variables except for symptom5 (P > 0.05). Additionally, there were no significant differences in clinical data between the – and PD-RBD subgroups in both the training and testing groups (P > 0.05), as detailed in Table 1.

Construction of the joint model

After dimensionality reduction, 10 features were extracted to construct the radiomic signature, including 3 features from white matter, 3 from gray matter, and 4

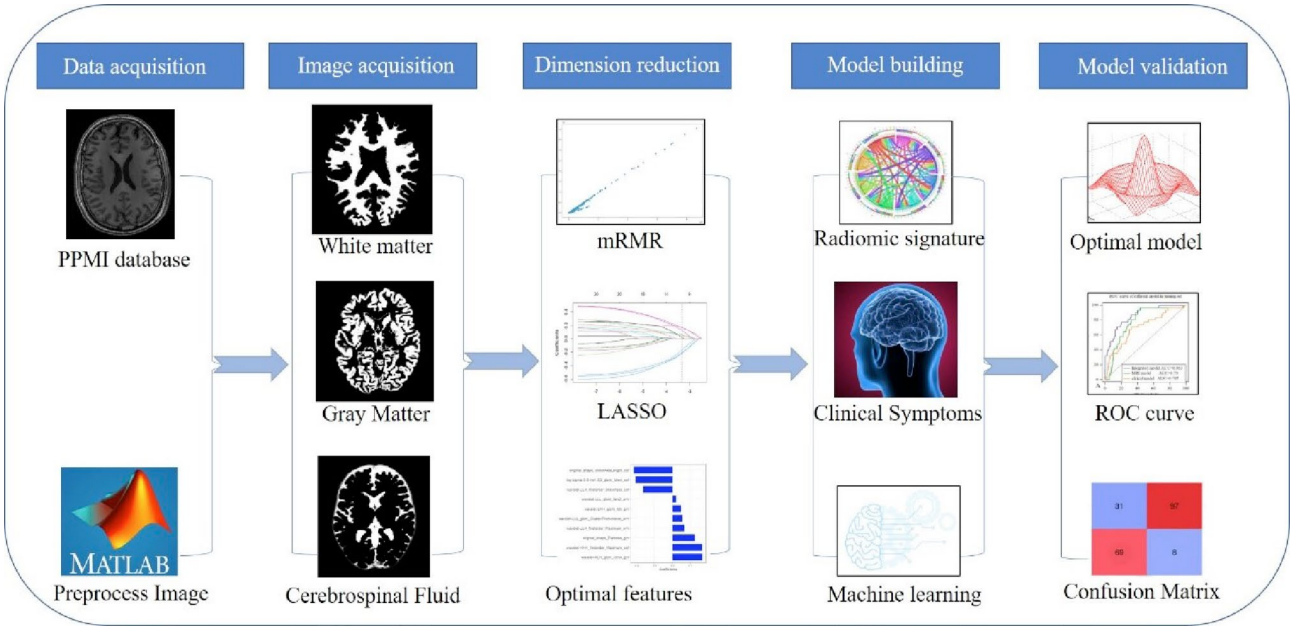


Fig. 1 Research flowchart shows the process of model construction and validation

from CSF regions. The radiomic signature constructed from these features achieved an AUC of 0.754 and 0.707 in the training and testing sets, respectively, with specificity values of 0.691 and 0.614, and sensitivity values of 0.733 and 0.75. Calibration curves further confirm the radiomics signature reliability without overfitting (see Fig. 2). Multivariate logistic regression analysis identified postural instability and the radiomic signature as independent predictors of PD-RBD and included them in the construction of the joint model (see Table 2). Further, joint models were built using SVM, Bayesian classifier, KNN, Decision tree, and random forest algorithms. The performance stability ratios were 31.705, 29.493, 30.115, 36.04, and 60.256, respectively. Compares the stability of five machine learning algorithms via a bee colony plot, demonstrating that the random forest model consistently achieved the highest AUC values across 200 iterations, underscoring its superior predictive robustness (see Fig. 3).

Validation of random forest models

The random forest model demonstrated excellent performance in both the training and testing sets. ROC curve analysis showed that the AUC values for the random forest model were 0.917 and 0.882 in the training and testing sets, respectively, with sensitivities of 0.933 and 0.889, and specificities of 0.786 and 0.722. The Hosmer–Lemeshow test indicated no overfitting in the random forest model ($P > 0.05$), and The DCA curve shows that the model demonstrated good clinical net benefits in both the training and testing cohort. Furthermore, the DeLong test revealed that the diagnostic performance of the random forest model was significantly superior to that of other independent predictors ($P < 0.05$) as detailed in Table 3, highlighting its significant improvement in predictive performance. The precision-recall curve further confirmed its strong diagnostic capability, with AUC values of 0.854 and 0.80 for the training and testing sets, respectively. The optimal threshold of 0.3772 was derived by maximizing the Youden index, defined as (sensitivity + specificity - 1). This metric identifies the value that best balances the model ability to correctly identify PD-RBD patients (sensitivity) and non-RBD individuals (specificity), dividing the subjects into high-risk and low-risk groups. Confusion matrix analysis showed that, both in the training and testing sets, there was a significant difference in the number of actual RBD-positive patients between the high-risk and low-risk groups ($P < 0.05$), further validating the clinical applicability of the random forest model (see Fig. 4). SHAP analysis shows that the contribution of radiomics signature is much greater than that of postural instability, Details can be found in the supplementary materials.

Discussion

This study developed a radiomics-based signature for predicting PD-RBD using whole-brain imaging features. The results demonstrated that the radiomics signature exhibited good predictive performance, indicating that changes in white matter, gray matter, and cerebrospinal fluid may be closely associated with the occurrence of RBD. Furthermore, by integrating relevant clinical features and utilizing the random forest machine learning algorithm to construct a combined model, the prediction

Table 1 Comparative analysis of clinical data for training cohort and test cohort

Characteristics		Training cohort (n = 129)				Test cohort (n = 54)				*Train VS Test* P value
		Combined cohort (n = 129)	PD-NC (n = 84)	PD-RBD (n = 45)	*P* value	Combined cohort (n = 54)	PD-NC (n = 36)	PD-RBD (n = 18)	*P* value	
Age(year)		64 ± 7.1	63 ± 7.6	64 ± 7.8	0.956[a]	63.4 ± 7	63.1 ± 7.9	64 ± 5.2	0.665[a]	0.614[a]
Gender (n,%)	Male	86 (66.7)	54 (64.3)	32 (71.1)	0.557	32 (59.3)	21 (58.3)	11 (61.1)	1.000	0.432
	Female	43 (33.3)	30 (35.7)	13 (28.9)		22 (40.7)	15 (41.7)	7 (38.9)		
Education (year)		15.8 ± 2.7	15.7 ± 2.9	16 ± 2.2	0.582[a]	16.2 ± 3.2	16 ± 2.9	16.7 ± 3.9	0.461[a]	0.410[a]
HY-stage	0–1	62 (48.1)	44 (52.4)	18 (40.0)	0.248	25 (46.3)	14 (38.9)	11 (61.1)	0.210	0.955
	2–3	67 (51.9)	40 (47.6)	27 (60.0)		29 (53.7)	22 (61.1)	7 (38.9)		
Symptom1	No	27 (20.9)	16 (19.0)	11 (24.4)	0.623	9 (16.7)	8 (22.2)	1 (5.6)	0.245	0.647
	Yes	102 (79.1)	68 (81.0)	34 (75.6)		45 (83.3)	28 (77.8)	17 (94.4)		
Symptom2	No	34 (26.4)	20 (23.8)	14 (31.1)	0.492	12 (22.2)	10 (27.8)	2 (11.1)	0.298	0.688
	Yes	95 (73.6)	64 (76.2)	31 (68.9)		42 (77.8)	26 (72.2)	16 (88.9)		
Symptom3	No	19 (14.7)	13 (15.5)	6 (13.3)	0.947	9 (16.7)	6 (16.7)	3 (16.7)	1.000	0.915
	Yes	110 (85.3)	71 (84.5)	39 (86.7)		45 (83.3)	30 (83.3)	15 (83.3)		
Symptom4	No	117 (90.7)	79 (94.0)	38 (84.4)	0.141	47 (87.0)	33 (91.7)	14 (77.8)	0.316	0.635
	Yes	12 (9.3)	5 (6.0)	7 (15.6)		7 (13.0)	3 (8.3)	4 (22.2)		
Symptom5	No	100 (77.5)	64 (76.2)	36 (80.0)	0.785	50 (92.6)	32 (88.9)	18 (100.0)	0.358	0.027*
	Yes	29 (22.5)	20 (23.8)	9 (20.0)		4 (7.4)	4 (11.1)	0 (0.0)		
ESS (n,%)	No	109 (84.5)	74 (88.1)	35 (77.8)	0.198	43 (79.6)	29 (80.6)	14 (77.8)	1.000	0.559
	Yes	20 (15.5)	10 (11.9)	10 (22.2)		11 (20.4)	7 (19.4)	4 (22.2)		
GDS (n,%)	No	117 (90.7)	78 (92.9)	39 (86.7)	0.403	49 (90.7)	33 (91.7)	16 (88.9)	1.000	1.000
	Yes	12 (9.3)	6 (7.1)	6 (13.3)		5 (9.3)	3 (8.3)	2 (11.1)		
MCI (n,%)	No	96 (74.4)	65 (77.4)	31 (68.9)	0.400	37 (68.5)	27 (75.0)	10 (55.6)	0.254	0.525
	Yes	33 (25.6)	19 (22.6)	14 (31.1)		17 (31.5)	9 (25.0)	8 (44.4)		

Note. RBD, rapid eye movement sleep behavior disorder; NC, Negative Control; Hy, Categorical Hoehn & Yahr (OFF State); Symptom1, Initial symptom (at diagnosis)-Resting Tremor; Symptom2, Initial symptom (at diagnosis)-Rigidity; Symptom3, Initial symptom (at diagnosis)-Bradykinesia; Symptom4, Initial symptom (at diagnosis)-Postural Instability; Symptom5, Initial symptom (at diagnosis)-Other; Ess_cat, Categorical Epworth Sleepiness Scale Score; GDS, Categorical Geriatric Depression Scale; MCI: Mild cognitive impairment; *, $P < 0.05$. a, indicates that independent sample t-test was used for this clinical variable

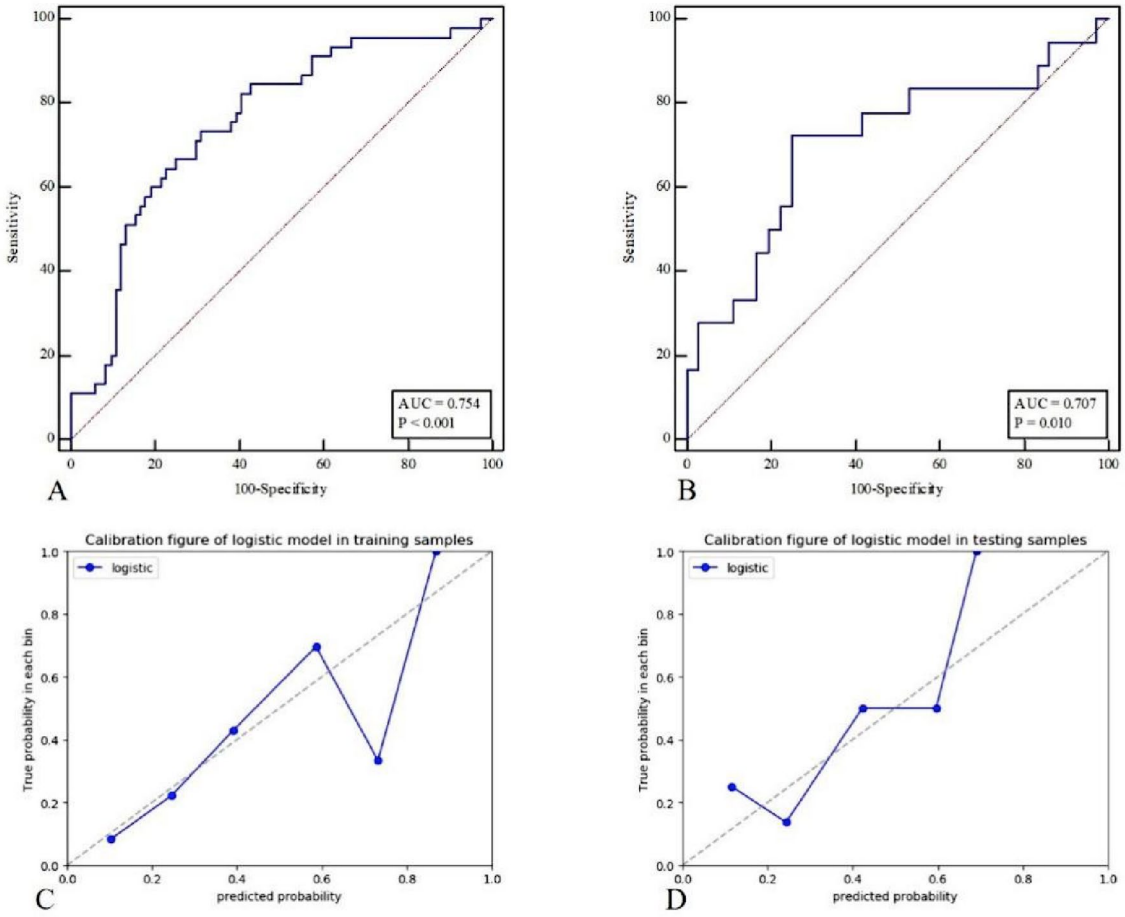


Fig. 2 Figures **A** and **B** display the ROC curves for the diagnostic performance of the radiomic signature in the training set (left) and testing set (right). Figures **C** and **D** show the calibration curves for the signature in the training set (left) and testing set (right)

performance for PD-RBD was significantly improved in both the training and validation sets. This finding further support the potential utility of machine learning techniques in predicting PD-RBD.

Our radiomic analysis identified ten radiomics features that effectively differentiate PD patients with RBD from those without. The white matter features likely reflect degenerative changes in dopaminergic neurons extending into axons that connect striatal and cortical regions. Previous studies have demonstrated that white matter alterations, particularly in brainstem pathways and limbic connections, correlate with RBD severity [31]. The gray matter features identified in our model may represent structural changes in key brainstem nuclei, especially the laterodorsal tegmental nucleus and locus coeruleus, which are directly implicated in REM sleep muscle atonia regulation [32]. The dysfunction of these regions is considered central to RBD pathophysiology, as they normally inhibit motor neurons during REM sleep. Regarding CSF features, they potentially capture inflammatory responses and neurotransmitter metabolite changes specifically associated with RBD. Alpha-synuclein pathology, which initially affects brainstem structures in PD, likely contributes to these radiomic signatures by disrupting neuronal integrity in regions critical for sleep regulation [33]. These structural alterations detected through our radiomic approach may precede clinical manifestations, highlighting the potential value of MRI radiomic features as early biomarkers for RBD in PD patients. On the other hand, studies have used baseline EEG features of isolated RBD patients to predict RBD, achieving an AUC of 0.901, which is notably higher than the diagnostic performance of the radiomics signature [34]. However, it is important to note that compared to isolated RBD patients, PD patients have more complex brain pathological features. For example, PD patients exhibit increased resting-state EEG complexity compared to healthy controls [35, 36], which may partly explain the relatively lower diagnostic performance of the radiomics signature. Nonetheless, unlike EEG features that directly reflect electrophysiological changes in the brain, radiomics features are based on structural changes and may contain more pathology-related information. Moreover, previous research using manually scored sleep EMG features to predict RBD

Table 2 Identification of Independent Predictors for RBD Based on Multivariate Logistic Regression

Variable	Univariate logistic regression		Multivariate logistic regression	
	OR(95%CI)	*P* value	OR(95%CI)	*P* value
Age(year)	0.924(0.86,0.993)	0.032*	NA	NA
Gen	0.481(0.172,1.345)	0.163	NA	NA
Education	0.939(0.783,1.126)	0.498	NA	NA
Symptom1	2.798(0.605,12.937)	0.188	NA	NA
Symptom2	0.28(0.086,0.906)	0.034*	NA	NA
Symptom3	1.89(0.437,8.174)	0.394	NA	NA
Symptom4	0.277(0.081,0.948)	0.041*	0.333(0.115,0.961)	0.042*
Symptom5	0.52(0.157,1721)	0.285	NA	NA
Hy	1.519(0.583,3.957)	0.392	NA	NA
Ess_cat	3.01(0.786,11.525)	0.108	NA	NA
Gds_cat	3.683(0.748,18.136)	0.109	NA	NA
MCI	1.712(0.589,4.982)	0.324	NA	NA
Radscore	3.911(0.831, 18.422)	<0.001*	1.784(0.519, 6.132)	<0.001*

Note. Symptom1, Initial symptom (at diagnosis)-Resting Tremor; Symptom2, Initial symptom (at diagnosis)-Rigidity; Symptom3, Initial symptom (at diagnosis)-Bradykinesia; Symptom4, Initial symptom (at diagnosis)-Postural Instability; Symptom5, Initial symptom (at diagnosis)-Other; Hy, Categorical Hoehn & Yahr (OFF State); Ess_cat, Categorical Epworth Sleepiness Scale Score; gds_cat, Categorical Geriatric Depression Scale; MCI, mild cognitive impairment; NA, not available because the variable is not included in multiple variables. The P value indicates whether the variable is an independent predictor of PD-REM. *, P < 0.05

Table 3 ROC curve analysis of the models in different cohorts

Features	Training Cohort (n = 129)	Testing Cohort (n = 54)
Random forest model		
AUC (95%CI)	0.917(0.836–0.998)	0.882(0.842–0.915)
Sensitivity	0.933	0.889
Specificity	0.786	0.722
PPV	0.701	0.615
F1 score	0.812	0.727
Rad socre		
AUC (95%CI)	0.754(0.628–0.880)	0.707 (0.650–0.760)
Sensitivity	0.733	0.75
Specificity	0.691	0.614
PPV	0.635	0.602
F1 score		
Postural Instability		
AUC (95%CI)	0.548(0.458–0.636)	0.569(0.428–0.703)
Sensitivity	0.156	0.222
Specificity	0.941	0.917
PPV	0.581	0.571
F1 score	0.246	0.320

Abbreviations: AUC: area under the curve; PPV: positive predictive value

reported an AUC of 0.699, which is significantly lower than the results of our study [37]. This further suggests that MRI-based radiomics features may capture more information related to RBD, thus showing greater potential for predicting RBD status. Furthermore, our findings highlight a critical distinction between functional and structural biomarkers. While EEG and EMG excel in detecting functional abnormalities during REM sleep, MRI-based radiomics captures structural neurodegeneration in regions critical for sleep regulation, such as brainstem nuclei and white matter pathways, which may precede clinical symptoms. These structural biomarkers provide early, stable indicators of RBD progression, complementing EEG/EMG real-time functional data. Future studies integrating both functional and structural biomarkers could further optimize RBD prediction in PD patients.

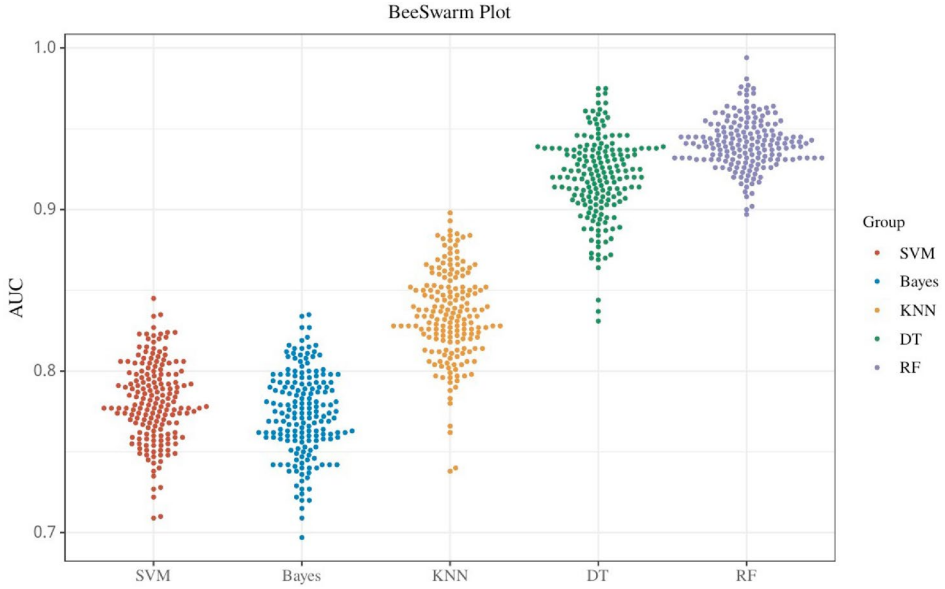


Fig. 3 The bee colony plot displays the distribution of AUC values generated from 200 cycles using 5 machine learning methods, with RF having the highest AUC value

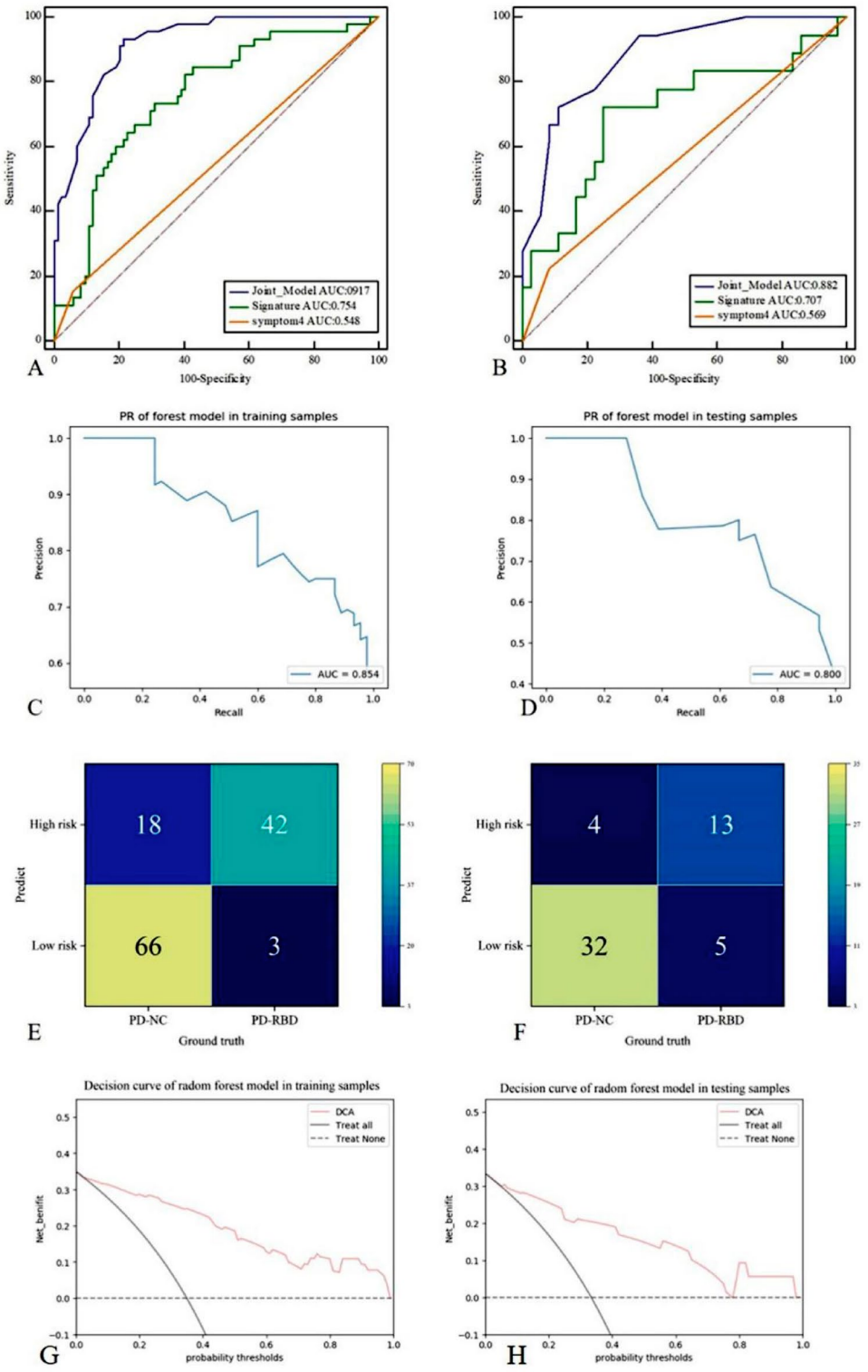


Fig. 4 Demonstrates the clinical applicability of the random forest model through ROC and Precision-Recall (PR) curves, alongside confusion matrices. Panels **A** and **B** compare the model diagnostic performance against individual predictors in training and testing cohorts, highlighting its superior accuracy. The PR curves (panels **C** and **D**) illustrate the balance between precision and recall for identifying PD-RBD patients. Confusion matrices (panels **E** and **F**) validate the model ability to stratify patients into low- and high-risk groups, with statistically significant differences in actual RBD cases between these groups ($P < 0.05$), underscoring its utility in clinical decision-making. panels G and H present Decision Curve Analysis (DCA) curves in training and testing cohort, quantifying the net benefit of the model across varying probability thresholds. The DCA curves compare the model clinical utility against "treat all" and "treat none" strategies, demonstrating superior net benefit within the clinically relevant threshold range (0.4–0.8)

In this study, we also found that postural instability is an independent predictor of PD-RBD. Postural instability typically reflects significant damage to basal ganglia function, particularly in the substantia nigra and striatum, where neurons undergo substantial degenerative changes [38]. The basal ganglia not only regulate motor function but also participate in the regulation of the autonomic nervous system. Previous studies have indicated that PD patients with RBD exhibit both structural and functional abnormalities in the nigrostriatal system, alongside changes in the limbic system and cortex, suggesting that RBD is a multi-system neurodegenerative process [39]. Against this background, we hypothesize that when the postural control system is impaired, the muscle inhibition function during rapid eye movement (REM) sleep may also be affected, thereby increasing the risk of RBD development. In fact, many PD patients, after the onset of motor symptoms, frequently develop RBD [40]. This suggests that postural instability in PD patients may reflect broader neurodegenerative changes, including dysfunction in the nigrostriatal system, which is associated with autonomic dysregulation. These alterations might correlate with disruptions in REM sleep motor inhibition, potentially elevating the likelihood of RBD development. Therefore, monitoring the occurrence of postural instability is crucial for the early identification of PD-RBD patients, aiding in timely intervention during critical stages of disease progression.

We also constructed a combined model using multiple machine learning algorithms, and the results demonstrated that the random forest algorithm performed the best, significantly outperforming other algorithms. This may be attributed to its advantages in handling high-dimensional data and non-linear features, a finding consistent with prior research evaluating PD-RBD status [41]. Compared to alternative approaches, such as serum miRNA assays (AUC: 0.98) [42] or olfactory dysfunction tests (AUC: 0.95) [43], and MRI-based radiomics offers distinct practical advantages. While miRNA and olfactory methods achieve high accuracy, their reliance on costly molecular assays or specialized olfactory tools limits scalability in routine clinical practice. In contrast, radiomics leverages widely accessible MRI to quantify structural neurodegeneration in brain regions linked to RBD pathophysiology, providing a cost-effective and non-invasive alternative. Overall, the clinical translation of our model could follow a phased implementation strategy, initially targeting specialized centers with advanced MRI and AI resources before expanding to broader clinical settings. By integrating radiomics biomarkers with clinical assessments, this approach may reduce diagnostic delays and prioritize high-risk PD patients for polysomnography, optimizing resource utilization. Such pathways mirror successes in neurodegenerative disease management, where AI tools have enhanced diagnostic precision. While challenges like multicenter validation persist, our model represents a pragmatic step toward scalable, precision-based neurology, balancing innovation with real-world applicability.

This study has several limitations. First, it is a retrospective design, and the data were sourced from a single-center database, lacking external datasets for model validation. Therefore, future multi-center studies with large sample sizes are needed to further validate the stability and generalizability of this model. Second, this study only performed a cross-sectional analysis to identify RBD, without conducting longitudinal research to assess the accuracy of the prediction model over time. Additionally, although the random forest algorithm performed excellently in this study, its "black-box" nature may limit clinicians' ability to interpret the model's predictions. Future studies will incorporate explainable algorithms to enhance the model's transparency, improving its acceptability and application value in clinical practice.

The integration of magnetic resonance imaging-based radiomics and machine learning offers new possibilities for the early diagnosis and intervention of Parkinson's disease and its related sleep disorders. With ongoing technological development and optimization, this approach may be widely applicable in the diagnosis of other neurodegenerative diseases, thereby advancing the development of personalized and precision medicine.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12880-025-01748-4.

Supplementary Material 1

Author contributions

YL and YX designed the research study. YW and LH analyzed the data. YX and YL performed the research and made manuscriptpreparation. All authors contributed to the drafting or important editorial changes in the manuscript. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Funding

This study is funded by the Jinhua Science and Technology Research Program (2021-3-060).

Data availability

No datasets were generated or analysed during the current study.

Declaration

Consent to participate

The case data used in this study came from the Parkinson's Progression Markers Initiative (PPMI) (http://www.PPMI-info.org) database, and data collection was approved by the institutional review board; For ethical review information on the data, please refer to the website

Clinical trial number:

Not applicable

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Radiology, The First Affiliated Hospital of Ningbo University, Ningbo, Zhejiang Province, China
[2]Department of Radiology, Jinhua Municipal Central Hospital, Jinhua, Zhejiang Province, China
[3]Advanced Analytics, GE Healthcare, Hangzhou, China
[4]Department of Magnetic Resonance, Heze Municipal Hospital, Heze, Shandong, China

Received: 27 January 2025 / Accepted: 26 May 2025
Published online: 01 July 2025

References

1. Kalia LV, Lang AE Parkinson's disease. Lancet. 2015, 386(9996):896–912.
2. Sveinbjornsdottir S The clinical symptoms of Parkinson's disease. J Neurochem. 2016, 139 Suppl 1:318–24.
3. Bruno V, Ruiz-Lopez M, Terroba-Chambi C, et al. Rapid Eye Movement Sleep Behavior Disorder in Parkinson's Disease: a Survey-Based Study. Can J Neurol Sci. 2023, 50(5):703–09.
4. Oliveira P, Cardoso F Impact of rapid eye movement sleep behavior disorder and autonomic disorders on Parkinson's disease: a review. Arq Neuropsiquiatr. 2021, 79(2):156–66.
5. St Louis EK, Boeve BF REM Sleep Behavior Disorder: diagnosis, Clinical Implications, and Future Directions. Mayo Clin Proc. 2017, 92(11):1723–36.
6. Jones BM, McCarter SJ Rapid Eye Movement Sleep Behavior Disorder: clinical Presentation and Diagnostic Criteria. Sleep Med Clin. 2024, 19(1):71–81.
7. Que Z, Zheng C, Zhao Z, et al. The treatment efficacy of pharmacotherapies for rapid eye movement sleep behavior disorder with polysomnography evaluation: a systematic review and meta-analysis. Heliyon. 2022, 8(11):e11425.
8. Yousaf T, Dervenoulas G, Politis M Advances in MRI Methodology. Int Rev Neurobiol. 2018, 141:31–76.
9. Mayerhoefer ME, Materka A, Langs G, et al. Introduction to Radiomics. J Nucl Med. 2020, 61(4):488–95.
10. Shahidi R, Baradaran M, Asgarzadeh A, et al. Diagnostic performance of MRI radiomics for classification of Alzheimer's disease, mild cognitive impairment, and normal subjects: a systematic review and meta-analysis. Aging Clin Exp Res. 2023, 35(11):2333–48.
11. Dragoș HM, Stan A, Pintican R, et al. MRI Radiomics and Predictive Models in Assessing Ischemic Stroke Outcome-A Systematic Review. Diagnostics (Basel). 2023,13(5):857.
12. Papanastasiou G, García Seco de Herrera A, Wang C, et al. Focus on machine learning models in medical imaging. Phys Med Biol. 2022, 68(1).
13. Greener JG, Kandathil SM, Moffat L, et al. A guide to machine learning for biologists. Nat Rev Mol Cell Biol. 2022, 23(1):40–55.
14. Panahi M, Hosseini MS Multi-modality radiomics of conventional T1 weighted and diffusion tensor imaging for differentiating Parkinson's disease motor subtypes in early-stages. Sci Rep. 2024, 14(1):20708.
15. Arnaldi D, Mattioli P, Raffa S, et al. International REM Sleep Behavior Disorder Study Group. Presynaptic Dopaminergic Imaging Characterizes Patients with REM Sleep Behavior Disorder Due to Synucleinopathy. Ann Neurol. 2024,95(6):1178–92.
16. Peng J, Wang W, Song Q, et al. ^{18}F-FDG-PET Radiomics Based on White Matter Predicts The Progression of Mild Cognitive Impairment to Alzheimer Disease: a Machine Learning Study. Acad Radiol. 2023,30(9):1874–84.
17. Zwanenburg A, Vallières M, Abdalah MA, et al. The Image Biomarker Standardization Initiative: standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. Radiology. 2020, 295(2):328–38.
18. Tavares V, Prata D, Ferreira HA Comparing SPM12 and CAT12 segmentation pipelines: a brain tissue volume-based age and Alzheimer's disease study. J Neurosci Methods. 2019, 334:108565.
19. Yushkevich PA, Gao Y, Gerig G ITK-SNAP: an interactive tool for semi-automatic segmentation of multi-modality biomedical images. Annu Int Conf IEEE Eng Med Biol Soc. 2016, 2016:3342–45.
20. van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational Radiomics System to Decode the Radiographic Phenotype. Cancer Res. 2017, 77(21):e104–e107.
21. Mu W, Schabath MB, Gillies RJ Images Are Data: challenges and Opportunities in the Clinical Translation of Radiomics. Cancer Res. 2022, 82(11):2066–68.
22. Zhang Y, Peng J, J L, et al. Preoperative Prediction of Perineural Invasion Status of Rectal Cancer Based on Radiomics Nomogram of Multiparametric Magnetic Resonance Imaging. Front Oncol. 2022,12:828904.
23. Suroor N, Jaiswal A, Sachdeva N Stack Ensemble Oriented Parkinson Disease Prediction Using Machine Learning Approaches Utilizing GridSearchCV-Based Hyper Parameter Tuning. Crit Rev Biomed Eng. 2022,50(5):39–58.
24. DeLong ER, DeLong DM, Clarke-Pearson DL Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988,44(3):837–45.
25. Surjanovic N, Loughin TM Improving the Hosmer-Lemeshow goodness-of-fit test in large models with replicated Bernoulli trials. J Appl Stat. 2023, 51(7):1399–411.
26. Landgrebe TC, Duin RP Efficient multiclass ROC approximation by decomposition via confusion matrix perturbation analysis. IEEE Trans Pattern Anal Mach Intell. 2008,30(5):810–22.
27. Ghasemi A, Zahediasl S Normality tests for statistical analysis: a guide for non-statisticians. Int J Endocrinol Metab. 2012 Spring;10(2):486–89.
28. Kim TK T test as a parametric statistic. Korean J Anesthesiol. 2015,68(6):540–46.
29. Fay MP, Proschan MA Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules. Stat Surv. 2010;4:1–39
30. McHugh ML The chi-square test of independence. Biochem Med (Zagreb). 2013;23(2):143–49.
31. Tan S, Wen J, J Q, et al. Wider and faster degeneration of white matter in Parkinson's disease with possible REM sleep behaviour disorder. Sleep Med. 2025, 126:97–106.
32. Kim HJ, Im HK, Kim J, et al. Brain Atrophy of Secondary REM-Sleep Behavior Disorder in Neurodegenerative Disease. J Alzheimers Dis. 2016,52(3):1101–09.
33. Concha-Marambio L, Weber S, Farris CM, et al. Accurate Detection of α-Synuclein Seeds in Cerebrospinal Fluid from Isolated Rapid Eye Movement Sleep Behavior Disorder and Patients with Parkinson's Disease in the DeNovo Parkinson (DeNoPa) Cohort. Mov Disord. 2023,38(4):567–78.
34. Jeong E, Woo Shin Y, Byun JI, et al. EEG-based machine learning models for the prediction of phenoconversion time and subtype in isolated rapid eye movement sleep behavior disorder. Sleep. 2024,47(5):zsae031.
35. Yao L, Brown P, Shoaran M Resting Tremor Detection in Parkinson's Disease with Machine Learning and Kalman Filtering. IEEE Biomed Circuits Syst Conf. 2019,2018:BIOCAS.2018.8584721.
36. Lee SB, Kim YJ, Hwang S, et al. Predicting Parkinson's disease using gradient boosting decision tree models with electroencephalography signals. Parkinsonism Relat Disord. 2022,95:77–85.
37. Cesari M, Portscher A, A S, et al. Machine Learning Predicts Phenoconversion from Polysomnography in Isolated REM Sleep Behavior Disorder. Brain Sci. 2024,14(9):871.
38. Debû B, De Oliveira Godeiro C, Lino JC, et al. Managing Gait, Balance, and Posture in Parkinson's Disease. Curr Neurol Neurosci Rep. 2018,18(5):23.
39. Valli M, Uribe C, Mihaescu A, et al. Neuroimaging of rapid eye movement sleep behavior disorder and its relation to Parkinson's disease. J Neurosci Res. 2022,100(10):1815–33.
40. Kataoka H, Sugie K Risk for later rapid eye movement sleep behavior disorder in Parkinson's disease: a 6-year prospective study. Int J Neurosci. 2020,130(3):237–42.
41. Chong-Wen W, Sha-Sha L, Xu E Predictors of rapid eye movement sleep behavior disorder in patients with Parkinson's disease based on random forest and decision tree. PLoS One. 2022,17(6):e0269392.
42. Soto M, Iranzo A, Lahoz S, et al. Serum MicroRNAs Predict Isolated Rapid Eye Movement Sleep Behavior Disorder and Lewy Body Diseases. Mov Disord. 2022,37(10):2086–98.
43. Lo C, Arora S, Ben-Shlomo Y, et al. Olfactory Testing in Parkinson Disease and REM Behavior Disorder: a Machine Learning Approach. Neurology. 2021,96(15):e2016–e2027.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.