

Multivariate radiomics models based on 18F-FDG hybrid PET/MRI for distinguishing between Parkinson's disease and multiple system atrophy

Xuehan Hu

Union Hospital, Tongji Medical College, Huazhong University of Science and Technology

Xun Sun

Union Hospital, Tongji Medical College, Huazhong University of Science and Technology

Fan Hu

Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology

Fang Liu

Union Hospital, Tongji Medical College, Huazhong University of Science and Technology

Weiwei Ruan

Union Hospital, Tongji Medical College, Huazhong University of Science and Technology

Tingfan Wu

Ge Healthcare

Rui An

Union Hospital, Tongji Medical College, Huazhong University of Science and Technology

Xiaoli Lan (✉ xiaoli_lan@hust.edu.cn)

Union Hospital, Tongji Medical College, Huazhong University of Science & Technology

<https://orcid.org/0000-0002-7263-7399>

Research Article

Keywords: Radiomics, PET/MRI, Parkinson's disease, multiple system atrophy, differential diagnosis

Posted Date: February 26th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-243430/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at European Journal of Nuclear Medicine and Molecular Imaging on April 7th, 2021. See the published version at <https://doi.org/10.1007/s00259-021-05325-Z>.

Abstract

Purpose To construct multivariate radiomics models using hybrid ^{18}F -FDG PET/MRI for distinguishing between Parkinson's disease (PD) and multiple system atrophy (MSA).

Methods Ninety patients (60 with PD and 30 with MSA) were randomised to training and validation sets in a 7:3 ratio. All patients underwent ^{18}F -Fluorodeoxyglucose (^{18}F -FDG) PET/MRI to simultaneously obtain metabolic images (^{18}F -FDG), structural MRI images (T1-weighted imaging [T1WI], T2-weighted imaging [T2WI] and T2-weighted fluid-attenuated inversion recovery [T2/flair]) and functional MRI images (susceptibility-weighted imaging [SWI] and apparent diffusion coefficient). Using PET and five MRI sequences, we extracted 1172 radiomics features from the putamina and caudate nuclei. The radiomics signatures were constructed with the least absolute shrinkage and selection operator algorithm in the training set, with progressive optimization through single-sequence and double-sequence radiomics models. Multivariable logistic regression analysis was used to develop a clinical-radiomics model, combining the optimal multi-sequence radiomics signature with clinical characteristics and SUV values. The diagnostic performance of the models was assessed by receiver operating characteristic and decision curve analysis (DCA).

Results The radiomics signatures showed favourable diagnostic efficacy. The optimal model comprised structural (T1WI), functional (SWI), and metabolic (^{18}F -FDG) sequences ($\text{Radscore}_{\text{FDG_T1WI_SWI}}$) with the area under curves (AUCs) of the training and validation sets of 0.971 and 0.957, respectively. The integrated model, incorporating $\text{Radscore}_{\text{FDG_T1WI_SWI}}$, three clinical symptoms (disease duration, dysarthria and autonomic failure) and SUV_{max} , demonstrated satisfactory calibration and discrimination in the training and validation sets (0.993 and 0.994, respectively). DCA indicated the highest clinical benefit of the clinical-radiomics integrated model.

Conclusions The radiomics signature with metabolic, structural, and functional information provided by hybrid ^{18}F -FDG PET/MRI may achieve promising diagnostic efficacy for distinguishing between PD and MSA. The clinical-radiomics integrated model performed best.

□

Introduction

Parkinson's disease (PD) is the second most common neurodegenerative disease in the world and is characterised by the progressive degeneration and death of dopaminergic neurons in the substantia nigra and neuron terminals in the basal ganglia [1]. At present, the diagnosis of PD mainly depends on clinical manifestations [2], as identified by the physicians relying on their clinical experience. A recent study of clinical pathology demonstrated that only 50% of early Parkinsonian patients were accurately diagnosed with PD [3]; others were mostly misdiagnosed as atypical parkinsonian syndromes (APS). Multiple system atrophy (MSA), as one of the most common APSs, is difficult to distinguish from PD for that the clinical symptoms always overlap, especially in the early stages [4]. However, MSA progresses aggressively and has a poorer prognosis, with a mean survival of 8–9 years [5]. Early and reliable distinction between PD and MSA is very important for developing individualized treatment plans as soon as possible, which not only can improve the patients'

survival and quality of life, but also can reduce psychosocial and economic burdens on patients and their families [6].

Neuroimaging has become an indispensable tool for the diagnosis of PD and other motor disorders. Magnetic resonance imaging (MRI) is used for distinguishing between PD and typical MSA, with putaminal atrophy and rim sign, the 'hot-cross-bun sign' in MSA [5], while there is generally no specific MRI signal in PD [7]. However, the aforementioned abnormalities are insufficient to visualise in the early stages [8], which means conventional MRI has a lower sensitivity for early differential diagnosis of PD and MSA. Although some application of neurofunctional MRI, such as resting-state functional MRI and diffusion tensor imaging, has been confirmed to help for discrimination between PD and MSA with high accuracy [9, 10], they are not routinely used because of with complicated reprocessing and Parkinsonian patients having difficulties to remain immobile for a relatively long scanning time. Compared with brain MRI, ^{18}F -fluorodeoxyglucose (^{18}F -FDG) positron emission tomography (PET) shows a higher diagnostic sensitivity [11]. Glucose metabolism is always normal or slightly higher in the basal ganglia in early PD, while it is lower in MSA because of putaminal microstructural damage [2, 12]. However, PET imaging can only help to support or refute clinical perceptions. It cannot be used for direct diagnosis without the establishment of regulatory guidelines [7, 13].

Early studies have shown the potential of radiomics for disease detection, diagnosis, prognosis, and therapeutic assessment [14-18]. Radiomics refers to converting medical images into mineable high-dimensional data via high-throughput extraction of abundant imaging features [14, 19], which can help for clinical decision-making through quantitative analysis of these data. Increasing studies have shown that radiomics and deep learning can obtain superior diagnostic performance in PD clinical assessment or distinction between PD and healthy controls or APSs via different brain imaging techniques [20-27]. However, there have been no studies to realise one-stop access to the structural, functional and metabolic images to maximize the advantages for distinguishing among PD and other diseases using radiomics analysis.

A hybrid PET/MRI system combines two optimal brain imaging modalities and acquires simultaneously with metabolic, functional and structural images, which ensures their consistency in time and space [28, 29]. In this study, we aimed to develop an optimal radiomics signature using hybrid PET/MRI for the first time to distinguish between PD and MSA. Furthermore, we constructed a clinical-radiomics integrated model to explore a more efficient strategy for diagnosing and identifying of PD and MSA.

Materials And Methods

Patients

Ninety patients, including 60 PD cases and 30 MSA cases from December 2017 to June 2019, were retrospectively enrolled in our study. Inclusion criteria were as follows: (1) the diagnosis of PD was based on the Movement Disorder Society PD Criteria [30], while MSA was diagnosed based on the second consensus statement on the diagnosis of multiple system atrophy [31] (which were international standard diagnostic criteria for PD and MSA for unavailable pathological diagnoses), and all fulfilling diagnoses were confirmed by neurological physicians with more than 10 years of experience; (2) disease duration was no more than 5 years and the average follow-up period was more than 1.5 years; and (3) all patients had undergone ^{18}F -FDG

PET/MRI examination with MRI sequences including T1-weighted imaging (T1WI), T2-weighted imaging (T2WI), T2-weighted fluid-attenuated inversion recovery (T2/FLAIR) imaging, diffusion-weighted imaging (DWI), and susceptibility-weighted imaging (SWI). Exclusion criteria were: (1) evidence of vascular disease confirmed on computed tomography (CT) or MRI; (2) organic lesions (such as trauma, tumours and infections) and other degenerative diseases; (3) severe motion artefacts on images and significant head movement during the scan; and (4) a blood glucose level was ≥ 11.1 mmol/L.

The patient screening process is shown in the study flowchart in **Fig. 1a**. Patients cohorts were randomly divided into a training set (n=63) and validation set (n=27) with a ratio of 7:3. This study is a retrospective study based on data from one of our clinical studies, in which all patients signed informed consent forms and ethical approval was obtained by the Ethics Committee of Tongji Medical College, Huazhong University of Science and Technology.

Image acquisition and reconstruction of ^{18}F -FDG PET/MRI

PET/MRI images were acquired on a 3.0T time-of-flight (TOF) Signa PET/MRI (GE Healthcare, Milwaukee WI, USA). All participants fasted for at least 6 h and stopped any drugs that could affect brain metabolism for at least 12 h before the ^{18}F -FDG PET/MRI acquisition. An ^{18}F -FDG dose of 0.1 mCi/kg (3.7 MBq/kg) was intravenously injected after ensuring the blood glucose level was < 11.1 mmol/L. The participants rested in a quiet and dimly lit room before and after the ^{18}F -FDG injection until the start of imaging. At 1h after intravenous injection of ^{18}F -FDG, an MRI and a 15-min PET scan in the three-dimensional mode were acquired. MR imaging protocols included T1WI (time of Repetition [TR]/time of echo [TE], 7.9 ms/3.0 ms; flip angles 12.0° ; bandwidth 41.67 MHz; matrix 288×224 ; section thickness 1 mm), T2/FLAIR (TR/TE, 9000 ms/100 ms; inversion time 2475; refocus flip angles 160.0° ; bandwidth 41.67 MHz; matrix 256×192 ; section thickness 5 mm; intersection gap 1.5 mm), T2WI (TR/TE, 5523 ms/105 ms; refocus flip angles 142.0° ; bandwidth 50.00 MHz; matrix 384×240 ; section thickness 5mm; intersection gap 1.5mm), SWI (TR/TE, 45.5 ms/4.0 ms; flip angles 15.0° ; bandwidth 41.67 MHz; matrix 384×320 ; section thickness 3 mm), and DWI (TR/TE, 6379.0 ms/70.0 ms; bandwidth 41.67 MHz; matrix 128×128 ; section thickness 5 mm; intersection gap 1.5mm). MRI images were acquired simultaneously with the PET acquisition without re-positioning. The apparent diffusion coefficient (ADC) map ($b = 1000$) was generated from DWI images. An atlas-based method was used for PET attenuation correction. For PET imaging, the ordered subsets expectation maximization (OSEM) iterative reconstruction algorithm with 28 subsets, 2 iterations, and 2.14 mm (full width at half maximum) post-filtering was used.

Region-of-interest segmentation, image pre-processing and feature extraction

The radiomics workflow is shown in **Fig. 1b**. The bilateral putamina and caudate nuclei were selected as the regions of interest (ROIs), which were segmented on T1WI images through the open-source software ITK-SNAP (version 3.6.0, www.itksnap.org). To minimize partial volume effects, these ROIs excluded the most inferior and most superior slices including these structures [32]. The ROIs were delineated manually by a nuclear medicine physician with 2-3 years' experience (Hu X) who was blinded to subject information, and repeated measurements were performed at an interval of 6 weeks. All ROIs were confirmed by two neuroradiologists

who had over 10 years of experience (Sun X and Liu F). The intra-observer differences were calculated by the intraclass correlation coefficient (ICC).

Based on the T1WI images, spatial registration of PET and other MR images was carried out using the SPM software package (Version 12.0, <http://www.fil.ion.ucl.ac.uk/spm/>) implemented in MATLAB 2016a (MathWorks, Natick, MA, USA) to provide the same spatial information (thickness, slice and interlamellar space). ^{18}F -FDG PET images were transformed into SUV maps by normalisation by injected dose and patients' weight using LIFEx software (version 6.20; www.lifexsoft.org), to provide the SUV values (SUVmax, SUVmean, SUVmin) in the corresponding ROIs.

The radiomics features were extracted using Anaconda Prompt (version 4.2.0) importing the feature package of pyradiomics (github.com/Radiomics/pyradiomics), according to the feature guidelines of the Image Biomarker Standardization Initiative (IBSI) [33, 34].

Optimal multimodal radiomics signature construction

To improve diagnostic performance and control scanning time to reduce the waste of medical resources, an optimal multimodal radiomics signature was constructed in the following three steps: (1) single-sequence radiomics signatures: ^{18}F -FDG, T1WI, T2WI, T2/FLAIR, SWI, and DWI, respectively; (2) double-sequence radiomics signatures: ^{18}F -FDG plus structural MRI (sMRI) sequences (^{18}F -FDG + T1WI, ^{18}F -FDG + T2WI, ^{18}F -FDG + T2/FLAIR), or ^{18}F -FDG plus functional MRI (fMRI) sequences (^{18}F -FDG + SWI, ^{18}F -FDG + DWI); (3) optimal multimodal radiomics signatures: ^{18}F -FDG + the best sMRI + the best fMRI, according to the results from the first and second steps.

We employed the same strategy of feature selection and model construction above all three steps. The datasets were randomly divided into a training set and a validation set with a case number ratio of 7:3. The minimal redundancy maximal relevance (mRMR) algorithm, which can considerably improve the accuracy of feature selection and classification [35], was performed for initial feature selection in the training set. The least absolute shrinkage and selection operator (LASSO) method, which is suitable for the regression of high-dimensional data, was used to select significant distinguishable features to construct the radiomics signature with 10-fold cross-validation.

Clinical-radiomics diagnostic model construction

The clinical variables, including clinical characteristics (age, sex, weight, pre-injection glucose levels, disease duration [DD], age at onset, hypermyotonia, asymmetric symptoms at onset, bradykinesia, limbs tremor, dysarthria, and autonomic failure [AF]) and SUV values, were collected and compared between PD and MSA in training and validation sets. The clinical-radiomics model was constructed by combining the clinical variables with the optimal multimodal radiomics signature obtained above. The clinical-radiomics model was built by a multivariate logistic regression model with 10-fold cross-validation to distinguish PD and MSA through a likelihood ratio test with back-ward step-down. A nomogram was then constructed on the basis of the clinical-radiomics model.

Model effectiveness evaluation

The area under the receiver-operating characteristic (ROC) curve (AUC) was used to evaluate the diagnostic performance of models constructed by the training set and validated by the validation set, whereby the radiomics score (Radscore) was calculated via the formula built in the training set. The accuracy of the radiomics signature was evaluated in both the training and validation sets. The models' calibration was assessed using calibration curves and the Hosmer–Lemeshow test; decision curve analysis (DCA) was performed to estimate the clinical utility of models.

Statistical analysis

Statistical analysis was performed by R 3.6.1 (www.Rproject.org). The packages in R used in this study were tidyverse, caret, pROC, glmnet, DMWR, rmda, ggpubr, ModelGood, rms, mRMRe, DescTOOLS and irr. The Delong test was applied to compare the differences in ROC curves between two arbitrary models by Medcalc (www.medcalc.org).

The differences in demographic and clinical variables were compared between patients with PD and MSA in both the training set and validation set by Graphpad prism 8 (www.graphpad-prism.cn). The Mann-Whitney U -test was used for non-normally distributed quantitative data; for normally distributed data, the independent sample t -test was used. Chi-squared tests were used for categorical data.

Results

Characteristics of the patients and clinical features

Demographic and clinical variables of the patients in the training and validation sets are summarised in **Table 1**. All the demographic data, including the age, sex, DD and age at onset, did not show significant differences between PD and MSA patients ($P > 0.05$). In the training group, the symptoms of age at onset, dysarthria, AF, and SUV_{max} showed statistically significant differences between PD and MSA ($P < 0.05$); but, except for dysarthria, were not statistically significantly different in the validation group.

Diagnostic performance of radiomics signatures

A total of 1127 features (first order statistics, shape-based texture, grey-level cooccurrence matrix, grey-level run length matrix, grey-level size zone matrix, grey-level dependence matrix, neighbouring grey tone difference matrix, and the details are provided in Online Resource 1) per sequence were extracted in all six PET/MRI sequences images. The intra-observer ICCs all indicated favourable feature extraction reproducibility (ICC > 0.7).

First, we investigated and compared the six single-sequence models of ^{18}F -FDG, T1WI, T2WI, T2/FLAIR, SWI and DWI separately. All the radiomics signatures showed significant differences between PD and MSA ($p < 0.05$, **Fig. 2a**) in the training cohort and validation cohort, with the AUCs > 0.8 and most ACCs > 0.7 (**Fig. 2d and e**). The equations of six single-sequence Radscores and more details of ROC analysis are presented in **Supplementary materials (Supple. Equation 1–6, Supple. Table 1)**. Although there were no statistically significant differences among individual Radscores of single-sequence models either in the training set or validation set ($P > 0.05$, **Fig. 2f and g**), ^{18}F -FDG had the highest diagnostic performance in the training set and

validation set (AUC = 0.900 vs 0.833, **Fig. 2d and e**). Therefore, we chose ^{18}F -FDG to combine with the other MRI sequences to optimize the models next step, which was consistent with our multi-modal radiomics design.

Next, we combined ^{18}F -FDG with one of the MRI sequences for bimodal radiomics optimisation. Five double-sequence composite models all demonstrated high diagnostic efficiency with AUCs > 0.9 in training set and > 0.8 in validation set (**Table 2**). In the analysis of the combinations of ^{18}F -FDG and sMRI, three models were constructed: (1) $\text{Radscore}_{\text{FDG_T1WI}}$; (2) $\text{Radscore}_{\text{FDG_T2WI}}$; (3) $\text{Radscore}_{\text{FDG_T2/FLAIR}}$. The $\text{Radscore}_{\text{FDG_T1WI}}$ had the highest AUC (training set: AUC = 0.958, 95% confidence interval [CI] = 0.875–0.992; validation set: AUC = 0.932, 95% CI = 0.766–0.993, **Table 2**). In the groups combining ^{18}F -FDG with fMRI, two models were built: (1) $\text{Radscore}_{\text{FDG_SWI}}$; (2) $\text{Radscore}_{\text{FDG_DWI}}$. The $\text{Radscore}_{\text{FDG_SWI}}$ showed a larger AUC (training set: AUC = 0.951, 95% CI = 0.865–0.950; validation set: AUC = 0.951, 95% CI = 0.792–0.997, **Table 2**). The equations of five double-sequence Radscores are presented in **Supplementary materials (Supple. Equation 7–11)**. No significant difference ($P > 0.05$) was detected when the five composite models were compared with each other, but the differences were statistically significant when comparing the double-sequence models with the single-sequence models using the DeLong test ($P < 0.05$).

Finally, we attempted to combined ^{18}F -FDG with T1WI (the best sMRI) and SWI (the best fMRI) to build a more optimized model. After feature dimension reduction, 10 features were used to build $\text{Radscore}_{\text{FDG_T1WI_SWI}}$ (**Fig. 3c**), when $\log(\lambda)$ was 0.032 (**Fig. 3a and b**). The features showed no significant collinearity which indicated that $\text{Radscore}_{\text{FDG_T1WI_SWI}}$ avoided overfitting (**Supplementary materials, Supple. Fig. 1**). The equation of $\text{Radscore}_{\text{FDG_T1WI_SWI}}$ was

The $\text{Radscore}_{\text{FDG_T1WI_SWI}}$ yielded an AUC of 0.971 (95% CI, 0.893–0.997) in the training set and 0.957 (95% CI, 0.801–0.998) in the validation set (**Table 3, Fig. 3d and e**), which was indicated statistical significance in comparison to arbitrary double-sequence composite models ($P < 0.05$).

Clinical-radiomics integrated model

In combination with clinical variables, we further constructed an integrated diagnosis model by logistic regression. The SUV_{max} and symptoms of DD, dysarthria, and AF were identified as independent factors in the clinical-radiomics integrated model by a multivariate logistic regression model. The variance inflation factors (VIFs) of the four clinical features were 1.036, 1.376, 1.217 and 1.087. The nomogram based on clinical factors and Radscore is shown in **Fig. 4a**. The Hosmer–Lemeshow test in the clinical-radiomics integrated model showed no significant differences in the goodness-of-fit for the training set ($P = 0.9862$).

We also evaluated the discriminatory efficiency of the clinical model and clinical-radiomics integrated model using ROC analyses (**Table 3, Fig. 4b and c**). The integrated model yielded the largest AUC of 0.993 (95% CI, 0.930–1.000) in the training group and 0.994 (95% CI, 0.861–1.000) in the validation group, which showed a statistically significant difference when compared with clinical models (training set: AUC = 0.895, 95% CI = 0.790–0.958; validation set: AUC = 0.858, 95% CI = 0.670–0.961; $P = 0.012$), but no statistical difference in comparison with $\text{Radscore}_{\text{FDG_T1WI_SWI}}$ ($p = 0.098$). At the same time, $\text{Radscore}_{\text{FDG_T1WI_SWI}}$ showed significant differences with single- and double-sequence Radscores ($p < 0.05$).

The DCAs for the clinical model, $\text{Radscore}_{\text{FDG_T1WI_SWI}}$, and clinical-radiomics integrated model in the training and validation sets are shown in **Fig. 4f**. The DCA indicated the threshold probability, in the range of 0 to 1, that could benefit from clinical-radiomics integrated model.

Discussion

In the present study, we progressively optimised the radiomics models from single- and double-sequences to multi-modality, and obtained an optimal multimodal radiomics signature to distinguish between PD and MSA using a hybrid PET/MRI, which could provide the metabolic, structural and functional information at the same time and space. To the best of our knowledge, this is the first study of 'radiomics' on a hybrid PET/MRI in the area of neurodegenerative diseases. Besides, the joint analysis of radiomics and clinical features may be of great significance in the differential diagnosis of other indistinguishable diseases.

The radiomics models were progressively optimised in three steps. In the monomodal radiomics analysis, the diagnostic efficiency of ^{18}F -FDG PET was better than that of five different MRI sequences both in the training (AUC = 0.900) and validation sets (AUC = 0.883), though there were no statistically significant differences between them. In fact, many published studies have shown that normal or slightly elevated ^{18}F -FDG uptakes were found in the striata of early PD patients [13, 36]; however, hypometabolism in the putamen, pons, and cerebellum has been demonstrated in different types of MSA [13]. The high diagnostic efficiency of ^{18}F -FDG is in accordance with the previous research by Brajkovic et al., whose diagnostic accuracy of ^{18}F -FDG PET was 93% for distinguishing PD and MSA [37]. Based on the fact that functional metabolic changes are often observed before structural changes, and nuances changes cannot regularly be observed by the naked eye, radiomics provides a highly sensitive opportunity to estimate the distribution of metabolism at the microscopic level [38]. Notably, no difference between ^{18}F -FDG and MRI sequences means that MRI sequences also have abilities to distinguish between PD and MSA, although it was indicated that ^{18}F -FDG PET tends better diagnostic ability compared with brain MRI; which is consistent with the research performed by Kwon et al. [11].

We combined ^{18}F -FDG PET with one MRI sequence for bimodal radiomics analysis in the next step for advanced model optimisation. The $\text{Radscore}_{\text{FDG_T1WI}}$ (training set: AUC = 0.958; validation set: AUC = 0.932) and $\text{Radscore}_{\text{FDG_SWI}}$ (training set: AUC = 0.951; validation set: AUC = 0.951) displayed the highest diagnostic efficiency in the combination of ^{18}F -FDG with sMRI and fMRI, respectively. Compared with PD, putaminal or pontocerebellar atrophy and a rim sign can be seen in typical MSA [7], where very thin T1WI sections are important for anatomical details [7]. The ROIs were sketched based on the T1WI images (section thickness 1 mm) in our study to reflect accurate anatomical structural information. Furthermore, many studies have shown that iron deposition plays an important role in neurodegenerative diseases [39-43], and higher iron deposition has been observed in the putamen and pulvinar thalamus of MSA compared with PD [39, 44]. SWI is based on a T2*-weighted gradient echo sequence, which is sensitive to iron deposition. Kraft et al. reported that hypointense putamina signal intensity was more often observed in MSA than in PD by using T2*-weighted sequences [45], which could improve diagnostic accuracy in distinguishing between PD and MSA. Nevertheless, the performance of $\text{Radscore}_{\text{FDG_ADC}}$ was satisfactory, for increased putaminal diffusivity in

some MSA patients compared with PD [7]. The lack of ADC images in the follow-up studies doesn't mean no sense it makes, just for the better selection (SWI) to reduce waste of health care resources.

In the third step, an optimised multimodal radiomics signature (training set: AUC = 0.971; validation set: AUC = 0.957) was developed by combining ^{18}F -FDG with T1WI (the best sMRI) and SWI (the best fMRI), which was significantly better than single-sequence and double-sequence models ($P < 0.05$). Interestingly, the AUC was increased whenever adding another sequence to construct a new combinatorial model. This means that multimodal imaging can achieve more accurate diagnostic capability by combining the synergistic advantages of different imaging modalities, and some studies also confirmed superior diagnostic or predictive accuracy with MRI and PET or SPECT [26, 27, 46]. However, there are some methodological challenges in multimodal imaging, such as image alignment problems and the risk of overfitting due to increased data [47]. In our study, ^{18}F -FDG-PET and MRI images were acquired using a hybrid PET/MRI system at the same time and space, which allowed for more accurate spatial alignment and reduced ROI errors to make the results more reliable. Additionally, we used 10-fold cross-validation to prevent overfitting as much as possible, which made for trustworthy models.

We further constructed the clinical-radiomics integrated model with the best diagnostic discrimination (training set: AUC = 0.993; validation set: AUC = 0.994) by combining the optimal $\text{Radscore}_{\text{FDG_T1WI_SWI}}$ with clinical features. As the selected clinical characteristics, shorter disease duration (rapid disease progression) and more severe symptoms of dysarthria and autonomic failure are more common in MSA patients [31]. However, it must be mentioned that the diagnostic performance and clinical utility of a purely clinical model were not satisfactory and that the difference was not significant ($p = 0.098$) between the clinical-radiomics integrated model and $\text{Radscore}_{\text{FDG_T1WI_SWI}}$. But it is still important to incorporate clinical characteristics within radiomics analysis, for which imaging is intended to complement and supplement, but not replace, clinical decision making. The clinical symptoms are important basics for differential diagnosis of PD and MSA, although they overlap in the early stages [7, 11]. A higher diagnostic efficacy also proves that multidisciplinary approaches, including radiomics analysis, make sense for the diagnosis and treatment of diseases [48].

Notably, we chose ^{18}F -FDG over dopamine system-related agents like the dopamine transporter (DAT) and the dopamine D2 receptors, because striatal DAT loss may overlap in most PD and MSA patients [49, 50]. Also, DAT imaging can only be used for a diagnosis of exclusion [51-53]. Compared with ^{18}F -FDG, dopamine-related agents are not widely used, as the conditions for manufacturing applications are more stringent.

There are some limitations to our study. First, there were no pathology results as a gold standard for PD and MSA diagnosis. We classified patients into PD and MSA groups strictly based on standard official diagnostic guidelines, and guaranteed the mean follow-up period was > 1.5 years, to ensure an accurate diagnosis of the two diseases. Second, the amount of data is an ongoing topic in radiomics research. Larger sample size will make the model more stable and give more reliable results. Therefore, we did not subtype the MSA patients in this study because motor symptoms occur sooner or later in any subtype of MSA [31]. However, subsequent studies could divide MSA patients into predominant parkinsonism (MSA-P) and predominant cerebellar ataxia (MSA-C) groups and add other APS such as progressive supranuclear palsy and corticobasal ganglionic degeneration. Additionally, multicentre cooperation and database establishment is preferred for further model testing. Finally, manually delineating ROIs is time-consuming and inevitably is accompanied by operator

dependence. Automatic sketching methods and brain zoning studies may achieve precise automatic segmentation in the future, and artificial intelligence screening is our ultimate goal.

Conclusion

The multivariable radiomics models based on the hybrid ^{18}F -FDG PET/MRI, which provides metabolic, structural, and functional information simultaneously, can be applied for the comparatively accurate identification of PD and MSA; and a clinical-radiomics integrated model displayed better performance.

Declarations

Acknowledgments

We thank Prof. Ping Han for her comments of revision, our friend Ling Ma for her help in comments of statistical analysis, and Libby Cone, MD, MA, from Liwen Bianji, Edanz Editing China (www.liwenbianji.cn/ac) for editing a draft of this manuscript.

Funding information

This work was supported by the National Natural Science Foundation of China (No. 81701759), the Key Project of Hubei Province Technical Innovation (2017ACA182), and the Clinical Research Physician Program of Tongji Medical College, Huazhong University of Science and Technology (No. 5001530008).

Conflicts of interest

The authors declare that they have no conflict of interest. **Ethics approval**

This study is a retrospective study based on data from one of our clinical studies, in which all patients signed informed consent forms and the ethical approval was obtained by the Ethics Committee of Tongji Medical College, Huazhong University of Science and Technology.

References

1. Stoessl AJ, Martin WW, McKeown MJ, Sossi V. Advances in imaging in Parkinson's disease. *Lancet Neurol.* 2011;10:987-1001. doi:10.1016/S1474-4422(11)70214-9.
2. Saeed U, Compagnone J, Aviv RI, Strafella AP, Black SE, Lang AE, et al. Imaging biomarkers in Parkinson's disease and Parkinsonian syndromes: current and emerging concepts. *Transl Neurodegener.* 2017;6:8. doi:10.1186/s40035-017-0076-6.
3. Adler CH, Beach TG, Hentz JG, Shill HA, Caviness JN, Driver-Dunckley E, et al. Low clinical diagnostic accuracy of early vs advanced Parkinson disease: clinicopathologic study. *Neurology.* 2014;83:406-12. doi:10.1212/WNL.0000000000000641.
4. Peralta C, BIAFORE F, Depetris TS, Bastianello M. Recent Advancement and Clinical Implications of ^{18}F FDG-PET in Parkinson's Disease, Atypical Parkinsonisms, and Other Movement Disorders. *Curr Neurol Neurosci Rep.* 2019;19:56. doi:10.1007/s11910-019-0966-3.

5. Fanciulli A, Wenning GK. Multiple-system atrophy. *N Engl J Med*. 2015;372:249-63. doi:10.1056/NEJMr1311488.
6. Zhou HY, Huang P, Sun Q, Du JJ, Cui SS, Hu YY, et al. The role of substantia nigra sonography in the differentiation of Parkinson's disease and multiple system atrophy. *Transl Neurodegener*. 2018;7:15. doi:10.1186/s40035-018-0121-0.
7. Berardelli A, Wenning GK, Antonini A, Berg D, Bloem BR, Bonifati V, et al. EFNS/MDS-ES/ENS [corrected] recommendations for the diagnosis of Parkinson's disease. *Eur J Neurol*. 2013;20:16-34. doi:10.1111/ene.12022.
8. Mahlknecht P, Hotter A, Hussl A, Esterhammer R, Schocke M, Seppi K. Significance of MRI in Diagnosis and Differential Diagnosis of Parkinson's Disease. *Neurodegener Dis*. 2010;7:300-18. doi:10.1159/000314495.
9. Peran P, Barbagallo G, Nemmi F, Sierra M, Galitzky M, Traon AP, et al. MRI supervised and unsupervised classification of Parkinson's disease and multiple system atrophy. *Mov Disord*. 2018;33:600-8. doi:10.1002/mds.27307.
10. Chen B, Fan G, Sun W, Shang X, Shi S, Wang S, et al. Usefulness of diffusion-tensor MRI in the diagnosis of Parkinson variant of multiple system atrophy and Parkinson's disease: a valuable tool to differentiate between them? *Clin Radiol*. 2017;72:610 e9- e15. doi:10.1016/j.crad.2017.02.005.
11. Kwon KY, Choi CG, Kim JS, Lee MC, Chung SJ. Comparison of brain MRI and 18F-FDG PET in the differential diagnosis of multiple system atrophy from Parkinson's disease. *Mov Disord*. 2007;22:2352-8. doi:10.1002/mds.21714.
12. Baudrexel S, Seifried C, Penndorf B, Klein JC, Middendorp M, Steinmetz H, et al. The value of putaminal diffusion imaging versus 18-fluorodeoxyglucose positron emission tomography for the differential diagnosis of the Parkinson variant of multiple system atrophy. *Mov Disord*. 2014;29:380-7. doi:10.1002/mds.25749.
13. Meyer PT, Frings L, Rucker G, Hellwig S. (18)F-FDG PET in Parkinsonism: Differential Diagnosis and Evaluation of Cognitive Impairment. *J Nucl Med*. 2017;58:1888-98. doi:10.2967/jnumed.116.186403.
14. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology*. 2016;278:563-77. doi:10.1148/radiol.2015151169.
15. Sun R, Limkin EJ, Vakalopoulou M, Derclé L, Champiat S, Han SR, et al. A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. *Lancet Oncol*. 2018;19:1180-91. doi:10.1016/S1470-2045(18)30413-3.
16. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol*. 2017;14:749-62. doi:10.1038/nrclinonc.2017.141.
17. Huang YQ, Liang CH, He L, Tian J, Liang CS, Chen X, et al. Development and Validation of a Radiomics Nomogram for Preoperative Prediction of Lymph Node Metastasis in Colorectal Cancer. *J Clin Oncol*. 2016;34:2157-64. doi:10.1200/JCO.2015.65.9128.
18. Xu X, Zhang HL, Liu QP, Sun SW, Zhang J, Zhu FP, et al. Radiomic analysis of contrast-enhanced CT predicts microvascular invasion and outcome in hepatocellular carcinoma. *J Hepatol*. 2019;70:1133-44. doi:10.1016/j.jhep.2019.02.023.

19. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. 2012;48:441-6. doi:10.1016/j.ejca.2011.11.036.
20. Tang Y, Meng L, Wan CM, Liu ZH, Liao WH, Yan XX, et al. Identifying the presence of Parkinson's disease using low-frequency fluctuations in BOLD signals. *Neuroscience Letters*. 2017;645:1-6. doi:10.1016/j.neulet.2017.02.056.
21. Long D, Wang JW, Xuan M, Gu QQ, Xu XJ, Kong DX, et al. Automatic Classification of Early Parkinson's Disease with Multi-Modal MR Imaging. *Plos One*. 2012;7. doi:ARTN e47714 10.1371/journal.pone.0047714.
22. Chen Y, Storrs J, Tan LR, Mazlack LJ, Lee JH, Lu LJ. Detecting brain structural changes as biomarker from magnetic resonance images using a local feature based SVM approach. *J Neurosci Meth*. 2014;221:22-31. doi:10.1016/j.jneumeth.2013.09.001.
23. Adeli E, Shi F, An L, Wee CY, Wu GR, Wang T, et al. Joint feature-sample selection and robust diagnosis of Parkinson's disease from MRI data. *Neuroimage*. 2016;141:206-19. doi:10.1016/j.neuroimage.2016.05.054.
24. Focke NK, Helms G, Scheewe S, Pantel PM, Bachmann CG, Dechent P, et al. Individual Voxel-Based Subtype Prediction can Differentiate Progressive Supranuclear Palsy from Idiopathic Parkinson Syndrome and Healthy Controls. *Human Brain Mapping*. 2011;32:1905-15. doi:10.1002/hbm.21161.
25. Martinez-Murcia FJ, Gorriz JM, Ramirez J, Moreno-Caballero M, Gomez-Rio M, Parkinsons Progression Markers I. Parametrization of textural patterns in I-123-ioflupane imaging for the automatic detection of Parkinsonism. *Medical Physics*. 2014;41. doi:10.1118/1.4845115.
26. Klyuzhin IS, Gonzalez M, Shahinfard E, Vafai N, Sossi V. Exploring the use of shape and texture descriptors of positron emission tomography tracer distribution in imaging studies of neurodegenerative disease. *Journal of Cerebral Blood Flow and Metabolism*. 2016;36:1122-34. doi:10.1177/0271678x15606718.
27. Rahmim A, Salimpour Y, Jain S, Blinder SAL, Klyuzhin IS, Smith GS, et al. Application of texture analysis to DAT SPECT imaging: Relationship to clinical assessments. *Neuroimage-Clinical*. 2016;12:E1-E9. doi:10.1016/j.nicl.2016.02.012.
28. Choi H, Cheon GJ, Kim HJ, Choi SH, Kim YI, Kang KW, et al. Gray matter correlates of dopaminergic degeneration in Parkinson's disease: A hybrid PET/MR study using (18) F-FP-CIT. *Hum Brain Mapp*. 2016;37:1710-21. doi:10.1002/hbm.23130.
29. Barthel H, Schroeter ML, Hoffmann KT, Sabri O. PET/MR in dementia and other neurodegenerative diseases. *Semin Nucl Med*. 2015;45:224-33. doi:10.1053/j.semnuclmed.2014.12.003.
30. Postuma RB, Berg D, Stern M, Poewe W, Olanow CW, Oertel W, et al. MDS clinical diagnostic criteria for Parkinson's disease. *Mov Disord*. 2015;30:1591-601. doi:10.1002/mds.26424.
31. Gilman S, Wenning GK, Low PA, Brooks DJ, Mathias CJ, Trojanowski JQ, et al. Second consensus statement on the diagnosis of multiple system atrophy. *Neurology*. 2008;71:670-6. doi:10.1212/01.wnl.0000324625.00404.15.
32. Li G, Zhai G, Zhao X, An H, Spincemaille P, Gillen KM, et al. 3D texture analyses within the substantia nigra of Parkinson's disease patients on quantitative susceptibility maps and R2*maps. *Neuroimage*. 2019;188:465-72. doi:10.1016/j.neuroimage.2018.12.041.

33. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res.* 2017;77:e104-e7. doi:10.1158/0008-5472.CAN-17-0339.
34. Zwanenburg A, Leger S, Vallières M, Löck S, Initiative f. Image biomarker standardisation initiative - feature definitions. 2016.
35. Peng H, Long F, Ding C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans Pattern Anal Mach Intell.* 2005;27:1226-38. doi:10.1109/TPAMI.2005.159.
36. Eckert T, Barnes A, Dhawan V, Frucht S, Gordon MF, Feigin AS, et al. FDG PET in the differential diagnosis of parkinsonian disorders. *Neuroimage.* 2005;26:912-21. doi:10.1016/j.neuroimage.2005.03.012.
37. Brajkovic L, Kostic V, Sobic-Saranovic D, Stefanova E, Jecmenica-Lukic M, Jesic A, et al. The utility of FDG-PET in the differential diagnosis of Parkinsonism. *Neurol Res.* 2017;39:675-84. doi:10.1080/01616412.2017.1312211.
38. Xu J, Zhang M. Use of Magnetic Resonance Imaging and Artificial Intelligence in Studies of Diagnosis of Parkinson's Disease. *ACS Chem Neurosci.* 2019;10:2658-67. doi:10.1021/acchemneuro.9b00207.
39. Wang Y, Butros SR, Shuai X, Dai Y, Chen C, Liu M, et al. Different iron-deposition patterns of multiple system atrophy with predominant parkinsonism and idiopathic Parkinson diseases demonstrated by phase-corrected susceptibility-weighted imaging. *AJNR Am J Neuroradiol.* 2012;33:266-73. doi:10.3174/ajnr.A2765.
40. Lee JH, Lee MS. Brain Iron Accumulation in Atypical Parkinsonian Syndromes: in vivo MRI Evidences for Distinctive Patterns. *Front Neurol.* 2019;10:74. doi:10.3389/fneur.2019.00074.
41. Wang JY, Zhuang QQ, Zhu LB, Zhu H, Li T, Li R, et al. Meta-analysis of brain iron levels of Parkinson's disease patients determined by postmortem and MRI measurements. *Scientific reports.* 2016;6:36669. doi:10.1038/srep36669.
42. Meyer E, Kurian MA, Hayflick SJ. Neurodegeneration with Brain Iron Accumulation: Genetic Diversity and Pathophysiological Mechanisms. *Annu Rev Genomics Hum Genet.* 2015;16:257-79. doi:10.1146/annurev-genom-090314-025011.
43. Healy S, McMahon JM, FitzGerald U. Modelling iron mismanagement in neurodegenerative disease in vitro: paradigms, pitfalls, possibilities & practical considerations. *Prog Neurobiol.* 2017;158:1-14. doi:10.1016/j.pneurobio.2017.08.004.
44. Ramli N, Nair SR, Ramli NM, Lim SY. Differentiating multiple-system atrophy from Parkinson's disease. *Clin Radiol.* 2015;70:555-64. doi:10.1016/j.crad.2015.01.005.
45. Kraft E, Trenkwalder C, Auer DP. T2*-weighted MRI differentiates multiple system atrophy from Parkinson's disease. *Neurology.* 2002;59:1265-7. doi:10.1212/01.wnl.0000032757.66992.3c.
46. Rahmim A, Huang P, Shenkov N, Fotouhi S, Davoodi-Bojd E, Lu L, et al. Improved prediction of outcome in Parkinson's disease using radiomics analysis of longitudinal DAT SPECT images. *Neuroimage-Clinical.* 2017;16:539-44. doi:10.1016/j.nicl.2017.08.021.
47. Teipel S, Drzezga A, Grothe MJ, Barthel H, Chetelat G, Schuff N, et al. Multimodal imaging in Alzheimer's disease: validity and usefulness for early detection. *Lancet Neurol.* 2015;14:1037-53. doi:10.1016/S1474-4422(15)00093-9.

48. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun.* 2014;5:4006. doi:10.1038/ncomms5006.
49. Jin S, Oh M, Oh SJ, Oh JS, Lee SJ, Chung SJ, et al. Differential Diagnosis of Parkinsonism Using Dual-Phase F-18 FP-CIT PET Imaging. *Nucl Med Mol Imaging.* 2013;47:44-51. doi:10.1007/s13139-012-0182-4.
50. Fazio P, Svenningsson P, Forsberg A, Jonsson EG, Amini N, Nakao R, et al. Quantitative Analysis of (1)(8)F-(E)-N-(3-Iodoprop-2-Enyl)-2beta-Carbofluoroethoxy-3beta-(4'-Methyl-Phenyl) Nortropane Binding to the Dopamine Transporter in Parkinson Disease. *J Nucl Med.* 2015;56:714-20. doi:10.2967/jnumed.114.152421.
51. Takaya S, Sawamoto N, Okada T, Okubo G, Nishida S, Togashi K, et al. Differential diagnosis of parkinsonian syndromes using dopamine transporter and perfusion SPECT. *Parkinsonism Relat Disord.* 2018;47:15-21. doi:10.1016/j.parkreldis.2017.11.333.
52. Bu LL, Liu FT, Jiang CF, Guo SS, Yu H, Zuo CT, et al. Patterns of dopamine transporter imaging in subtypes of multiple system atrophy. *Acta Neurol Scand.* 2018;138:170-6. doi:10.1111/ane.12932.
53. Brooks DJ. Molecular imaging of dopamine transporters. *Ageing Res Rev.* 2016;30:114-21. doi:10.1016/j.arr.2015.12.009.

Tables

Table 1 Demographic and clinical characteristics of patients in the training and validation groups

	Training group			Validation group		
	PD	MSA	P value	PD	MSA	P value
Age, year	56.5 ± 9.8	55.5 ± 8.5	0.584	52.5 ± 10.0	60.0 ± 9.0	0.182
Sex			0.216			0.705
Female	27	10		10	6	
Male	15	11		8	3	
Weight, kg	62.3 ± 10.8	64.4 ± 9.3	0.293	59.9 ± 10.2	66.7 ± 12.5	0.380
Pre-injection glucose, mmol/L	5.9 (0.7)	5.8 (1.2)	0.313	6.0 (0.8)	5.9 (1.2)	0.940
Disease duration, year	2.0 (2.0)	1.5 (1.0)	0.084	2.3 (2.0)	1.5 (1.9)	0.592
Age at onset, year	54.3 ± 10.1	53.9 ± 8.3	0.688	50.3 ± 9.9	58.3 ± 8.6	0.129
Hypermyotonia	81.0% (34/42)	61.9% (13/21)	0.113	88.9% (16/18)	66.7% (6/9)	0.384
Asymmetric symptoms at onset	83.3% (35/42)	38.1% (8/21)	0.0004*	72.2% (13/18)	28.6% (2/9)	0.096
Bradykinesia	73.8% (31/42)	85.7% (18/21)	0.302	66.7% (12/18)	88.9% (8/9)	0.384
Limbs tremor	69.0% (29/42)	57.1% (12/21)	0.363	72.2% (13/18)	11.1% (1/9)	0.039*
Dysarthria	9.5% (4/42)	52.4% (11/21)	0.0003*	16.7% (3/18)	77.8% (7/9)	0.042*
Autonomic failure	7.1% (3/42)	52.4% (11/21)	0.0001*	5.6% (1/18)	44.4% (4/9)	0.142
SUV						
Maximum	14.28 ± 3.99	11.48 ± 3.75	0.017*	12.42 ± 2.64	12.36 ± 3.21	0.756
Mean	7.96 ± 2.73	7.08 ± 2.77	0.545	6.42 ± 1.97	8.07 ± 1.60	0.021*
Minimum	2.53 (1.26)	2.91 (1.22)	0.560	2.24 (0.87)	3.14 (0.61)	0.144

PD, Parkinson's disease; MSA, Multiple System Atrophy

The data are presented as mean ± SD or median (IQR)

* the clinical features of PD and MSA showed statistically significant differences

Table 2 Diagnostic performance of radiomics signatures in double-sequence models

	PET_sMRI						PET_fmRI			
	FDG_T1WI Radscore		FDG_T2/Flair Radscore		FDG_T2WI Radscore		FDG_SWI Radscore		FDG_ADC Radscore	
	Training group	Validation group	Training group	Validation group	Training group	Validation group	Training group	Validation group	Training group	Validation group
re	1.505 ± 3.080	0.254 ± 4.928	1.047 ± 1.667	0.833 ± 2.166	0.834 ± 1.048	0.509 ± 1.003	1.024 ± 1.714	0.840 ± 1.181	1.083 ± 1.918	0.733 ± 2.134
y	0.958	0.932	0.941	0.926	0.895	0.889	0.951	0.951	0.932	0.920
	0.841	0.815	0.937	0.667	0.889	0.889	0.873	0.741	0.921	0.667
	0.809	0.778	0.810	0.778	0.643	0.722	0.833	0.778	0.810	0.834
	0.875	0.766	0.851	0.757	0.791	0.833	0.865	0.792	0.839	0.749
r	0.992	0.993	0.985	0.991	0.958	0.889	0.950	0.997	0.980	0.989
ity	0.945	0.882	0.975	0.764	0.907	0.857	0.972	0.823	0.951	0.714
ity	0.704	0.700	0.869	0.500	0.850	1.000	0.741	0.600	0.864	0.500
	0.810	0.833	0.928	0.722	0.928	1.000	0.833	0.778	0.929	0.833
	0.905	0.777	0.952	0.556	0.809	0.666	0.952	0.667	0.905	0.333
	12	12	11	11	6	6	11	11	10	10

PET, positron emission tomography; sMRI, structural magnetic resonance imaging; fmRI, structural magnetic resonance imaging; FDG, ¹⁸F-fluorodeoxyglucose; T1WI, T1-weighted imaging; T2WI, T2-weighted imaging; T2/FLAIR, T2-weighted fluid-attenuated inversion recovery imaging; SWI, susceptibility-weighted imaging; ADC, apparent diffusion coefficient; AUC, area under the curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value

Radscore data are presented as mean ± SD

Table 3 Diagnostic performance of radiomics signatures in combinational models

	FDG_SWI_T1WI Radscore		Clinical model		Clinical-radiomics integrated model	
	Training group	Validation group	Training group	Validation group	Training group	Validation group
AUC	0.971	0.957	0.895	0.858	0.993	0.994
Accuracy	0.920	0.889	0.809	0.704	0.952	0.963
Youden	0.857	0.833	0.667	0.667	0.923	0.944
95% CI						
lower	0.893	0.801	0.791	0.670	0.930	0.861
Upper	0.997	0.998	0.958	0.961	1.000	1.000
Sensitivity	0.905	0.833	0.762	0.889	0.929	0.944
Specificity	0.952	1.000	0.905	0.778	1.000	1.000

FDG, ¹⁸F-fluorodeoxyglucose; T1WI, T1-weighted imaging; SWI, susceptibility-weighted imaging; AUC, area under the curve; CI, confidence interval

Figures

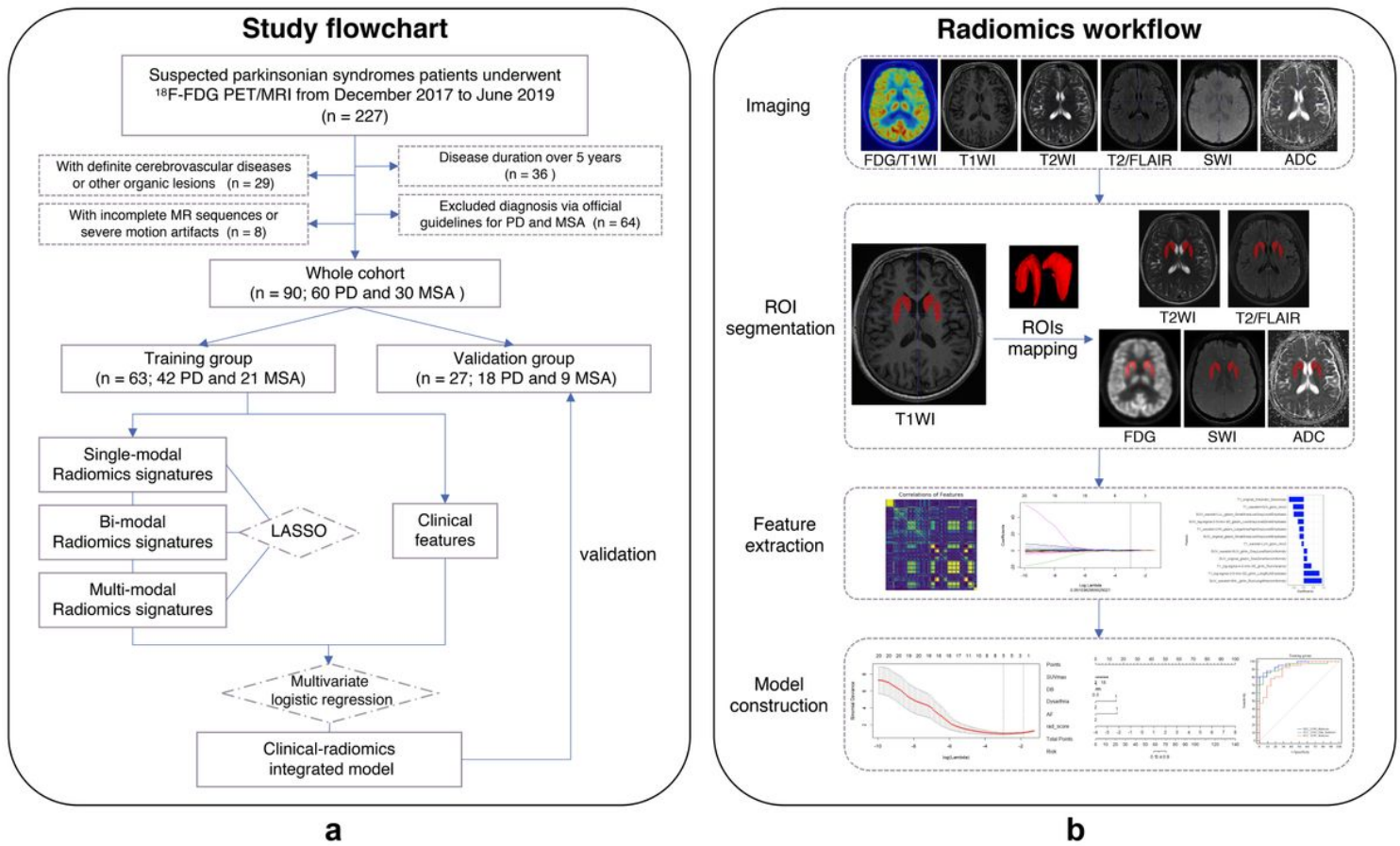


Figure 1

(a) study flowchart and (b) radiomics workflow

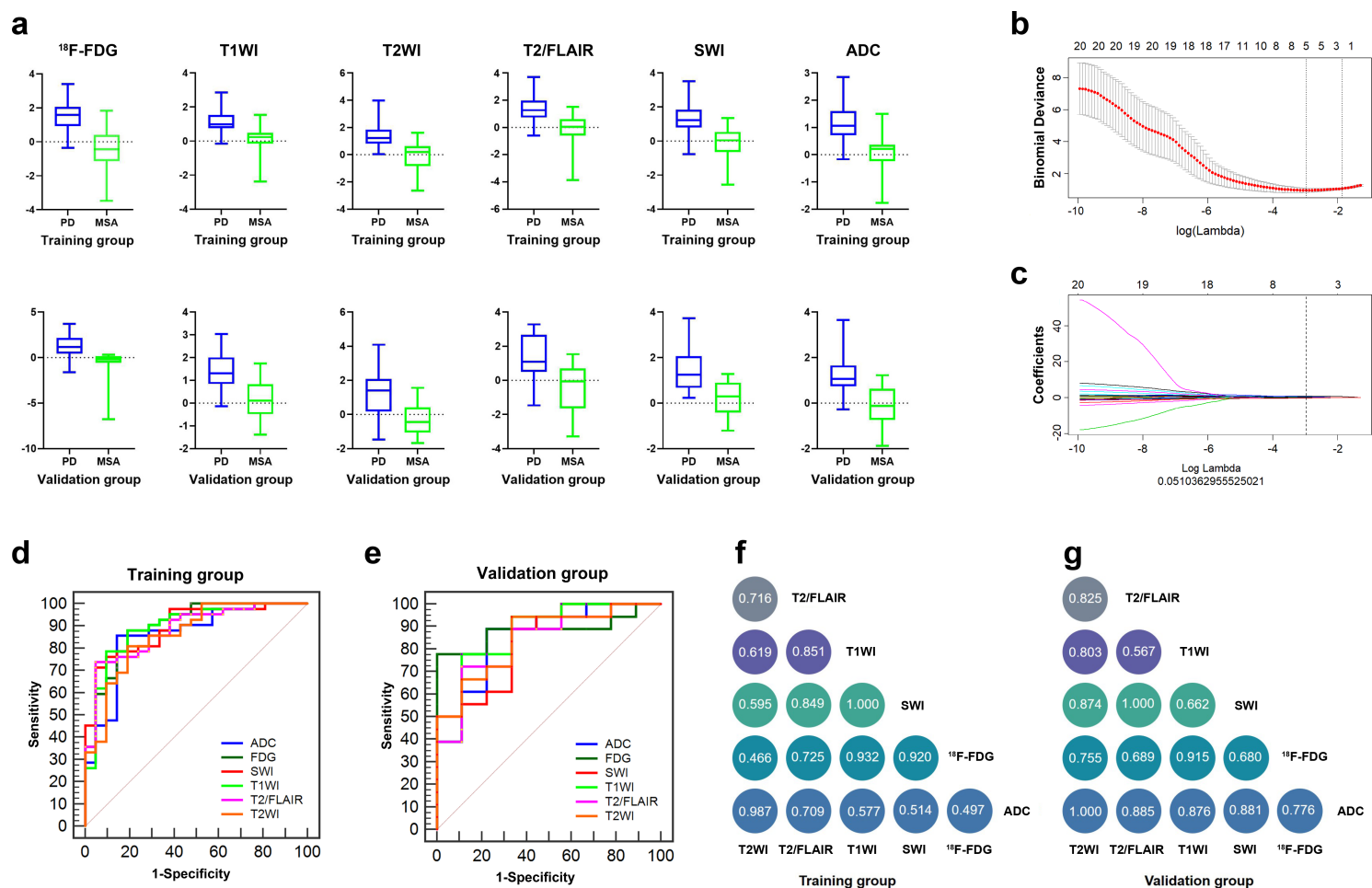


Figure 2

Diagnostic performance and DeLong test of radiomics signatures in single-sequence models. (a) The radiomics signatures of all six sequences showed significant differences between PD and MSA ($p < 0.05$) in the training cohort and validation cohort. (b) Texture reduction and selection was using LASSO based on the minimum I with 10-fold cross validation. Y-axis indicates binomial deviances, and X-axis means the number of features. (c) The value of I with number of features. (d) The ROC curves of single-sequence models in training set and (e) validation set, and FDG performed best (AUC of training set and validation set, 0.971 vs 0.957). (f) No statistically significant differences ($P > 0.05$) among Radscores of single-sequence models both in training set and (g) validation set. LASSO, least absolute shrinkage and selection operator. ROC, receiver operating characteristic. AUC, area under the curve

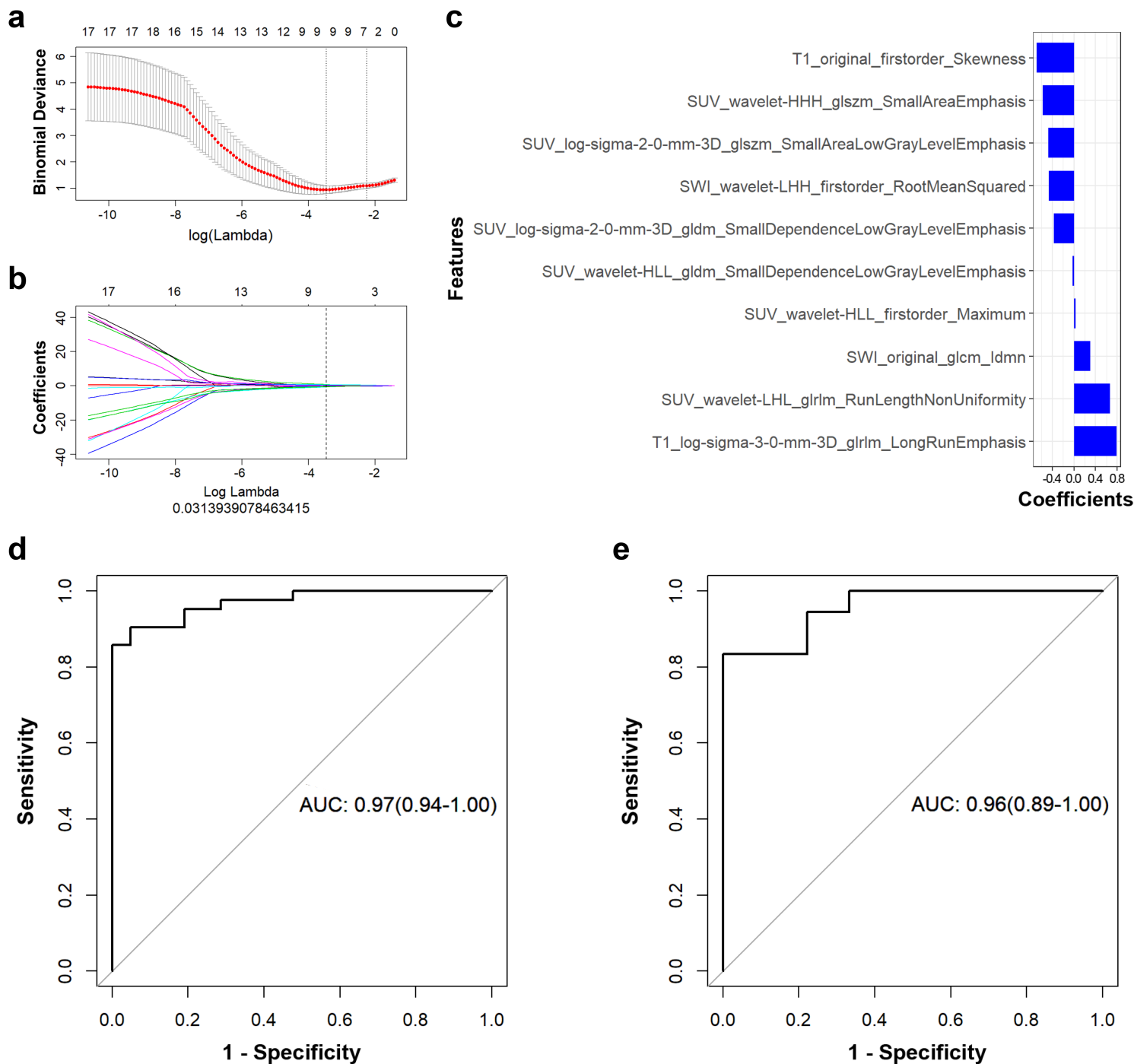


Figure 3

Diagnostic efficacy of radiomics signatures in optimal multi-modal models combining with FDG, T1WI and SWI. (a) (b) LASSO. (c) Coefficients of features in RadscoreFDG_T1WI_SWI. (d) The ROC curves of modelFDG_T1WI_SWI in training set (AUC = 0.971) and (e) validation set (AUC = 0.957). ROC, receiver operating characteristic. AUC, area under the curve. LASSO, least absolute shrinkage and selection operator

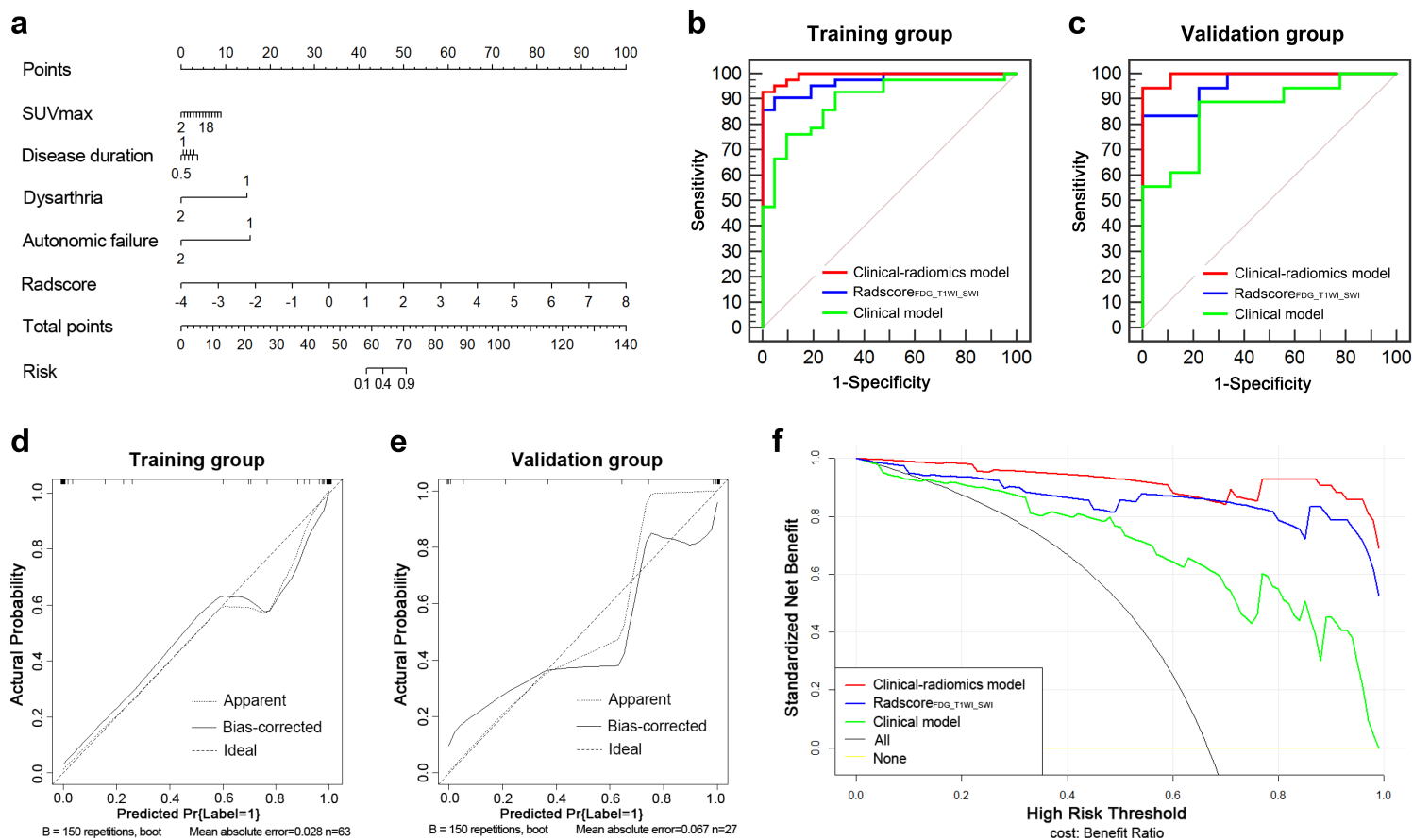


Figure 4

Diagnostic evaluation and validation of clinical-radiomics integrated models for differentiation of PD and MSA. (a) Nomogram based on clinical characteristics and Radscore_{FDG_T1WI_SWI}. (b) The ROC curves of Radscore_{FDG_T1WI_SWI}, clinical model and clinical-radiomics integrated model in the training set and (c) validation set; and clinical-radiomics integrated model performed best (AUC of the training set and validation set, 0.993 vs 0.994). (d) Calibration curves for clinical-radiomics integrated model in training sets ($P = 0.9862$) and (e) validation sets ($P = 0.8095$). (f) Decision curves for Radscore_{FDG_T1WI_SWI}, clinical model and clinical-radiomics integrated model; the y-axis shows the model benefit. The red line represents nomogram. The blue line represents Radscore_{FDG_T1WI_SWI}, the green line represents the clinical model. The x-axis means the threshold probability. ROC, receiver operating characteristic. AUC, area under the curve

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [ESM2.docx](#)