

# $^{18}\text{F}$ -FDG-PET Radiomics Based on White Matter Predicts The Progression of Mild Cognitive Impairment to Alzheimer Disease: A Machine Learning Study

Jiaxuan Peng, MMed, Wei Wang, MMed, Qiaowei Song, MMed, Jie Hou, MMed, Hui Jin, MMed, Xue Qin, MMed, Zhongyu Yuan, MMed, Yuguo Wei, MD, Zhenyu Shu, MD

**Rationale and Objectives:** To build a model using white-matter radiomics features on positron-emission tomography (PET) and machine learning methods to predict progression from mild cognitive impairment (MCI) to Alzheimer disease (AD).

**Materials and Methods:** We analyzed the data of 341 MCI patients from the Alzheimer's Disease Neuroimaging Initiative, of whom 102 progressed to AD during an 8-year follow-up. The patients were divided into the training (238 patients) and test groups (103 patients). PET-based radiomics features were extracted from the white matter in the training group, and dimensionally reduced to construct a psychoradiomics signature (PS), which was combined with multimodal data using machine learning methods to construct an integrated model. Model performance was evaluated using receiver operating characteristic curves in the test group.

**Results:** Clinical Dementia Rating (CDR) scores, Alzheimer's Disease Assessment Scale (ADAS) scores, and PS independently predicted MCI progression to AD on multivariate logistic regression. The areas under the curve (AUCs) of the CDR, ADAS and PS in the training and test groups were 0.683, 0.755, 0.747 and 0.737, 0.743, 0.719 respectively, and were combined using a support vector machine to construct an integrated model. The AUC of the integrated model in the training and test groups was 0.868 and 0.865, respectively (sensitivity, 0.873 and 0.839, respectively; specificity, 0.784 and 0.806, respectively). The AUCs of the integrated model significantly differed from those of other predictors in both groups ( $p < 0.05$ , Delong test).

**Conclusion:** Our psych radiomics signature based on white-matter PET data predicted MCI progression to AD. The integrated model built using multimodal data and machine learning identified MCI patients at a high risk of progression to AD.

**Key Words:** Radiomics;  $^{18}\text{F}$ -FDG-PET; Mild cognitive impairment; Alzheimer's disease; White matter.

© 2022 The Association of University Radiologists. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Acad Radiol 2022; ■:1–11

From the Jinzhou medical university, Jinzhou, Liaoning Province, China (J.P., J.H., Z.Y.); Department of Radiology, The First Affiliated Hospital of Chongqing Medical and Pharmaceutical College, Chongqing, China (W.W.); Center for Rehabilitation Medicine, Department of Radiology, Zhejiang Provincial People's Hospital, Affiliated People's Hospital, Hangzhou Medical College, Hangzhou, Zhejiang, China (Q.S., Z.S.); Bengbu medical college, Bengbu, China (H.J., X.Q.); Department of Pharmaceuticals Diagnosis, GE Healthcare, Hangzhou, China (Y.W.). Received October 27, 2022; revised December 11, 2022; accepted December 18, 2022. Jiaxuan Peng and Wei Wang contributed equally to this work. **Address correspondence to:** Z.S. e-mail: [cooljuty@hotmail.com](mailto:cooljuty@hotmail.com)

© 2022 The Association of University Radiologists. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>) <https://doi.org/10.1016/j.acra.2022.12.033>

## INTRODUCTION

Alzheimer's disease (AD) is a highly disabling, progressive, neurodegenerative disease, for which there is currently no effective treatment (1). Mild cognitive impairment (MCI), an intermediate stage between normal age-related cognitive changes and dementia, is considered to be a precursor to AD. Approximately 10%–12% of patients with MCI progress to AD every year, and early interventions for MCI may delay this progression or even prevent the occurrence of AD entirely (2). Therefore, reliable biomarkers are urgently required for identifying MCI patients who are at a high risk for progression to AD and implementing individualized treatments that prevent or slow the development of AD.

Positron-emission computed tomography (PET) is a functional imaging technique that is widely used to evaluate the functions of affected brain regions for the clinical assessment and diagnosis of neurodegenerative diseases (3).  $^{18}\text{F}$ -Fluoro-2-deoxyglucose ( $^{18}\text{F}$ -FDG)-PET is an excellent predictor of cognitive impairment, as it can measure the local cerebral glucose metabolism and reveal typical metabolic patterns in suspected AD patients (4).

The whole cerebellum and the cerebellar gray matter are commonly used as reference regions for analyzing glucose metabolism on  $^{18}\text{F}$ -FDG-PET; however, considerable variability has been observed in these regions on longitudinal PET measurements, raising concerns about the suitability of these structures as reference regions (5). Animal experiments have shown that white-matter destruction and myelin degeneration are the earliest pathological changes in AD (6), and in AD patients, abnormal amyloid- $\beta$  metabolism in the cerebrospinal fluid is closely related to pathological changes in the white matter (7). Therefore, the subcortical white matter has now become the preferred reference region for monitoring longitudinal changes in AD (8). This method of monitoring neurodegenerative changes may have inherent robustness for revealing longitudinal changes in cerebral perfusion and disease progression, as the results produced are more stable and may more accurately reflect the degree of disease progression (9). However, this specific pattern of local glucose metabolism changes is not limited to AD but also occurs in other degenerative dementias (10). Therefore, the pattern of brain metabolic changes alone cannot accurately identify MCI patients who are at a high risk for progression to AD.

Radiomics is an emerging technology that combines medicine and engineering, and is capable of quantifying the inherent heterogeneity or homogeneity of tissues by extracting a large number of radiomics features, thereby aiding prognostic predictions and clinical decision-making. The early applications of radiomics were mainly limited to the field of oncology, but the technique is currently also used in the field of neurology due to its great universality (11). We previously reported that radiomics based on magnetic resonance imaging (MRI) of the whole-brain white matter can help predict the progression of Parkinson disease (12). Another study found that the application of an artificial intelligence-based algorithm to multimodal data can greatly improve the prediction of MCI progression to AD (13). Accordingly, we hypothesized that PET-based radiomics features of the white matter may predict an individual's risk of progression from MCI to AD, and that machine learning algorithms may improve the efficiency of this prediction.

The present study aimed to extract radiomics features associated with progression from MCI to AD from PET images of the white matter region and to use machine learning methods to combine multimodal data, such as clinical information and semiquantitative PET data, to construct an integrated model for predicting which patients are at a high risk of progressing from MCI to AD.

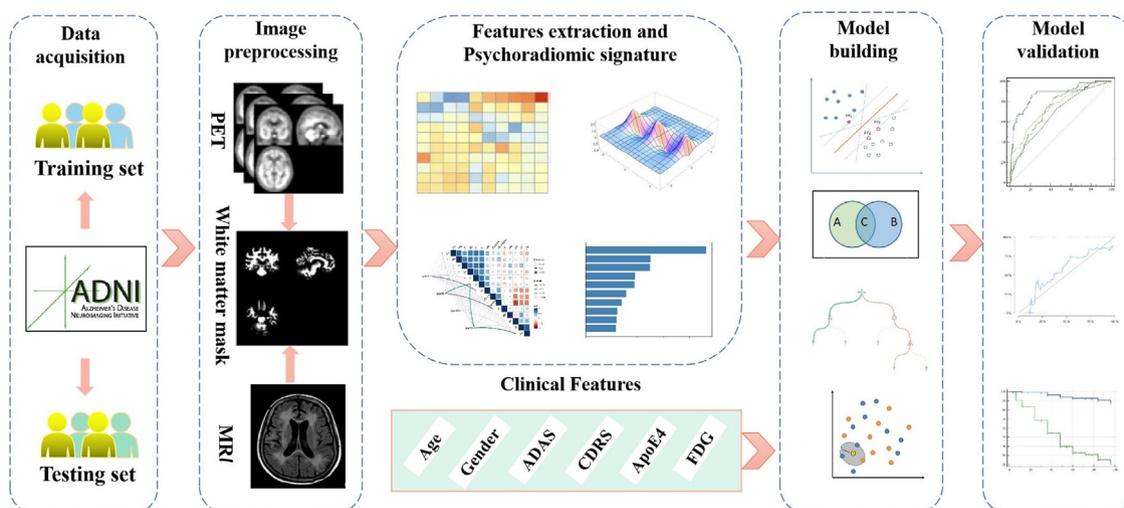
## MATERIALS AND METHODS

### Research Subjects

The case data included in this study were obtained from the ANDI-2 and ANDI-GO datasets on the official website of Alzheimer's Disease Neuroimaging Initiative (<https://adni.loni.usc.edu/>). For information about the ethics review related to the ADNI data, please refer to the website. The ADNI is a longitudinal multicenter study designed to develop clinical, imaging, biochemical, and genetic biomarkers for the early detection and monitoring of AD. The ADNI was launched as a public-private partnership in 2004 by the National Institute on Aging, the Foundation for the National Institutes of Health, and several private companies and nonprofit organizations. Given that the main goal of ADNI is to determine whether imaging technology (PET, MRI), clinical and neuropsychological assessments, and other biomarkers can be combined to assess the progression of MCI and early-stage AD, we also collected the above data for this study. Specifically, we analyzed the subjects' general demographic information, APOE4 gene status (positive *vs.* negative), Mini-Mental State Examination (MMSE) score, Alzheimer's Disease Assessment Scale (ADAS) score, and Clinical Dementia Rating (CDR). The inclusion criteria were as follows: (1) initial diagnosis of MCI and follow-up evaluations, and (2) baseline MRI, PET, and clinical data. The Exclusion criteria were as follows: (1) lack of relevant research data, and (2) poor MRI image quality and unable to register with PET image, and (3) the patient will have bidirectional transformation (MCI progresses to AD and then to MCI). In total, we enrolled 341 patients with MCI, and 102 of these patients progressed to AD during a follow-up period of 8 years, and the exact numbers of patient and the specific ADNI database were found in the Supplementary materials (Table S2). We divided these patients according to their entry numbers in the ADNI database as follows: training group, 238 patients and test group, 103 patients. We used data from the training group to construct a predictive model of AD, and verified the performance of this model in the test group. The research protocol is depicted in Figure 1. For details of the MRI and  $^{18}\text{F}$ -FDG-PET data used in this study, please visit the image protocol section of the ADNI dataset on their official website.

### FDG-PET Image Acquisition and Preprocessing

The PET images in the ADNI database (<https://adni.loni.usc.edu/>) were acquired using General Electric, Philips, and Siemens PET scanners. Each subject was injected with  $185 \pm 18.5$  MBq FDG, and 30 min later, a dynamic three-dimensional scan consisting of six frames of 5 min each was obtained. All frames were merged to obtain a single image after motion-correction of all subsequent frames to the first frame. To construct 3D white-matter regions from the PET images, we used structural magnetic resonance T1-weighted images for rigid registration with the PET images. Image processing was performed using



**Figure 1.** Research flow chart.

SPM12 (<http://www.fil.ion.ucl.ac.uk/spm/>) in MATLAB 2016b (<https://ww2.mathworks.cn/products/matlab.html>). First, the original baseline FDG-PET images were registered with the corresponding structural MRI scans. We used T1-weighted images as a template for rigid registration, and ensured that the PET images had identical spacing, voxels, and origin. Using the unified segmentation method, we segmented the MRI scans into gray matter, white matter, and cerebrospinal fluid tissue-probability maps. The registered PET images were normalized to the MNI space by using transformation parameters. After registration, the PET images could share regions of interest (ROIs) in the white matter segmented from the structural T1-weighted images. Pretreatment process can be found in Supplementary materials.

### Radiomics Feature Extraction and Selection

Before extracting the radiomics features, we performed the following preprocessing steps on the PET images using A.K. software (Quantitative Analysis Kit, v1.2; GE Healthcare): each sequence of images was resampled at a  $1 \times 1 \times 1 \text{ mm}^3$  resolution by linear interpolation, and the gray level of the images was discretized and normalized to 32 orders. After preprocessing, the baseline PET images and white-matter ROIs were imported into the Pyradiomics library (<https://github.com/Radiomics/pyradiomics>)(v2.1.1) (14) for Image Biomarker Standardization Initiative-compliant feature extraction in the nii format. In all, 279 features were extracted for each subject, and detailed information on the features extracted can be found in Supplementary Table S1. Subsequently, redundant and irrelevant features were excluded from the training dataset, and only a small set of radiomics features that were closely related to patient outcomes were selected in order to obtain good predictive performance. To ensure the accuracy and stability of the features extracted from the ROIs during the process of feature-dimensionality reduction, two physicians (radiologists A and

B) corrected the automatically segmented white-matter regions by removing the cerebellum, brainstem, and any non-brain tissue from the automatic segments, and modifying the white-matter segmentation as needed, and we will verify the segmentation results by importing the PET images first in ITK-SNAP (<http://www.itksnap.org/pmwiki/pmwiki.php>) and then import the images of the white matter region images. When it is found that there is a slight deviation in the white matter image on the PET image, such as white matter areas overlaying gray matter or cerebrospinal fluid areas in the PET image, two radiologists will manually adjust it at this time, and these corrections has proved to be necessary. Thus, two sets of features were obtained, one from the segments marked by radiologist A and the other from the segments marked by radiologist B. The Spearman rank correlation test was used to calculate the correlation coefficient of each feature between the two sets of features. Features with correlation coefficients  $> 0.8$  were considered as robust (15), and were further processed using the dimension-reduction methods minimum redundancy maximum relevance (mRMR) and gradient boosting decision trees (GBDT) (16). The specific dimension-reduction process is shown in the Supplementary materials.

### Integrated Model Construction

Radiomics biomarkers were selected using logistic regression analysis of the features remaining after dimensionality reduction. The biomarkers thus selected were termed the “psychoradiomics signature,” and a score was calculated using this signature for each patient. This score, termed the rad score, reflected the probability of progression of MCI to AD. Potential clinical predictors of MCI progression were selected using the backward stepwise selection method with a stopping rule based on Akaike information criterion in the training dataset. A model integrating the radiomics signature and selected clinical features was constructed using machine

learning classifiers, including support vector machine (SVM), naive Bayes, random forest, and k-nearest neighbor. To avoid reporting biased results and limit overfitting, we used 10-fold cross-validation while constructing the integrated model. Finally, optimal classifiers were selected using accuracy and kappa values. The output of the integrated model was a binary prediction of the course of MCI, defined as MCI stabilization or MCI progression.

The kappa value during machine learning was defined as  $\text{kappa} = (\text{observed accuracy} - \text{expected accuracy}) / (1 - \text{expected accuracy})$  (17).

### Validation of the Integrated Model

To quantify the performance of the model, we used receiver operating characteristic (ROC) curves to assess the performance of the integrated model and the Delong test to verify the difference in diagnostic performance between the integrated model and other clinical predictors (18). The goodness of fit of the integrated model was analyzed using the Hosmer-Lemeshow test, and the consistency between the predicted and actual progression of MCI was visualized using calibration curves. The clinical net benefit of the integrated model was determined using decision curve analysis (19). To analyze the clinical efficacy of the integrated model, we calculated a prognostic index (PI) for each subject. We took the threshold value of the Youden index of the ROC curve as the optimal cutoff value (20), and used it to divide the patients into a progressive MCI group and a stable MCI group. Then the Kaplan–Meier survival curves based on ranked PI values were used to compare the rates of MCI progression to AD between the above groups (21).

### Statistical Analysis

Statistical analysis was conducted using MedCalc software (version 11.2), Python (version 3.5), and SPSS software (version 24.0). The *t*-test or the Mann–Whitney *U* test (22) was used to compare measurement data, while the chi-square test was used to compare rank data. ROC curve-based metrics (23), such as accuracy, sensitivity, and specificity, were used to assess model performance. All statistics were two-way, and differences with *p* values of  $<0.05$  were considered to be statistically significant.

## RESULTS

### Comparison of Clinical Factors

The clinical data did not significantly differ between the training and test groups ( $p > 0.05$ ). In both the training and test groups, APOE4 gene status; MMSE, CDR, and ADAS scores; and FDG values significantly differed between the stable and progressive MCI groups ( $p < 0.05$ ), whereas the other clinical data, include the age, gender, education, did

not show significant differences between the stable and progressive MCI groups ( $p > 0.05$ ; Table 1).

### Establishment and Validation of Psychoradiomics Model

A total of 279 features (first-order, grey-level cooccurrence matrix, grey-level run length matrix, grey-level size zone matrix, grey-level dependencematrix, neighboring grey tone difference matrix, log sigma matrix) were extracted in PET sequence images. According to the condition that the correlation coefficient is greater than 0.8, we retained a total of 237 features. Then, mRMR and GBDT algorithm was used to reduce the dimension of the remaining features. Thereafter, 10 features were screened to construct the psychoradiomics model, including three original features and 7 log features (Fig 2). The specific dimensionality-reduction process can be found in the Supplementary materials. The areas under the curve (AUCs) of the psychoradiomics model in the training and test groups were 0.747 and 0.719, respectively; the corresponding sensitivity was 0.831 and 0.613, and the corresponding specificity was 0.581 and 0.778 (Fig 3).

### Integrated Model Construction and Validation

The CDR score, ADAS score, and psychoradiomics signature were identified as independent predictors of MCI progression to AD on multivariate logistic regression analysis, and used to build an integrated model (Table 2). The independent predictive factors were processed using four machine learning methods, and the results showed that the model built using the SVM and KNN method had the better performance (Fig 4). When the Linear of SVM is 0.1, the model has the best accuracy and Kappa value. Details of the machine learning screening results are provided in the Supplementary materials. The machine learning results show that the classification results of the four models of stable MCI and progression MCI. The integrated model has the best classification performance, and the AUC values of the training group and the test group are 86.8% and 86.5%, respectively. The remaining three models performed obviously lower than the integrated model with accuracy of 71.9%, 74.2% and 73.7%, and the sensitivity of 61.3%, 64.5%, 80.7%, respectively; and specificity of 77.8%, 75% and 56.9%, respectively. To further showcase the superiority of the integrated model. Calibration curves showed that in both the training and test groups, the integrated model had no overfitting as compared to the actual predicted value, and the integrated model had the best diagnostic performance among all the predictors assessed. The accuracy of the Integrated model, psychoradiomics signature, CDR score and ADAS score in the training and test groups were 0.868 and 0.819, 0.755 and 0.741, 0.626 and 0.669, 0.596 and 0.551, respectively. In addition, the AUC of the integrated model significantly differed from those of other predictors in the training and test groups ( $p < 0.05$ , Delong test). Decision curve analysis showed excellent net benefit of the integrated model (Table 3, Fig 5).

TABLE 1. Comparative Analysis of Clinical Data Between the Training and Test Groups

Characteristic	Training Cohort (n = 238)			Test Cohort (n = 103)			Training vs. Test cohort p value
	All Patients	Stable MCI (n = 167)	Progressive MCI (n = 71)	All Patients	Stable MCI (n = 72)	Progressive MCI (n = 31)	
Gender (n, %)							
Male	127 (53.36%)	88 (52.69%)	39 (54.93%)	51 (49.51%)	39 (54.17%)	12 (38.71%)	0.514
Female	111 (46.64%)	79 (47.31%)	32 (45.07%)	52 (50.49%)	33 (45.83%)	19 (61.29%)	
APOE4 (n, %)							
Negative	115 (48.32%)	89 (53.29%)	26 (36.62%)	47 (45.63%)	38 (52.78%)	9 (29.03%)	0.026*
Positive	123 (51.68%)	78 (46.71%)	45 (63.38%)	56 (54.37%)	34 (47.22%)	22 (70.97%)	
Age (years)	72.29 ± 7.64	72.15 ± 7.92	72.63 ± 7.00	71.75 ± 7.41	70.95 ± 7.60	73.62 ± 6.70	0.093
MMSE score	28.02 ± 1.67	28.20 ± 1.67	27.61 ± 1.63	28.02 ± 1.81	28.36 ± 1.61	27.23 ± 2.01	0.008*
CDR	1.56 ± 0.94	1.37 ± 0.85	1.99 ± 1.01	1.48 ± 0.82	1.26 ± 0.69	1.98 ± 0.89	<0.001*
ADAS	15.33 ± 6.40	13.54 ± 5.52	19.59 ± 6.38	15.50 ± 6.87	13.42 ± 5.78	20.31 ± 6.85	<0.001*
Education (years)	16.19 ± 2.61	16.17 ± 2.72	16.25 ± 2.35	16.22 ± 2.85	16.25 ± 2.88	16.16 ± 2.82	0.886
FDG (SUV)	1.25 ± 0.14	1.28 ± 0.13	1.18 ± 0.14	1.24 ± 0.13	1.29 ± 0.11	1.14 ± 0.12	<0.001*

MCI, mild cognitive impairment; APOE4, apolipoprotein E gene; MMSE, mini mental state examination; CDR, clinical dementia rating; ADAS, Alzheimer's disease assessment scale; FDG, fluorodeoxyglucose; SUV, standardized uptake value.  
\* indicates significant p value.

Clinical Evaluation of the Integrated Model

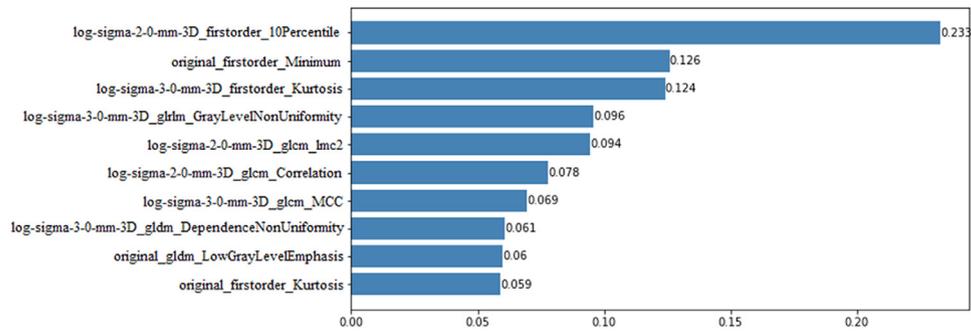
In the training group, ROC curve analysis showed that the optimal cutoff value corresponding to the threshold value of the Youden index was 0.2329. Using this cutoff, we divided the subjects in the training and test groups into those with a low risk and those with a high risk of MCI progression to AD. The timing of MCI progression significantly differed between the low- and high-risk groups ( $\chi^2 = 94.12$  and  $35.05$ , respectively;  $p < 0.0001$ ) according to the log-rank test and Kaplan–Meier survival curves analysis. The hazard ratios (95% confidence interval [CIs]) in the training and test groups were 12.42 (7.4666–20.6655) and 10.11 (4.7005–21.7502), respectively, indicating that the risk of MCI progression in the high-risk group was 12.42 times and 10.11 times that in the low-risk group (Fig 6).

DISCUSSION

This study showed that a psychoradiomics signature constructed using PET imaging of white-matter regions significantly differed between patients with stable MCI and those with progressive MCI. This indicates that white-matter regions contain relevant information that may predict the progression of MCI to AD, which may provide a reliable method for the clinical analysis of disease progression. We also found that the rad-score was superior to clinical scale scores (Fig 5C, 5D), which further confirmed that the PET-based white-matter radiomics signature could serve as a biomarker for identifying MCI progression. The combination of the radiomics signature and clinical scale scores by using machine learning methods further improved the prediction efficiency of the model, which reflects the performance advantages of machine learning.

Thus far, most of the radiomics-based research on AD has focused on the hippocampus (24–26) and amygdala (27). Feng et al. concluded that hippocampal texture is a potential biomarker for AD, and extracted radiomics features from the hippocampal region in patients with AD, amnesic MCI, and normal cognition (28). However, few studies have analyzed MCI progression based on PET data of white-matter regions. Shao et al. reported that white-matter damage occurs early in the course of AD, and is highly likely to be involved in the pathogenesis of AD (29). Currently, radiomics-based research on AD, specific white matter fiber tract injury is also a potential biomarker for AD (30). Cognitive function is affected by changes in the white matter in patients with MCI (31). Hence, it is necessary to comprehensively analyze the whole brain and especially, the white matter. Our results confirmed that PET-based radiomics features of the white matter can accurately identify MCI patients who are at a high risk for progression to AD. This finding may be attributable to the high amount of information contained in PET images.

Amyloid- $\beta$  and tau protein deposition in the brain is known to accelerate MCI progression, and both biomarkers effectively predict cognitive decline. Although PET can be

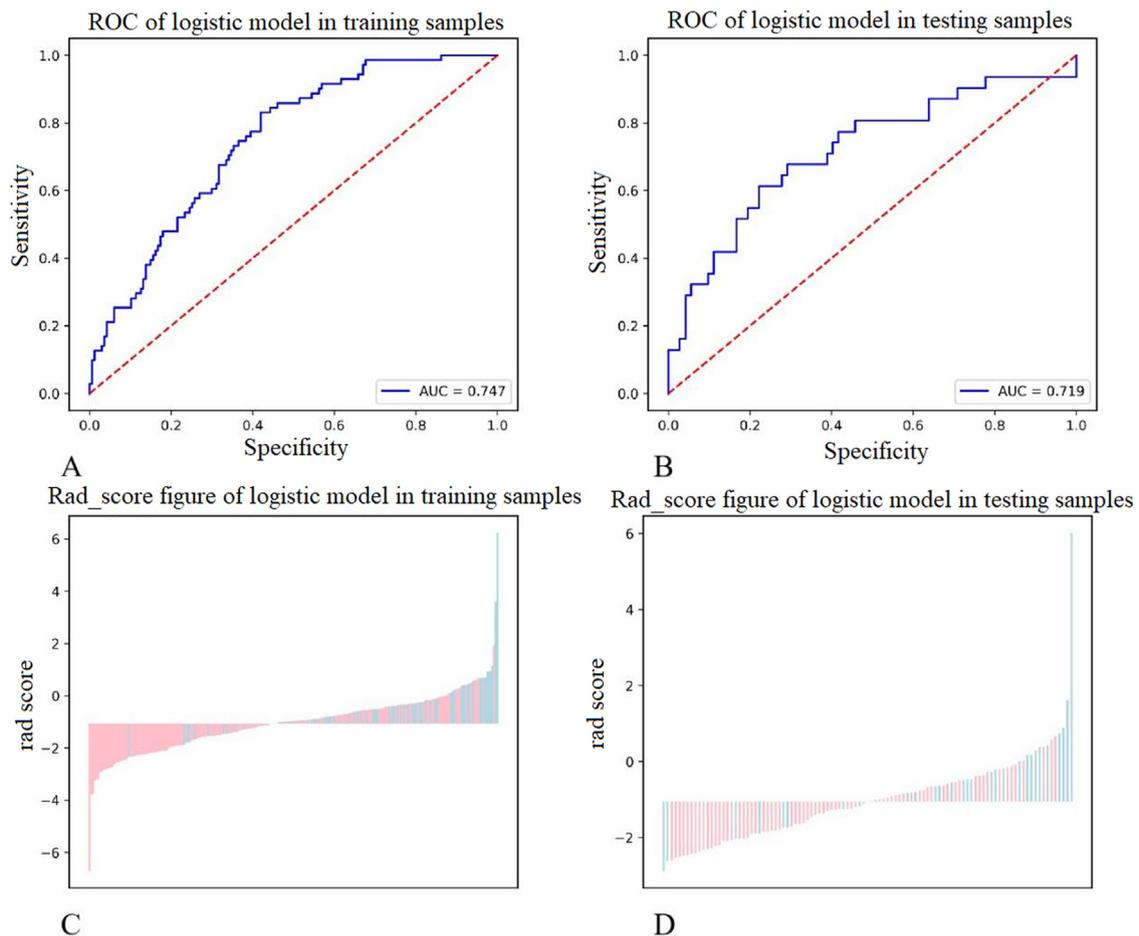


**Figure 2.** The 10 radiomics features remaining after dimensionality reduction. The abscissa represents the weighted value of each feature.

used to quantify both these indicators, the high cost-effectiveness of amyloid- and tau-based PET has limited its clinical utility. In contrast,  $^{18}\text{F}$ -FDG-PET is easier to perform and shows better predictive performance. Shaffer et al. reported that among MRI, FDG-PET, and cerebrospinal fluid analysis, only FDG-PET significantly improved the prediction of MCI progression to AD (32), which further confirms the clinical application value of FDG-PET.

Combining radiomics and clinical features provides more reliable information than either feature type on its own for

disease classification and prediction models. Moreover, neurological diseases such as AD have specific clinical scales that can detect abnormal cognitive function in the early stage when CT, MRI, and other imaging studies still do not show lesions. In this study, we identified the CDR and ADAS scores as the independent predictors of MCI progression and used them for the construction of an integrated model. Ranjbar et al. grouped subjects according to their CDR scores, and found that the CDR showed good performance in distinguishing subjects with no cognitive impairment (CDR



**Figure 3.** (A, B) Diagnostic performance of the psychoradiomics signature in the training and test groups. (C, D) Classification performance of the psychoradiomics signature in the training and testing groups. Values less than 0 indicate stable cases of MCI, and values greater than 0 indicate progressive cases of MCI. Blue indicates progressive cases, and red indicates stable cases.

TABLE 2. Independent Predictors of MCI Progression on Multivariate Logistic Regression Analysis

Variable	Univariate Logistic Regression		Multivariate Logistic Regression	
	OR (95% CI)	p value	OR (95% CI)	p value
Gender	1.208 (0.582, 2.508)	0.611	NA	NA
APOE4	1.429 (0.706, 2.891)	0.321	NA	NA
Age (years)	0.984 (0.937, 1.035)	0.536	NA	NA
MMSE score	1.076 (0.863, 1.341)	0.517	NA	NA
CDR score	1.521 (1.041, 2.223)	0.03	1.605 (1.122, 2.296)	0.01*
ADAS score	1.134 (1.059, 1.215)	<0.001*	1.132 (1.067, 1.201)	<0.001*
Education (years)	1.012 (0.885, 1.156)	0.864	NA	NA
FDG	0.391 (0.019, 7.969)	0.542	NA	NA
Psychoradiomics signature	2.193 (1.417, 3.395)	<0.001*	2.277 (1.52, 3.409)	<0.001*

MCI, mild cognitive impairment; OR, odds ratio; CI, confidence interval; APOE4, apolipoprotein e gene; MMSE, mini mental state examination; CDR, clinical dementia rating; ADAS, alzheimer’s disease assessment scale; FDG, fluorodeoxyglucose.

\* indicates significant p values.

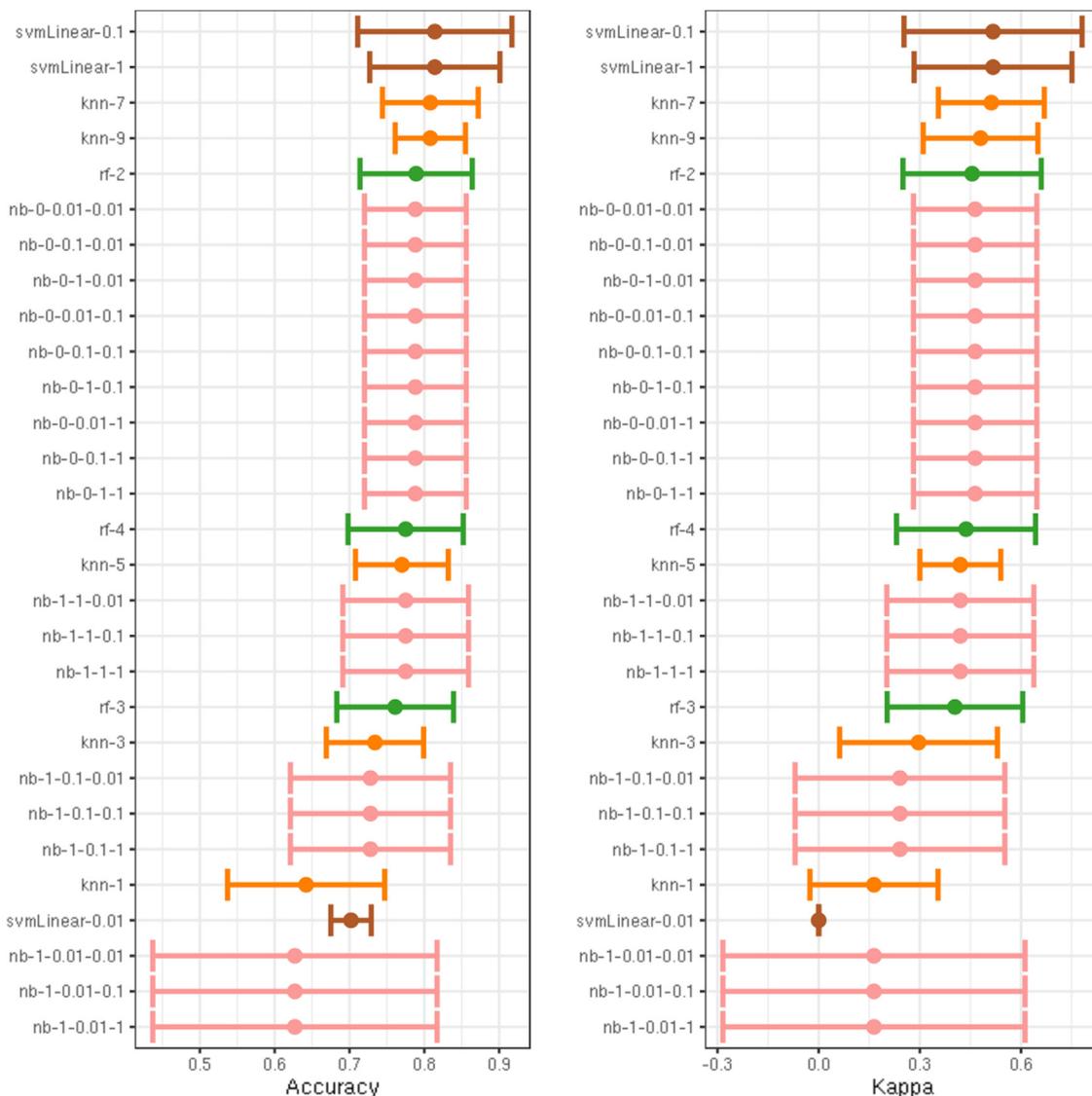
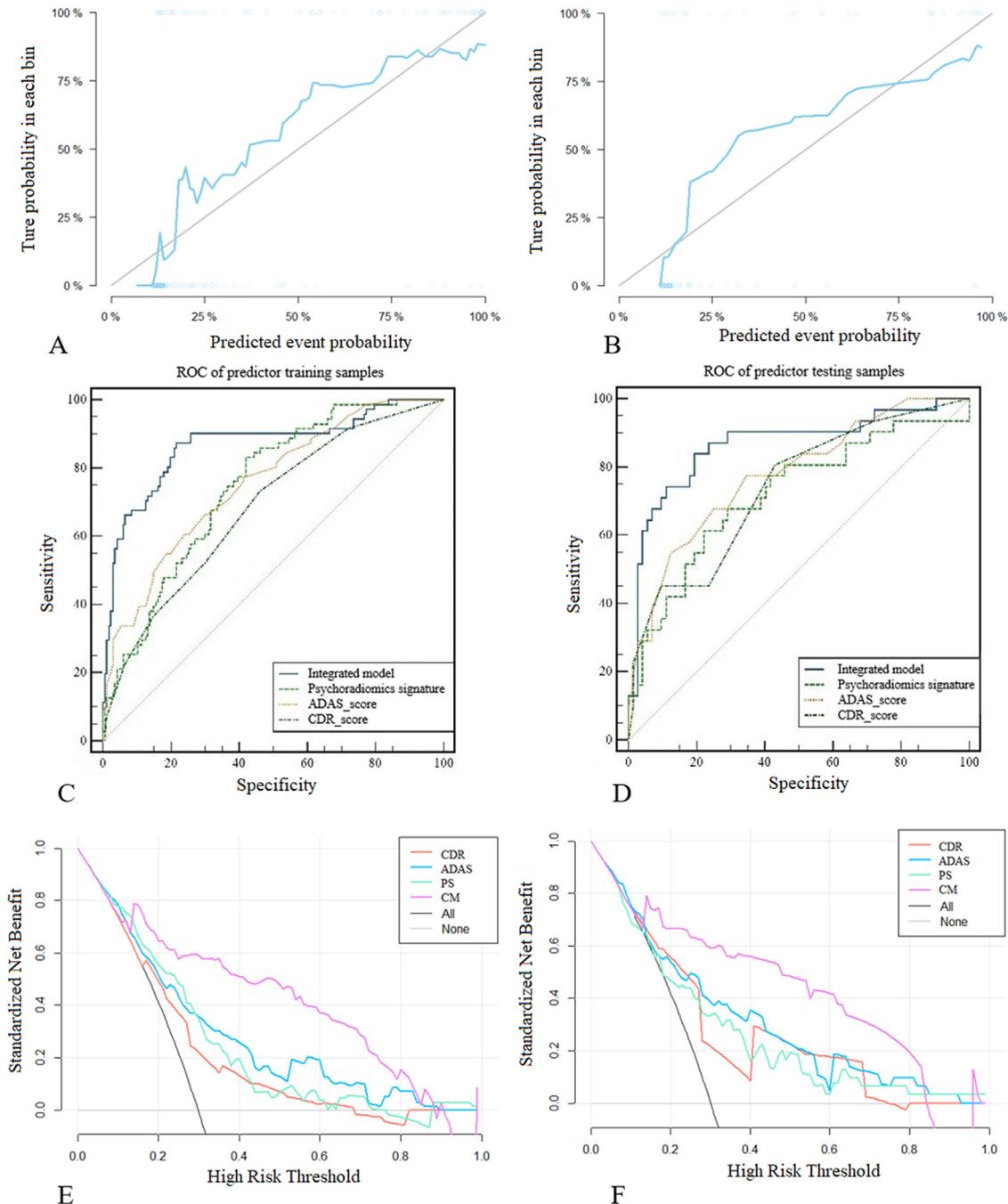


Figure 4. Diagnostic performance of models built using different machine learning methods combined with different hyperparameters. The X-axes represent the accuracy and kappa value, and the Y-axes represent different machine learning methods and hyperparameters.

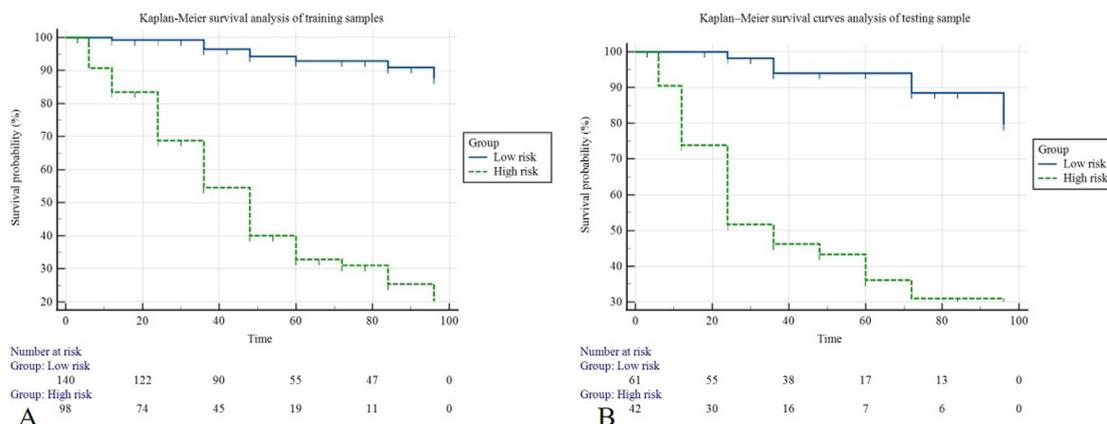
**TABLE 3. Diagnostic Performance of Each Model in the Training and Test Groups**

Predictive Model/Factor	Training Cohort			Test Cohort		
	AUC	Sensitivity	Specificity	AUC	Sensitivity	Specificity
Integrated model	0.868	0.873	0.784	0.865	0.839	0.806
Psychoradiomics signature	0.747	0.831	0.581	0.719	0.613	0.778
ADAS score	0.755	0.606	0.761	0.742	0.645	0.75
CDR score	0.683	0.732	0.539	0.737	0.807	0.569

AUC, area under the curve; ADAS, alzheimer's disease assessment scale; CDR, clinical dementia rating.



**Figure 5.** (A, B) Validation curves of the integrated model in the training (A) and test (B) groups. (C, D) Diagnostic performance of the integrated model and individual predictors in the training (C) and test (D) groups. (E, F) Net clinical benefit of the integrated model and individual predictors in the training (E) and test (F) groups.



**Figure 6.** Survival curve analysis of low-risk and high-risk patients in the (A) training and (B) test groups, based on classification according to the integrated model.

score, 0) from those with mild dementia (CDR score, 1; AUC, 0.95), but poor performance in distinguishing between subjects with CDR scores of 1 and 2 (AUC, 0.56;  $p = 0.46$ ) (33). In our study, the AUC of the CDR for predicting MCI progression was 0.683. A reasonable explanation for the difference between our finding and that of Ranjbar et al. is that their sample size was 173, and approximately two-thirds of their subjects had CDR scores of  $<1$ , which is very different from the distribution of CDR scores in our sample. Furthermore, Aisen et al. argued that clinical scales rely on making a distinction between stages, whereas AD does not present in discrete, well-defined stages, but rather as a multifaceted continuous process (34); hence, clinical scales can only play a supportive role in assessing AD progression, and cannot serve as a decisive factor.

The good performance of our radiomics-based model is attributable to not only the inclusion of clinical scales but also the application of machine learning methods during model building. Grueso et al. analyzed the data of 116 AD by using a machine learning algorithm, which only 24 are related to the prediction of MCI progression (35). To provide clinicians with more substantial assistance in treatment planning, we limited our predictions to whether MCI patients would remain stable or progress to AD, rather than attempting to classify the subjects into the normal cognition, MCI, and AD groups. Plant C et al. found that during an average follow-up of 2.5 years, nine of 24 MCI patients progressed to AD, and reported that their voting feature intervals-based classifier achieved approximately 75% prediction accuracy in both the training and test sets (36). It should be noted that 2 or 3 years of follow-up may not be sufficient to detect progression to AD and dementia. In addition, the authors adopted a new feature-selection algorithm. First, they selected brain regions with the highest classification accuracy (92%) between AD and NC and then, extracted the features that predicted MCI progression from these regions. Finally, they used the Bayesian and voting feature intervals algorithms to improve their prediction model. These differences probably led to differences in the final optimal classifiers between our study and the

above study. The above feature-selection method is also instructive; it achieves relatively considerable accuracy with a small sample size. Using structural MRI data from the ADNI, Misra et al. established a model with an accuracy of 0.75–0.80 and an AUC of 0.77 in the progressive MCI group ( $n = 27$ ) and stable MCI group ( $n = 76$ ) (37). In comparison, our SVM-based model achieved an AUC of 0.865. This may be related to two reasons. First, we chose <sup>18</sup>F-FDG-PET for image acquisition. Second, each brain region has its unique histological characteristics or underlying mechanisms. We selected the white-matter region, which is large and contains a considerable amount of information. Our results also show that compared to other classifiers, SVM is associated with higher kappa and accuracy values, which is in agreement with the findings of Li et al. (38) and Zhang and Shen (39). These results suggest that SVM may be more suitable for predicting the progression of neurodegenerative diseases. But we also need to note that the best kappa of SVM was still middle moderate, and not yet really in the highly reproducible range ( $\text{Kappa} \geq 0.75$ ) (17). In fact, general retrospective analysis may lead to certain data imbalance, and the model will be somewhat biased leading to low kappa values. In this study, there were 239 negative cases (stable group) and 102 positive cases (progressive group), respectively, which may be due to the imbalance of data distribution resulted in the moderate efficacy of kappa value.

Our study has some limitations. First, PET images use steps such as smoothing during preprocessing, and some texture information is inevitably lost in these processes. Whether this affects the extracted radiomics features remains to be determined. Second, this study is a retrospective analysis of publicly available datasets, and requires validation using external datasets. Won et al. summarized 26 radiomics analyses of MCI patients in a radiomics quality score system, and found that external validation was performed in only one study (3.8%) (41). This reflects the limited access to AD data, and we are making efforts to improve feature selection, clinical applicability, and model performance, which is still needs to be prospective validation from another datasets from set. At

last, the study population was divided into training and test groups according to the entry numbers of ANDI database, and the entry numbers were arranged according to the time when patients were recruited into the ANDI database. However, that absolutely can introduce a time-based bias as equipment and imaging processing methods have changed with time.

In conclusion, the radiomics signature that we constructed based on white-matter regions in PET images can be used as a biomarker for identifying people with MCI who may progress to AD. Our psychoradiomics model combined with multimodal data could identify high-risk MCI patients, which not only improves our understanding of the spectrum of AD but may also help identify biomarkers for the early detection of AD and for the monitoring of disease progression.

## REFERENCES

- Satizabal CL, Beiser AS, Chouraki V, et al. Incidence of dementia over three decades in the framingham heart study. *N Engl J Med* 2016; 374(6):523–532. doi:10.1056/NEJMoa1504327.
- Langa KM, Levine DA. The diagnosis and management of mild cognitive impairment: a clinical review. *JAMA* 2014; 312(23):2551–2561. doi:10.1001/jama.2014.13806.
- Kreisl WC, Kim MJ, Coughlin JM, et al. PET imaging of neuroinflammation in neurological disorders. *Lancet Neurol* 2020; 19(11):940–950. doi:10.1016/S1474-4422(20)30346-X.
- Khosravi M, Peter J, Winterson NA, et al. 18F-FDG is a superior indicator of cognitive performance compared to 18F-florbetapir in alzheimer's disease and mild cognitive impairment evaluation: a global quantitative analysis. *J Alzheimers Dis* 2019; 70(4):1197–1207. doi:10.3233/JAD-190220.
- van Berckel BN, Ossenkoppele R, Tolboom N, et al. Longitudinal amyloid imaging using 11C-PIB: methodologic considerations. *J Nucl Med* 2013; 54(9):1570–1576. doi:10.2967/jnumed.112.113654.
- Desai MK, Sudol KL, Janelsins MC, et al. Triple-transgenic Alzheimer's disease mice exhibit region-specific abnormalities in brain myelination patterns prior to appearance of amyloid and tau pathology. *Glia* 2009; 57(1):54–65. doi:10.1002/glia.20734.
- Kalheim LF, Selnes P, Bjørnerud A, et al. Amyloid dysmetabolism relates to reduced glucose uptake in white matter hyperintensities. *Front Neurol* 2016; 7:209. doi:10.3389/fneur.2016.00209.
- Landau SM, Fero A, Baker SL, et al. Measurement of longitudinal  $\beta$ -amyloid change with 18F-florbetapir PET and standardized uptake value ratios. *J Nucl Med* 2015; 56(4):567–574. doi:10.2967/jnumed.114.148981.
- Blautzik J, Brendel M, Sauerbeck J. Alzheimer's disease neuroimaging initiative. Reference region selection and the association between the rate of amyloid accumulation over time and the baseline amyloid burden. *Eur J Nucl Med Mol Imaging* 2017; 44(8):1364–1374. doi:10.1007/s00259-017-3666-8.
- Bohnen NI, Djang DS, Herholz K, et al. Effectiveness and safety of 18F-FDG PET in the evaluation of dementia: a review of the recent literature. *J Nucl Med* 2012; 53(1):59–71. doi:10.2967/jnumed.111.096578.
- Li F, Wu D, Lui S, et al. Clinical strategies and technical challenges in psychoradiology. *Neuroimaging Clin N Am* 2020; 30(1):1–13. doi:10.1016/j.nic.2019.09.001.
- Shu ZY, Cui SJ, Wu X, Xu Y, Huang P, Pang PP, Zhang M. Predicting the progression of Parkinson's disease using conventional MRI and machine learning: An application of radiomic biomarkers in whole-brain white matter. *Magn Reson Med* 2021; 85(3):1611–1624. doi:10.1002/mrm.28522.
- Lee G, Nho K, Kang B. for Alzheimer's disease neuroimaging initiative. predicting alzheimer's disease progression using multi-modal deep learning approach. *Sci Rep* 2019; 9(1):1952. doi:10.1038/s41598-018-37769-z.
- van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res* 2017; 77(21):e104–e107. doi:10.1158/0008-5472.CAN-17-0339.
- Wu J, Aguilera T, Shultz D, et al. Early-stage non-small cell lung cancer: quantitative imaging characteristics of (18F) Fluorodeoxyglucose PET/CT allow prediction of distant metastasis. *Radiology* 2016; 281(1):270–278. doi:10.1148/radiol.2016151829.
- Zhang B, He X, Ouyang F, et al. Radiomic machine-learning classifiers for prognostic biomarkers of advanced nasopharyngeal carcinoma. *Cancer Lett* 2017; 403:21–27. doi:10.1016/j.canlet.2017.06.004.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33(1):159–174.
- Demler OV, Pencina MJ, Sr D'Agostino RB. Misuse of DeLong test to compare AUCs for nested models. *Stat Med* 2012; 31(23):2577–2587. doi:10.1002/sim.5328.
- Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ* 2016; 352:i6. doi:10.1136/bmj.i6.
- Fluss R, Faraggi D, Reiser B. Estimation of the Youden Index and its associated cutoff point. *Biom J* 2005; 47(4):458–472. doi:10.1002/bimj.200410135.
- Goel MK, Khanna P, Kishore J. Understanding survival analysis: Kaplan-Meier estimate. *Int J Ayurveda Res* 2010; 1(4):274–278. doi:10.4103/0974-7788.76794.
- McGee M. Case for omitting tied observations in the two-sample t-test and the Wilcoxon-Mann-Whitney Test. *PLoS One* 2018; 13(7):e0200837. doi:10.1371/journal.pone.0200837.
- Hajian-Tilaki K. Receiver operating characteristic (roc) curve analysis for medical diagnostic test evaluation. *Caspian J Intern Med* 2013; 4(2):627–635.
- Wang L, Feng Q, Ge X, et al. Textural features reflecting local activity of the hippocampus improve the diagnosis of Alzheimer's disease and amnesic mild cognitive impairment: a radiomics study based on functional magnetic resonance imaging. *Front Neurosci* 2022; 16:970245. doi:10.3389/fnins.2022.970245.
- Feng Q, Wang M, Song Q, et al. Correlation between hippocampus MRI radiomic features and resting-state intrahippocampal functional connectivity in alzheimer's disease. *Front Neurosci* 2019; 13:435. doi:10.3389/fnins.2019.00435.
- Li Y, Jiang J, Shen T, et al. Radiomics features as predictors to distinguish fast and slow progression of mild cognitive impairment to alzheimer's disease. *Annu Int Conf IEEE Eng Med Biol Soc* 2018; 2018:127–130. doi:10.1109/EMBC.2018.8512273.
- Feng Q, Niu J, Wang L, et al. Comprehensive classification models based on amygdala radiomic features for Alzheimer's disease and mild cognitive impairment. *Brain Imaging Behav* 2021; 15(5):2377–2386. doi:10.1007/s11682-020-00434-z.
- Feng F, Wang P, Zhao K, et al. Radiomic features of hippocampal subregions in alzheimer's disease and amnesic mild cognitive impairment. *Front Aging Neurosci* 2018; 10:290. doi:10.3389/fnagi.2018.00290.
- Shao W, Li X, Zhang J, et al. White matter integrity disruption in the pre-dementia stages of Alzheimer's disease: from subjective memory impairment to amnesic mild cognitive impairment. *Eur J Neurol* 2019; 26(5):800–807. doi:10.1111/ene.13892.
- Yu B, Ding Z, Wang L, et al. Application of diffusion tensor imaging based on automatic fiber quantification in alzheimer's disease. *Curr Alzheimer Res* 2022; 19(6):469–478. doi:10.2174/1567205019666220718142130.
- Power MC, Su D, Wu A, et al. Association of white matter microstructural integrity with cognition and dementia. *Neurobiol Aging* 2019; 83:63–72. doi:10.1016/j.neurobiolaging.2019.08.021.
- Shaffer JL, Petrella JR, Sheldon FC. Alzheimer's disease neuroimaging initiative. predicting cognitive decline in subjects at risk for alzheimer disease by using combined cerebrospinal fluid, MR imaging, and PET biomarkers. *Radiology* 2013; 266(2):583–591. doi:10.1148/radiol.12120010.
- Ranjbar S, Velgos SN, Dueck AC. Alzheimer's disease neuroimaging initiative. brain mr radiomics to differentiate cognitive disorders. *J Neuropsychiatry Clin Neurosci* 2019; 31(3):210–219. doi:10.1176/appi.neuropsych.17120366.
- Aisen PS, Cummings J, Jack Jr CR, et al. On the path to 2025: understanding the Alzheimer's disease continuum. *Alzheimers Res Ther* 2017; 9(1):60. doi:10.1186/s13195-017-0283-5.
- Grueso S, Viejo-Sobera R. Machine learning methods for predicting progression from mild cognitive impairment to Alzheimer's disease

- dementia: a systematic review. *Alzheimers Res Ther* 2021; 13(1):162. doi:[10.1186/s13195-021-00900-w](https://doi.org/10.1186/s13195-021-00900-w).
36. Plant C, Teipel SJ, Oswald A, et al. Automated detection of brain atrophy patterns based on MRI for the prediction of Alzheimer's disease. *Neuroimage* 2010; 50(1):162–174. doi:[10.1016/j.neuroimage.2009.11.046](https://doi.org/10.1016/j.neuroimage.2009.11.046).
37. Misra C, Fan Y, Davatzikos C. Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: results from ADNI. *Neuroimage* 2009; 44(4):1415–1422. doi:[10.1016/j.neuroimage.2008.10.031](https://doi.org/10.1016/j.neuroimage.2008.10.031).
38. Li Y, Jiang J, Lu J, et al. Radiomics: a novel feature extraction method for brain neuron degeneration disease using <sup>18</sup>F-FDG PET imaging and its implementation for Alzheimer's disease and mild cognitive impairment. *Ther Adv Neurol Disord* 2019; 12. doi:[10.1177/1756286419838682](https://doi.org/10.1177/1756286419838682).
39. Zhang D, Shen D, Alzheimer's Disease Neuroimaging Initiative. Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease. *Neuroimage* 2012; 59(2):895–907. doi:[10.1016/j.neuroimage.2011.09.069](https://doi.org/10.1016/j.neuroimage.2011.09.069).
40. Won SY, Park YW, Park M. Quality reporting of radiomics analysis in mild cognitive impairment and alzheimer's disease: a roadmap for moving forward. *Korean J Radiol* 2020; 21(12):1345–1354. doi:[10.3348/kjr.2020.0715](https://doi.org/10.3348/kjr.2020.0715).

## SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found in the online version at [doi:10.1016/j.acra.2022.12.033](https://doi.org/10.1016/j.acra.2022.12.033).