Peer Review File

Article information: https://dx.doi.org/10.21037/atm-21-4349

Reviewer Comments

Comment 1: In this manuscript, Yang and colleagues applied radiomics-based feature extraction algorithms to FDG PET and MRI data of patients from ADNI and Huashan Hospital cohorts to predict progression from mild cognitive impairment (MCI) to Alzheimer's disease dementia (AD). By reading regional values from both modalities, they extracted 68800 features for each subject, which served as the basis for PET, MRI, and combined Cox models. Moreover, a clinical model based on demographic and cognitive data was constructed and compared with each imaging model and their combinations. The authors demonstrate no superiority of the combined model in its predictive value over single modality models.

Reply 1: Thanks for your summary and comments.

Comment 2: This study replicates many existing published studies but does not necessarily expand the understanding of the mechanisms leading to the development of AD. The proposed methods seem to be suitable for solving the research question and are of interest compared to many existing tools for feature extraction. However, the results themselves do not provide new information about the underlying biology or clinical mechanisms of progression and do not increase diagnostic confidence.

Reply 2: Thank you very much for your suggestion.

According to your suggestion, we made the following explanations and changes:

1. Up to now, there are indeed some published studies in the field of AD that use radiomics method (1-6), including ours (7). From our perspectives, the radiomics method is a technology for extracting features, which is different from the traditional analysis method. In the final analysis, it is just a technique, and it cannot explore the pathogenesis in depth. We totally understand and support your suggestion, "*The proposed methods seem to be suitable for solving the research question*…". Identical to previous studies that mentioned above, the current research is a practical study, and the purpose is for clinical application, not focusing on the pathogenesis or mechanism.

2. From another perspective, recent studies in radiomics analysis have suggested that the abnormalities of textural features occur early than traditional volumetric or morphometric features (3,5,6), highlighting the potential use of radiomics analysis as a measure of neurodegenerative processes, which may contain unique information about changes at the microscopic level that can occur before changes at the macroscopic level, such as atrophy. Therefore, we think that there have been subtle changes in the brain regions from which the radiomics features come, and this point explores the mechanism to a certain extent. In other words, we think that some brain regions of patients with "converted MCI" have changed compared with patients with "non-converted MCI", and these regions are potential therapeutic targets, performing specific interventions in these regions

may be able to delay the progression of the disease.

3. As shown in Figure 2 of the main text and Supplementary Table 1, most of the preserved brain regions are obviously related to AD. For example, the medial temporal lobe (hippocampus and hippocampal gyrus), inferior parietal lobule (include the angular gyrus and etc.), precuneus, and cingulated gyrus. These regions are the brain areas of early deposition of AD-related pathological proteins (amyloid-beta or hyperphosphorylated tau proteins (8), and are also brain regions with early atrophy or thickness reduction or metabolism reduction (9-11).

In the third paragraph of the "discussion" section (see Page 14, line 7-17), we have changed the original sentences from "Moreover, when scrutinizing of the feature correlation between different modalities, correlations were observed, included in related features located in the same brain region. For example, there is a correlation between the features of the hippocampus between the two modalities. This suggests that the features reflect true underlying pathophysiological processes within the hippocampus. The study also found that there is a correlation between the hippocampus and other brain areas such as precuneus and the medial cingulate gyrus." to "Consistent with our assumptions, most of the conserved features were identified in regions that were significantly associated with AD. Specifically, the medial temporal areas, inferior parietal lobe, precuneus, and cingulate gyrus are all brain regions that experience early pathological protein (amyloid-β and hyperphosphorylated tau) deposition (8). They also experience early atrophy, thickness reduction, or metabolism reduction (12-14). Moreover, the correlation analyses revealed that some features form different modalities have obvious correlations, further indicating the important roles of some regions. For example, there is a correlation between the hippocampal features of the two modalities, and this region also correlated with other areas including the precuneus and the medial cingulate gyrus; the hippocampus is a typical area of AD, a recent multi-center study suggested that the hippocampal radiomic features can serve as robust biomarkers for clinical application of MCI conversion (15).".

4. For your suggestion "However, the results themselves do not provide new information about the underlying biology or clinical mechanisms of progression and do not increase diagnostic confidence.", we have the following opinions:

Firstly, also taking into consideration of the "comment 3", we have added some results. Compared with the traditional analysis method, such as the average FDG SUVR, the radiomics method can improve the distinguish ability to some extent.

Secondly, we totally understand and agree with your suggestion, our results did not obviously increase the diagnostic confidence when compared with previous results of radiomics-related studies (16,17). However, as we have mentioned in the manuscript (Page 5, line 9-13), "The primary objective of the present study was to explore the potential of a combination of 18F-FDG PET and MRI. To do so, the predictive performance for MCI conversion was compared between these single modalities and a dual-modality approach combining the two. We also investigated the degree to which imaging-derived biomarkers were comparable between these modalities.".

Changes in the text: We have modified our text as advised (see Page 14, line 7-17).

The modifications were marked in the new version.

Comment 3: The study lacks a comparison with existing generally accepted methods for evaluating MRI and PET data. It is unclear if radiomics analysis provides additional value for predicting

conversions. Different results from different cohorts can also indicate a selection of unrepresentative traits. Unfortunately, there is no clear conclusion about which modality should be preferred.

Reply 3: Thank you for your suggestion.

According to your suggestion, we constructed a traditional PET model by calculating the average FDG standardized uptake value ratio (SUVR) and a traditional MRI model by calculating the global gray matter volume (GMV). We compared the superiority of the prediction performance of different modalities.

We have described it with "The traditional PET model was constructed by calculating the average FDG standardized uptake value ratio (SUVR), and an MRI model was constructed by calculating the global gray matter volume (GMV)." in the "Comparison classification" part of the "Materials and Methods" section (see Page 9, line 12-14).

In the fourth paragraph of the "results" section (see Page 11, line 15-20), we have changed the original sentences from "Four prediction models were constructed, including PET model, MRI model, combined model and clinical model. Generally, image-based models (PET/MRI/combined model) were superior to clinical model, combined model was superior to single-modality imaging models (PET and MRI model), MRI model and PET model had comparable prediction performance." to "Six prediction models were constructed, including the SUVR_PET model, GMV_MRI model, radiomics PET model, radiomics MRI model and radiomics_combined model and clinical models. Generally, image-based models (SUVR_PET, GMV_MRI, and radiomics PET/MRI/combined models) were superior to the clinical model, and the radiomics combined model was superior to the single-modality imaging models (radiomics PET and MRI model); conversely, the MRI model and PET model had comparable prediction performances.".

In the second paragraph of the "discussion" section (see Page 13, line 10-13), we have added the sentences as "In this study, four Lasso-Cox models were constructed based on the source of the features, and the performance of the image-based prediction model was superior to the clinical model." to "In this study, six LASSO models were constructed based on the sources of the features, and the performance of the image-based prediction models was superior to that of the clinical model. Notably, radiomics PET/MRI models had better prediction performance compared to traditional PET/MRI models.". Similarly, we have added the C-index results in the third and fourth rows of the table 2.

Based on the above experiments, the radiomics analysis provided additional values for predicting conversions compared to the traditional measures. It is noteworthy that similar results were obtained in the two different cohorts. In conclusion, radiomics combined model showed the most superior prediction performance.

Furthermore, previous studies have reached similar conclusions, that is, the accuracy of radiomics models in predicting the transition from MCI to dementia is higher than that of traditional models (5-7).

Changes in the text: We have modified our text as advised (see Page 9, line 12-14; Page 11, line 15-20; Page 13, line 10-13).

The modifications were marked in the new version.

Comment 4: The feature selection process needs to be revised: by testing 68,800 variables on a patient in a cohort that is significantly smaller than the number of patients, a comprehensive statistical approach is important (e.g. a generalized linear model with repeated measures design). The Tukey HSD is a post hoc multiple comparison procedure, but it does not account for a number of independent tests performed. Moreover, the authors mention that they used ANOVA to test for differences between multiple groups (three or more). What other groups that were tested besides converters and non-converters?

Reply 4: Thank you very much for your suggestion.

According to your suggestion, we made the following changes and explanations:

1. In the previous submission, we described the feature selection process in detail in the supplementary materials. In view of the "comment 11", we have moved the "feature selection" section to the "Radiomic feature extraction and selection" part of the "Materials and Methods" section (see Page 7, line 11-14). As we have mentioned, "Two steps were performed before feature selection: 1) For a 10-fold cross-validation, 90% of the data were set as the training dataset and 10% were set as the test dataset in each fold and 2) Unit restriction for each feature value was eliminated through normalization to zero mean and unit standard deviation.".

2. The ANOVA and Tukey HSD test were not applied in this study. This is an unintentional fault, which does not affect the results and conclusions of this study. We are very sorry for this mistake. In the last paragraph of the "Materials and Methods" section (see Page 10), we have deleted the original sentences "One-way analysis of variance was used to examine the differences between multiple groups (three or more) of continuous variables. Tukey's test was applied to correct for multiple comparisons.".

Changes in the text: We have modified our text as advised (see Page 7, line 11-14; Page 10).

The modifications were marked in the new version.

Comment 5: What was the rationale for including healthy controls in the feature selection algorithm, and how does this relate to the differences between converters and non-converters? What were the inclusion criteria for the controls?

Reply 5: Thank you for your suggestion.

According to your suggestion, we have made the following explanations:

1. The objective of the healthy controls was to select stable features. We collected the first and follow-up MRI and PET images of healthy controls from Cohort A (average interval was two years) to extract the radiomic features at two-time points. Cronbach's alpha coefficient was calculated to investigate the repeated features (i.e. stable features), we added some detailed descriptions in "Radiomic feature extraction and selection" part of "Materials and Methods" section (see Page 7, line 15-18) as "Feature reliability analysis based on the HCs from Cohort A: the Cronbach's alpha coefficient was used to evaluate the stability of the features, and stable features with a Cronbach's alpha coefficient greater than 0.75 were selected. Each HC subject had feature sets at two time points.".

2. About your suggestion "What were the inclusion criteria for the controls?", we have added the inclusion criteria for the controls in the Supplementary Appendix (see Page 1, line 12-13). The

specific sentences are as follows: "The inclusion criteria for the healthy controls were as follows: 1) MMSE scores between 24-30; 2) a CDR of 0; 3) non-depressed, non-MCI, and non-demented.".

Changes in the text: We have modified our text as advised (see main text-Page 7, line 15-18; Supplementary Appendix-Page 1, line 12-13).

The modifications were marked in the new version.

Comment 6: The authors did not discuss the selected features, their interpretation, or the direction of metabolic or anatomical changes, respectively. This should be added to the discussion.

Reply 6:

Thank you very much for your professional comments.

As we have mentioned in the "Conserved features in different modalities" part of the "Results" section, there were "13 conserved features in the PET model, 12 conserved features in the MRI model; and 14 conserved features in the combined model". After removing the duplicate features, there were 16 different features. For space reasons, we did not describe the significance of these features in the previous version. According to your suggestion and the actual situation, we suggested that this information be improved in the supplementary materials, instead of the discussion section.

1. In the "Conserved features in different modalities" part of the "Results" section (see Page 11, line 12-13), we have added a sentence at the end, "In addition, we have explained the meaning of these conserved features and the direction of changes in the Supplementary Materials.".

2. In the supplementary materials, we have added a paragraph, "Conserved features in different modalities" (see Page 6, line 17-22-Page 7, line 1-8), in this section, we have described the meaning of these features, specifically, "After removing the duplicate features, there were 16 different conserved features. The meanings of these features are as follows: the Variance feature extracted from the GLCM category is an indicator of dispersion of the unit values around the mean (18); the Coarseness feature extracted from the NGTDM has been likened to granularity within an image, that is, coarseness is higher in images of larger granularity and lower in those with a smaller granularity (19); the Contrast and Busyness features were both derived from the neighborhood grey-level difference matrix, they describe the local texture features based on the differences between each voxel and the neighboring voxels, specifically, the Contrast correlates with the difference between neighboring regions of voxel intensities, a high Contrast within an image indicates that voxel intensity significantly differs between the neighboring voxels, and the Busyness relates to the change rate between neighborhood intensities weighted by the difference in intensities, and a busy texture is characterized by rapid intensity changes in adjacent voxels (20).".

3. Also considering the second question, we have made changes of the third paragraph of the discussion section of the main text (see Page 14, Line 7-17). We emphasized that the conserved features "all reflect a subtle neurodegeneration", after all, in essence, both structural MRI and FDG-PET reflect nerve injury. Specifically, we have changed the sentences from "Moreover, when scrutinizing of the feature correlation between different modalities, correlations were observed, included in related features located in the same brain region. For example, there is a correlation between the features of the hippocampus between the two modalities. This suggests that the features reflect true underlying pathophysiological processes within the hippocampus. The study also found that there is a correlation between the hippocampus and other brain areas such as precuneus and the

medial cingulate gyrus." to "Consistent with our assumptions, most of the conserved features were identified in regions that were significantly associated with AD. Specifically, the medial temporal areas, inferior parietal lobe, precuneus, and cingulate gyrus are all brain regions that experience early pathological protein (amyloid-β and hyperphosphorylated tau) deposition (8). They also experience early atrophy, thickness reduction, or metabolism reduction (12-14). Moreover, the correlation analyses revealed that some features form different modalities have obvious correlations, further indicating the important roles of some regions. For example, there is a correlation between the hippocampal features of the two modalities, and this region also correlated with other areas including the precuneus and the medial cingulate gyrus; the hippocampus is a typical area of AD, a recent multi-center study suggested that the hippocampal radiomic features can serve as robust biomarkers for clinical application of MCI conversion (15).".

Changes in the text: We have modified our text as advised (see Main text-Page 11, Line 12-13; Page 14, line 7-17; Supplementary Appendix-Page 6, line 17-22-Page 7, line 1-8).

The modifications were marked in the new version.

Comment 7: The clinical model includes MMSE and education (I assume the number of years of study). Have any other cognitive tests been tested? Why does this particular model include age and sex correction rather than imaging models?

Reply 7: Thank you very much for your suggestion.

For the first question, we revisited the ADNI database and the clinical information of enrolled subjects from Huashan Hospital. We found that both the cohort A and B subjects also performed other neuropsychological tests, including the Montreal Cognitive Assessment (MoCA), and Alzheimer' s Disease Assessment Scale-Cognitive 13 (ADAS-Cog 13). We have added other cognitive assessments information in table 1.

As for the second question, age and gender play a key role in clinical models. According to previous studies (7,16), clinical models usually take age and gender as corrected covariates rather than imaging models. In addition, there are significant differences in age in our study, while other cognitive tests do not show any significant differences. Therefore, we do not pay attention to the information of other cognitive tests.

Changes in the text: We have modified our text as advised (see Page 19).

The modifications were marked in the new version.

Comment 8: In reporting the performance of the models, the author points out that the MRI model outperformed PET (C = 0.76 versus 0.734, respectively). Was this difference in model performance significant? On the basis of which test did the authors conclude that the increase in the concordance index of the combined model (C = 0.80 versus 0.76 for MRI, p =?) is not significant and there is no benefit from undergoing both examinations?

Reply 8: Thank you very much for your suggestion

According to your suggestion, we made the following changes and explanations:

We used a two-sample t-test to investigate whether there is a significant difference in C-index between different models. As the result shows, significant differences in C-index were found

between MRI and PET (C-indices of 0.760 versus 0.734, P<0.001), combined model and MRI (C-indices of 0.798 versus 0.760, P<0.001), combined model and PET (C-indices of 0.798 versus 0.734, P<0.001). Thank you for your correction.

In the second paragraph of the "discussion" section (see Page 13, line 17-21), we have changed the original sentences from "However, when comparing the MRI, PET and combined models, it can be seen that for the Huashan test cohort (B) MRI outperformed PET (C-indices of 0.760 vs. 0.734) and that the combined model resulted in only a modest improvement (PET/MRI C-index 0.798)." to "However, comparison of the MRI, PET, and combined models revealed that for the Huashan test cohort (cohort B), MRI outperformed PET (C-index: 0.760 vs. 0.734, P<0.001) and that the combined model resulted in only a modest improvement (C-index for PET/MRI vs. C-index for MRI: 0.798 vs. 0.760, P<0.001; C-index for PET/MRI vs. C-index for PET: 0.798 vs. 0.734, P<0.001).".

Changes in the text: We have modified our text as advised (see Page 13, line 17-21).

The modifications were marked in the new version.

Comment 9: Why did the partial correlation analysis between MRI and PET include educational level as a covariate? Were there any associations observed in these cohorts? Also, a p-value of 0.05 might be too liberal.

Reply 9: Thank you very much for your suggestion.

According to your suggestion, we made the following changes and explanations:

For the first question, there is no significant differences in educational level between converters and non-converters in the experiment, so we removed the covariate of educational level according to your suggestion. Thank you for your comment. In the last paragraph of the "Materials and Methods" section (see Page 10, line 5-7), we have changed the sentences from "To control the age, gender and education effects, partial correlation coefficients were used to evaluate the correlation between PET and MRI radiomic features." to "To control the age and gender effects, partial correlation coefficients were used to evaluate the correlation between PET and MRI radiomic features.".

As for the second question, we did observe associations between the features of different modalities by analyzing correlation between the conserved features from PET modality and MRI modality. For example, there is a correlation between the features of the hippocampus between the two modalities. This suggests that the features reflect true underlying pathophysiological processes within the hippocampus.

In order to make the filter conditions more stringent, we adjust the p-value to 0.01. According to the adjusted experimental results, 83 showed significant correlations in Cohort A (P < 0.01) (Fig.4A); 4 showed significant correlations in Cohort B (P < 0.01) (Fig.4B), among which, there were four features pairs that were both related between different cohorts and modalities. We have updated corresponding figures (Fig.4A &Fig.4B).

In the seventh paragraph of the "results" section (see Page 12, line18-22-Page 13, line 1-2), we have changed the original sentences from "Of these paired correlation matrices, 79 showed significant correlations in Cohort A (P < 0.05) (Fig.4A); 15 showed significant correlations in Cohort B (P < 0.05) (Fig.4B). Among them, there were 11 features pairs that were both related between different cohorts and modalities, as shown in the red box in the figure. These relevant features were from PET

images located in the precuneus (ZSV), hippocampus (ZP, LZLGE) and cingulate cortex (Skewness), and from MRI images located in the hippocampus (ZP, Coarseness), angular gyrus (RP), parietal and inferior edge angular gyrus (Skewness), and olfactory cortex (Coarseness)." to "Of these paired correlation matrices, 83 showed significant correlations in Cohort A ($P < 0.01$) (Fig.4A); four showed significant correlations in Cohort B ($P < 0.01$) (Fig.4B). Among them, there were four features pairs that were both related between different cohorts and modalities, as shown in the red box in the figure. In the PET images, these relevant features were located in the ParaHippocampal (Busyness), Cingulum_Mid (ZP, LZLGE), and Cingulum_Mid (LZE). In the MRI images, these relevant features were located in the angular gyrus (RP), hippocampus (coarseness), and olfactory (coarseness).".

Changes in the text: We have modified our text as advised (see Page 10, line 5-7; Page 12, line18-22-Page 13, line 1-2).

The modifications were marked in the new version.

Comment 10: There are multiple studies that show that partial volume correction increases the differences between groups. Have the authors tried to reproduce their results on uncorrected data? Are the selected features stable?

Reply 10: Thank you very much for your suggestion. We totally agree with you that the PVC step may influence the final result. In general, we chose PVC or not based on the final result. In this study, we also tried without the PVC step. The results are as follows:

The Cox model based on corrected data showed better prediction performance in distinguishing converters from non-converters (0.871 in training set; 0.753 in testing set) when compared with uncorrected results (0.862 in training set; 0.734 in testing set). Therefore, we used the corrected data to build the model.

Because there are already too many experiments in our article, we did not add these results in the main text.

For your second question: we have compared selected features before and after correction by using t-test, here are some changes in the selected features before and after correction, but there is no statistical difference ($p > 0.05$). Therefore, a conclusion can be made that the selected features are basically stable.

The modifications were marked in the new version.

Comment 11: Supplemental material largely repeats the text of the manuscript and may be shortened. In turn, the most important aspect of data processing needs to be moved to the main text.

Reply 11: Thank you for your suggestion. We have moved the feature selection and Cox model construction aspect to the main text.

1. In the "Radiomic feature extraction and selection" part of the "Materials and Methods" section (see Page 7, line 15-22-Page 13, line 1-5), we have changed the original sentences from "The feature selection step was then performed using data from Cohort A. The feature selection steps were as follows: 1) Feature reliability analysis was applied for HCs from Cohort A by calculating Cronbach's alpha coefficient; 2) Statistical test: the most discriminative features between MCI-c group and MCI-nc group were selected using T-test and rank sum test; 3) The top features were selected, and

L1-penalized Cox model was constructed from the training dataset using the least absolute shrinkage and selection operator (LASSO). It is worth noting that after feature selection by LASSO method, the construction of prediction model was also completed. Particularly, 10-fold cross-validation was used with the repetitions of 200 times in feature selection step. Details on feature extraction and selection processes were described in supplementary material." to "The feature selection steps were as follows: 1) Feature reliability analysis based on the HCs from Cohort A: the Cronbach's alpha coefficient was used to evaluate the stability of the features, and stable features with a Cronbach's alpha coefficient greater than 0.75 were selected. Each HC subject had feature sets at two time points. 2) Statistical test: The most discriminative features between the MCI-c group and MCI-nc groups were selected. Thereafter, the t-test and rank sum test were used to identify the features with significant differences between MCI converters and MCI non-converters ($P < 0.01$). 3) The top features were selected, and the L1-penalized Cox model was constructed from the training dataset using the least absolute shrinkage and selection operator (LASSO) regression. LASSO is a robust method that is especially suitable for the regression of high-dimensional features in a radiomics strategy and patient features could be selected based on the associations with the survival endpoints and time (21). It is worth noting that after feature selection by the LASSO method, the construction of the prediction model was also completed. In particular, the 10-fold cross-validation was performed with 200 repetitions in the feature selection step.".

We added some details in our text as "Two steps were performed before feature selection: 1) For a 10-fold cross-validation, 90% of the data were set as the training dataset and 10% were set as the test dataset in each fold and 2) Unit restriction for each feature value was eliminated through normalization to zero mean and unit standard deviation.".

2. We divided the "Cox model construction and comparison classification" into two aspects, including "Cox model construction" and "Comparison classification" (see Page 8, line 6-22-Page 9, line 1-17). The revised text is as follows:

Cox model construction

Our prediction model was the L1-penalized Cox regression model. The Cox model was constructed during the training phase while selecting typical features, and the final model was constructed based on the selected typical features. Cox regression is a statistical analysis method that combines the clinical outcomes with the time of outcome appearance. In this study, the clinical outcome was the conversion of MCI to AD. The time of outcome appearance was defined as the interval between the baseline timepoint and the endpoint. For each subject, the baseline timepoint was the time of MRI and PET data collection and the endpoint was the time of AD diagnosis for MCI-c or the last follow-up time point for MCI-nc.

The Cox model was constructed in R (http://www.R-project.org/) using the glmnet and survival packages FRIDMAN 2010; SIMON 2011 THERNEAU 2013. The prediction performance of the model was evaluated using the Harrell's consistency coefficient (C-index); the C-index was calculated for the training and test datasets. The Cox model was used in the test dataset to calculate the prognostic index (PI) for each subject. PI is a linear combination of the selected features and their coefficients. The C-index of the test dataset was calculated according to the PI. To evaluate the prediction performance of the Cox model more reasonably and in an unbiased manner, the 10-fold cross-validation was repeated 20 times, and the average C-index (with the standard deviation values) was calculated thereafter. In addition, we also counted the number of times each feature participated

in model construction, and selected features that were included in the prediction model for more than two-thirds of the time. These “conserved” features were used for further analyses.

To further compare the prediction performance of these features from the imaging modality (PET and MRI), the PI of each subject was calculated according to the corresponding model. Using the median PI, the individuals were then stratified into high-risk and low-risk prognostic groups. The log-rank test was employed to examine the survival differences between the different risk groups, and Kaplan–Meier survival curves were also plotted.

Comparison classification

To compare the prediction performances of the different modalities, the single-modality Cox model was constructed using features from MRI or PET alone, while the dual-modality Cox model was constructed using combined features. The traditional PET model was constructed by calculating the average FDG standardized uptake value ratio (SUVR), and an MRI model was constructed by calculating the global gray matter volume (GMV). Further, a clinical Cox model was constructed using the available clinical variables (age, sex, education [number of years of study], and MMSE scores). A clinical model was established to compare the effects of the imaging factors and basic clinical factors on the probability of MCI conversion.

Changes in the text: We have modified our text as advised (see Page 7, line 15-22-Page 13, line 1-5; Page 8, line 6-22-Page 9, line 1-17).

The modifications were marked in the new version.

References

1. de Oliveira MS, Balthazar ML, D'Abreu A, et al. MR imaging texture analysis of the corpus callosum and thalamus in amnestic mild cognitive impairment and mild Alzheimer disease. AJNR Am J Neuroradiol 2011;32:60-6.

2. Li S, Yuan X, Pu F, et al. Abnormal changes of multidimensional surface features using multivariate pattern classification in amnestic mild cognitive impairment patients. J Neurosci 2014;34:10541-53.

3. Feng F, Wang P, Zhao K, et al. Radiomic Features of Hippocampal Subregions in Alzheimer's Disease and Amnestic Mild Cognitive Impairment. Front Aging Neurosci 2018;10:290.

4. Luk CC, Ishaque A, Khan M, et al. Alzheimer's disease: 3-Dimensional MRI texture for prediction of conversion from mild cognitive impairment. Alzheimers Dement (Amst) 2018;10:755-63.

5. Sørensen L, Igel C, Liv Hansen N, et al. Early detection of Alzheimer's disease using MRI hippocampal texture. Hum Brain Mapp 2016;37:1148-61.

6. Lee S, Lee H, Kim KW. Magnetic resonance imaging texture predicts progression to dementia due to Alzheimer disease earlier than hippocampal volume. J Psychiatry Neurosci 2020;45:7-14.

7. Li TR, Wu Y, Jiang JJ, et al. Radiomics Analysis of Magnetic Resonance Imaging Facilitates the Identification of Preclinical Alzheimer's Disease: An Exploratory Study. Front Cell Dev Biol 2020;8:605734.

8. Long JM, Holtzman DM. Alzheimer Disease: An Update on Pathobiology and Treatment Strategies. Cell 2019;179:312-39.

9. Kato T, Inui Y, Nakamura A, et al. Brain fluorodeoxyglucose (FDG) PET in dementia. Ageing Res Rev 2016;30:73-84.

10. Pini L, Pievani M, Bocchetta M, et al. Brain atrophy in Alzheimer's Disease and aging. Ageing Res Rev 2016;30:25-48.

11. Risacher SL, Saykin AJ. Neuroimaging in aging and neurologic diseases. Handb Clin Neurol 2019;167:191-227.

12. Kato T, Inui Y, Nakamura A, et al. Brain fluorodeoxyglucose (FDG) PET in dementia. Ageing Research Reviews 2016;30:73-84.

13. Pini L, Pievani M, Bocchetta M, et al. Brain atrophy in Alzheimer’s Disease and aging. Ageing Research Reviews 2016;30:25-48.

14. Risacher SL, Saykin AJ. Chapter 12 - Neuroimaging in aging and neurologic diseases. In: Dekosky ST, Asthana S, editors. Handbook of Clinical Neurology. Elsevier; 2019. p. 191-227.

15. Zhao K, Ding Y, Han Y, et al. Independent and reproducible hippocampal radiomic biomarkers for multisite Alzheimer’s disease: diagnosis, longitudinal progress and biological basis. Science Bulletin 2020;65:1103-13.

16. Zhou H, Jiang J, Lu J, et al. Dual-Model Radiomic Biomarkers Predict Development of Mild Cognitive Impairment Progression to Alzheimer's Disease. Front Neurosci 2018;12:1045.

17. Feng Q, Ding Z. MRI Radiomics Classification and Prediction in Alzheimer's Disease and Mild Cognitive Impairment: A Review. Curr Alzheimer Res 2020;17:297-309.

18. Pantic I, Jeremic R, Dacic S, et al. Gray-Level Co-Occurrence Matrix Analysis of Granule Neurons of the Hippocampal Dentate Gyrus Following Cortical Injury. Microscopy and Microanalysis 2020;26:166-72.

19. Cheng N-M, Dean Fang Y-H, Tung-Chieh Chang J, et al. Textural Features of Pretreatment ¹⁸F-FDG PET/CT Images: Prognostic Significance in Patients with Advanced T-Stage Oropharyngeal Squamous Cell Carcinoma. Journal of Nuclear Medicine 2013;54:1703.

20. Ahn HK, Lee H, Kim SG, et al. Pre-treatment 18F-FDG PET-based radiomics predict survival in resected non-small cell lung cancer. Clinical Radiology 2019;74:467-73.

21. Gui J, Li H. Penalized Cox regression analysis in the high-dimensional and low-sample size settings, with applications to microarray gene expression data. Bioinformatics 2005;21:3001-8.