Independent and reproducible hippocampal radiomic biomarkers for multisite Alzheimer's disease: diagnosis, longitudinal progress and biological basis

Kun Zhao

Shandong normal university

Yanhui Ding

Shandong normal university

Ying Han

Xuanwu Hospital

Yong Fan

University of Pennsylvania

Aaron F. Alexander-Bloch

Yale University School of Medicine

Tong Han

Tianjin Huanhu Hospital

Dan Jin

Institute of Automation Chinese Academy of Sciences

Bing Liu

Institute of Automation Chinese Academy of Sciences

Jie Lu

Xuanwu Hospital

Chengyuan Song

Shandong University Qilu Hospital

Pan Wang

Tianjin Huanhu Hospital

Dawei Wang

Shandong University Qilu Hospital

Qing Wang

Shandong University Qilu Hospital

Kaibin Xu

Institute of Automation Chinese Academy of Sciences

Hongwei Yang

Xuanwu Hospital

Hongxiang Yao

Chinese PLA General Hospital

Yuanjie Zheng

Shandong normal university

Chunshui Yu

Tianjin Medical University General Hospital

Bo Zhou

Chinese PLA General Hospital

Xinqing Zhang

Xuanwu Hospital

Yuying Zhou

Tianjin Huanhu Hospital

Tianzi Jiang

Institute of Automation Chinese Academy of Sciences

Xi Zhang

Chinese PLA General Hospital

Yong Liu (✉ yliu@nlpr.ia.ac.cn)

Institute of Automation Chinese Academy of Sciences <https://orcid.org/0000-0002-1862-3121>

Research

Keywords: Hippocampal radiomic features, Multisite Alzheimer's disease MRI, Independent cross-validation, Brain biomarker, Biological basis

Posted Date: November 27th, 2019

DOI: <https://doi.org/10.21203/rs.2.17749/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 Science Bulletin on July 1st, 2020. See the published version at <https://doi.org/10.1016/j.scib.2020.04.003>.

Abstract

Background: Hippocampal morphological change is one of the main hallmarks of Alzheimer's disease (AD). The primary aim of this study is to explore whether radiomics feature is a robust biomarker for AD and further explore the biological basis of those features.

Methods: Hippocampal radiomic features were extracted for classification and prediction using a support vector machine (SVM) with 1943 subjects (multi-site).

Results: Multivariate classifier-based SVM analysis provided individual-level predictions for distinguishing AD patients (N=261) from normal controls (NCs; N=231) with accuracy=88.21% with inter-site cross-validation. Further analyses of a large, independent ADNI dataset (N=1228) reinforced these findings. In mild cognitive impairment (MCI) groups, a systemic analysis demonstrated that the identified features were significantly associated with clinical features (e.g., apolipoprotein E (APOE) genotype, polygenic risk scores, cerebrospinal fluid (CSF) A β , CSF Tau), and longitudinal changes in cognition ability; more importantly, the radiomic features have a consistently altered pattern with changes in the MMSE scores over 5 years of follow-up.

Conclusion: These comprehensive results suggest that hippocampal radiomic features can serve as robust biomarkers for clinical applications in AD/MCI.

1. Background

Convergence magnetic resonance imaging (MRI)-based biomarkers that target gray matter atrophy or shape alterations are the most commonly used measures for early biomarker detection of Alzheimer's disease (AD) [1–3]. These markers have been used to perform classification analyses that distinguish AD patients from normal controls (NCs) with 80–90% accuracy that has reached 95% in several small sample studies [4]. However, due to the limited sample size, the reproducibility and generalizability of these results are debatable; nevertheless this kind of robustness is the most fundamental property for clinical translation [5–8].

Hippocampal atrophy, or shape change, is one of the main hallmarks of AD [9]. However, the volume and/or shape are only crude proxies for the complex anatomical changes that occur in AD, and studies of atrophy often ignore the fact that this process is not uniform across different disease phases [10]. The hippocampus undergoes microstructural changes before severe atrophy, and pathological changes such as neurofibrillary tangles (NFTs) and amyloid- β (A β) are not directly detectable at the current resolution of clinical MRI [11]. Thus, novel MRI analyses that yield greater information about subtle changes in the hippocampus would be a significant contribution.

Radiomics, a method of texture analysis, provides information about first-, second-, and higher-order morphological features [12–14]. Texture analysis includes a variety of image analysis techniques that quantify the variations in surface intensity or patterns, including some that are imperceptible to the

human visual system [15], and is a useful way to extract detail information from brain images, increase the precision of diagnosis, and assess prognosis [11, 14, 16–19].

The apolipoprotein E (APOE) gene is known as a major genetic risk factor for AD [20]. Polygenic risk scores (PGRSs), comprehensive indicators that combine multiple risk alleles, provide a quantitative measure of genetic disease risk [21]. Previous large-scale genome-wide association studies found an association between the PGRSs for AD and the structural indices of certain brain regions [22–24]. During the progression of AD, A β plaques are considered to occur in the early stage, while Tau accumulation is considered to be the main factor underlying later dysfunction [25, 26]. However, the association between the imaging biomarkers and these neurobiological measures is not yet clear.

Inspired by these studies, the first aim of this study is to explore whether hippocampal radiomic features with reproducible change patterns can serve as MRI biomarkers of AD using multisite MRI data (715 subjects from 6 sites). To further assess the reproducibility and generalization of the findings, an independent dataset from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) (N = 1228) was included. The second hypothesis is that these hippocampal radiomic features or classification outputs have a solid neurobiological basis, thereby improving the power for individual diagnosis or prediction. For this purpose, we investigate the neurobiological basis of the identified hippocampal radiomic features by relating these features to other variables, including APOE, PGRS, CSF A β , CSF Tau, and the progression of disease status of MCI subjects. The results confirm that hippocampal radiomic features are robust neuroimaging biomarkers with a solid neurobiological basis that are useful for diagnosing AD and predicting the likelihood of progression from MCI; therefore, they have translation potential.

2. Materials And Methods

2.1 Subjects

This study included 715 subjects (231 NCs, 223 MCI and 261 AD patients) as in-house data from six different MRI scanners. The replicated data used in this work were obtained from the ADNI (www.loni.ucla.edu/ADNI). Of these 1228 subjects (356 NCs, 550 MCI and 322 AD), 708 (182 NCs, 328 MCI and 198 AD) had genetic information as well as A β and Tau pathology. Table 1 lists the demographic and clinical information, and additional details are provided in the supplementary material S01.

2.2 Hippocampal segmentation and feature extraction

For each subject, T1 MRI scans were registered to the Montreal Neurological Institute (MNI) space using affine registration and resampled to 1 mm×1 mm×1 mm. The hippocampus was then segmented bilaterally using an revised segmentation method based on multi-atlas based local label learning (LLL) (<https://www.nitrc.org/projects/locallabel>), an automatic segmentation method including the N3 bias correction of imaging inhomogeneity [27, 28]. For each side, we then extracted 495 features, including

intensity-based features, shape-based features and texture-based features, across 8 wavelet-based frequency domains. The definitions and detailed descriptions can be found in previous publications[19, 29], and are listed in the supplementary material S11.

2.3 Redundancy removal and statistical analysis

Because shape features are the same in different frequency domains of the wavelet transform, duplicate features were removed before subsequent analysis. To reduce site effects, we first tested for group differential analysis at each site and then performed meta-analytic tests to integrate the multisite results [30]. At each site, statistical significance was determined using two-sample two-sided t-tests for any two groups after the age and gender effects were regressed out using a linear regression model. We then used the Liptak-Stouffer z-score method [31] to integrate the results [32, 33]. Briefly, the p-value of each feature at the *i*-th site was converted to the corresponding Z_i score using the following formula: $Z_i = \Phi^{-1}(1-p_i)$, where Φ^{-1} is the inverse of the standard normal cumulative density function. Then, a combined z-score for each feature was obtained using the Liptak-Stouffer formula as follows:

[Due to technical limitations, the formula could not be displayed here. Please see the supplementary files section to access the formula.]

where w_i is the inverse of the variance in z_i , which represents a relative measure of the statistical power compared with the other datasets. The z-scores were expected to follow a standard normal distribution under the null hypothesis. Using this method, we calculated the p-value according to the corresponding z-scores for the tests of group difference between the NC, MCI and AD groups. The correlation between the t-score of each feature in the ADNI dataset and the z-score of each feature in the in-house dataset was computed to verify the reproducibility of the results.

2.4 Classification analysis, validation and generalizability

To assess the multivariate performance of the radiomic features, we established a support vector machine (SVM) model to classify the AD patients and NCs. Specifically, for each feature in each center, we first introduce a common min-max feature normalization scheme, a nonlinear SVM with a radial basis function (RBF) kernel was constructed using LIBSVM (<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>), and SVM-recursive feature elimination (SVM-RFE) was used for feature selection. The classification analysis was evaluated with the inter- and intra-site cross-validation methods [34]. We next tested the generalizability of the classification with radiomic features in the ADNI data, two additional independent cross-validation steps were performed. 1) In-house data were used as the training set, and ADNI data were applied as the testing set (CV3), 2) Conversely, the ADNI data was used as the training set, and in-house data was used as the testing set (CV4) (Table 2, supplementary material S02-S04).

The classification performance was evaluated by means of several accuracy metrics [accuracy (ACC), sensitivity (SEN) and specificity (SPE)] and the areas under the receiver operating characteristic (ROC)

curves (AUCs) [35]. To further verify the clinical relevance of the radiomic-based classification, we also investigated the correlations between the classifier output (decision score) and the cognitive ability scores of individual subjects in the test sets.

2.5 Relationship between radiomic features and cognitive ability, APOE genetics, A β , Tau, and longitudinal changes in the cognitive abilities in MCI group

In this study, the conservative features, which were defined as the overlap of the altered and conserved radiomic features, were obtained from group meta-analysis and classification analysis (Figure 1). Later, we continue to explore the relationship between the clinical information and the conservative features as well as the uniformity with longitudinal changes in cognitive ability. To assess the association between the radiomic features and cognitive ability, Pearson's correlation coefficients were calculated between the features and the MMSE scores in individual subjects, both combining AD and MCI and treating each group separately. We also perform these correlation analyses in the ADNI and in-house cohorts to further evaluate the generalizability of the identified features.

We also evaluated the differences in the identified features between A β ⁺ and A β ⁻ in the MCI subgroups (also between Tau⁺ and Tau⁻ and between APOE ϵ 4⁺ and APOE ϵ 4⁻). A β and Tau occasionally coexist, which greatly increases the risk for AD [25, 36, 37]. Thus, we also evaluated the differences in the identified features in three MCI subgroups (A β ⁺&Tau⁺, A β ⁻&Tau⁺ or A β ⁺&Tau⁻, and A β ⁻&Tau⁻). To explore the relationship between A β /Tau and the radiomic features, Pearson's correlations between the radiomic features and A β /Tau were also evaluated in the MCI groups (supplementary material S05).

To explore the relationship between the radiomic features and PGRSs, we used 533 subjects with genome-wide single-nucleotide polymorphisms. For each subject, we used the “score” utility in PLINK [38] and recent summary statistics for AD [22] to compute the polygenic AD risk score. The Spearman's correlations between the radiomic features and the PGRSs were also calculated after accounting for the regression group and race effects (supplementary material S06).

We also evaluated whether the changes in these hippocampal radiomic features were associated with changes of MMSE in progressive MCI (PMCI; N = 168, interscan interval: 33.01 $\pm$ 24.57 months) and stable MCI (SMCI; N = 134, interscan interval: 38.61 $\pm$ 29.22 months) subjects (Table S11) to assess the feasibility of using these features to predict disease progression and to predict whether the MCI subject would progress to AD. To continue exploring the different patterns of the changes in the radiomic features of the PMCI and SMCI subjects, 35 PMCI subjects and 29 SMCI subjects with more than five points of longitudinal data and a time interval between each visit greater than 1 year were selected for further analyses of the relationship between the changes in the MMSE scores and the changes in the identified radiomic feature (supplementary material S07-S08).

3. Results

3.1 Hippocampal radiomic features differing among the AD, MCI and NC subjects

After redundancy removal, 431 hippocampal features were used in each hemisphere. Of these 862 features, 310 were significantly altered in the AD patients compared to those in the NCs. In summary, 13 intensity features, 7 shape features and 22 textural features differed in at least one frequency domain (supplementary material S09). In addition, 108 features were significantly different between the MCI and NC subjects, and 177 features were significantly different between the AD and MCI subjects ($P < 0.05$, Bonferroni-corrected with $N = 862$) (Figure 2a-c). The results showed that the difference maps were significantly correlated between the ADNI and in-house databases (all $P < 10^{-114}$), which validates the reproducibility of the radiomic features (Figure 2d-f).

3.2 Classification performance

The inter-site cross-validation showed an AUC = 0.95 (ACC = 0.88) with the default parameters ($c = 2$, $g = 1/N$, N : number of features) that was obtained by the SVM-RFE with the 57 selected features (when selecting 50–250 features, $ACC = 0.87 \pm 0.01$; Figure S6). To evaluate the accuracy in each pair of sites, we used the same method as CV1 to validate the robustness of the radiomic features, revealing a mean ACC higher than 0.84 (supplementary material S04). With the intra-site cross-validation, the mean AUC = 0.95 ± 0.02 (Figure 3a–3b, Table 2).

When using the in-house data as the training data and the ADNI data as the testing data, we still achieved a higher AUC = 0.84 (ACC = 0.79) (Figure 3c, Table 2); when selecting 50–250 features, $ACC = 0.78 \pm 0.01$. Conversely, when using the in-house data as the testing set, we obtained an AUC = 0.92 (ACC = 0.84) (Figure 3d, Table 2); when selecting 50–250 features, $ACC = 0.82 \pm 0.005$. Moreover, the classifier output in the test data was significantly negatively correlated with the MMSE score in the AD and MCI groups, in only the AD group and in only the MCI group in the testing datasets (Figure 3e–3f).

3.3 Radiomic features and cognitive ability, APOE genotype, A β , Tau, and longitudinal cognitive changes in the MCI subjects

For the inter-site cross-validation, we chose features that were included in the classifier more than two thirds of the time. This resulted in 47 ‘conserved’ features for further analysis. Of these features, the 40 that were significantly different between the AD and NC groups were defined as conservative features ($P < 0.05$, Bonferroni correction $N = 47$). Of these identified features, 33 showed significant correlations with the MMSE scores in the AD and MCI groups ($P < 0.05$) (Figure 4a). More importantly, almost all of these results (>95%) were replicated in the ADNI datasets (Figure 4c), and the correlation scores for the in-house data were significantly correlated with those for the ADNI dataset ($P = 4.86 \times 10^{-104}$).

Of the 33 identified features, 8 were significantly different between the APOE $\epsilon 4+$ subjects (N = 177) and the APOE $\epsilon 4-$ subjects (N = 151) ($P < 0.05$) (Figure 5, Tables S3, S5). Twenty-one of the 33 features were significantly associated with the PGRSs after accounting for the regression group and the influence of race (Figure 5, Table S10). For A β , we defined the A β + group as A β < 675 (mean - 0.5std, N = 86) and the A β - group as A β > 1017 (mean + 0.5std, N = 129). Nineteen of the 33 features were significantly different between the A β - and A β + subjects, and 20 features showed significant correlations with the A β scores in the MCI subjects (N = 328, $P < 0.05$) (Figure 5a, Tables S5, S8). For the Tau pathology, Tau+ was defined as Tau > 349 (N = 88), and Tau- was defined as Tau < 221 (N = 116). Unsurprisingly, 19 features were significantly different between the Tau+ subjects and Tau- subjects, and 22 features had a significant correlation with the Tau scores in the MCI subjects ($P < 0.05$) (Figure 5a, Tables S5, S8). When comparing the subjects who were A β +&Tau+ (N = 43) to those who were A β -&Tau- (N = 52), 21 of the 33 features were significantly different. The subjects who were A β + / Tau- or A β - / Tau+ were intermediate between these two (Figure 5b, Table S7).

For the 168 PMCI subjects, 24 features showed significant correlations with the changes in feature scores and the changes in MMSE ($P < 0.05$, uncorrected). For the 134 SMCI subjects included, these features demonstrated significant but weaker correlations compared with those of the PMCI subjects ($P < 0.05$, uncorrected) (Figure 5a, Table S12). In addition, AUC improved from 0.65 to 0.82 by using logistical regression to predict the conversion of MCI subjects by adding radiomic features one by one until the AUC did not increase (Figure 5c, S07). For subjects with >5 visits, the changes in the radiomic features had a highly uniform variation trend with the changes in the MMSE scores in the PMCI subjects (N = 34) and SMCI subjects (N = 29) (Figure 5d, Figure S10).

4. Discussion

In a large-scale analysis of data pooled across sites, we demonstrated that radiomic features appear to be a robust, reproducible and generalizable imaging signature of AD using four types of cross-validations with widely used machine learning techniques. Moreover, these features are related to neurobiological substrates underlying the progress of AD and cognitive decline progression in MCI. This is of great significance for early clinical diagnosis or prognostic follow-up in AD.

The multisite MRI findings of altered hippocampal radiomic features in AD confirm and extend previous neuroimaging findings [11, 14, 19, 28]. Our results show robust and reproducible AD-related alterations in intensity features, such as kurtosis, mean, mad, median, entropy and uniformity; these features reflect the properties or distribution of gray matter within the hippocampus in AD patients [11], which may indicate atrophy or gray matter loss in the hippocampus [39, 40]. Interestingly, some shape features, such as the area, compactness and surface-to-volume ratio, were also altered in AD/MCI; these findings replicate those of previous studies [16, 41–43], and these phenomena indicate that atrophy of the hippocampus does not occur collaterally [44, 45]. Textural features, such as long run emphasis (LRE), gray level nonuniformity (GLN), and run length nonuniformity (RLN), were the most significantly different features between the two groups. Although pathological features of AD, such as NFTs and A β plaques, cannot be

detected by MRI, these microstructural changes might lead to altered textural patterns detectable via texture analysis [11, 46–48]. In support of this, the 27 identified features showed significant associations with the positron emission tomography (PET) amyloid value in the hippocampus (supplementary material S10). A subset of radiomic features was also significantly associated with high-risk genetic status, A β and Tau deposition, and changes in cognitive ability. These results highlight that the altered radiomic features have a solid neurobiological basis and confirm that with “radiomics: images are more than pictures, as they are also data” [13].

Neuroimaging genetics has moved from establishing heritable phenotypes to finding genetic markers associated with imaging phenotypes. APOE polymorphic alleles are the main genetic determinants of AD risk [49, 50]. As expected, this study shows significant radiomic differences between APOE ϵ 4+ and APOE ϵ 4- in MCI subjects. Unlike a single gene focus, the PGRS has been proposed to have improved predictive ability and statistical power. The PGRS of AD has been found to be associated with cognitive decline and brain imaging measures, highlighting the fact that elevated genetic risk influences traits even among individuals without dementia [23, 24, 51–53]. Hence, the association between the PGRS and radiomic features highlights that hippocampal textures can be used to predict whether a subject has an accelerated genetic basis for progressing from MCI to AD. A β plaques and Tau NFTs are the hallmark lesions of AD, and the strong association between radiomic features and the progression of A β and/or Tau in the MCI subjects indicates the strong neurobiological substrate of these features [2, 11, 16, 54]. It is still a significant challenge to evaluate and predict the progress of MCI [28, 55]. A particularly important finding is that the changes in the pattern of the identified features was strongly correlated with the changes in the MMSE scores of the PMCI and SMCI groups over five years. This provide powerful evidence for using features to follow the progress of high-risk groups, which has important clinical implications, although there is no definite marker to determine this eventual progress at the present time.

MRI studies aimed at identifying robust disease biomarkers require large numbers of samples, which can be difficult to obtain from a single site [56], and the generalizability of models to larger multisite datasets is an important step to increasing the statistical power necessary to identify biomarkers in translational neuroimaging [6, 8]. It should be noted that the ADNI cohort, which has a large sample (more than 1500), was collected from more than 50 centers [57, 58], and the mean sample size is relatively small (Figure S1). Importantly, because these results were validated with data from independent sites, our approach is expected to have greater generalizability to future datasets compared to a single site study using internal validation methods such as leave-N-out cross-validation [4, 34]. To some degree, better classification might potentially be achieved via parameter tuning, feature selection/combination, or other methods, such as deep learning. However, these benefits must be balanced with the risk of overfitting and reduction of the generalizability of the classification approach to novel data [35]. More significantly, the negative correlation between the classifier’s output and the MMSE implies that the more severe the disease is, the easier it is to recognize, which implies a potential bridge between medical imaging and personalized medicine [59].

Limitations and caveats

Despite these advances, this study has several limitations that should be considered. First, the most obvious advantage of a multisite study is the large number of subjects with more generalizable information than a single site. Most but not all of them are MCI subjects, and these large pools of data come from different sites, which inevitably leads to the results being affected by the inhomogeneity of the subjects. Despite the increasing number of multisite MRI studies, the problem of site/scanner confounds has not been entirely addressed, although a recent study provides a statistical framework to approach this issue by correcting distributional shifts between datasets [60]. Although the classification results are encouraging, the accuracy may improve with the use of a more powerful classification approach, such as deep learning, with more samples [61]. We should rethink how the normalization and segmentation methods affect the radiomics measures, although we have confirmed the test-retest probability of these measures based on the LLL method [62]. Finally, combining markers from other brain regions will help identify robust and reproducible biomarkers for clinical applications [4, 8, 63–65].

5. Conclusion

This systematic study highlights the presence of hippocampal textural abnormalities in AD and the possibility that textures can serve as neuroimaging biomarkers for AD for further clinical applications.

6. Declarations

6.1 Ethics approval and consent to participate

These studies were approved by the Medical Ethics Committee of the local Hospitals. All subjects or their legal guardians signed written consent forms.

6.2 Consent for publication

Not applicable.

6.3 Availability of data and materials

Individual hippocampal maps and scripts are available at <https://github.com/YongLiulab>, and other information is available from the corresponding author.

6.4 Competing interests

The authors declare that they have no competing interests.

6.5 Funding

This work was partially supported by the National Key Research and Development Program of China (grant no. 2016YFC1305904), the Strategic Priority Research Program (B) of the Chinese Academy of Sciences (grant no. XDB32020200), the National Natural Science Foundation of China (grant nos. 81871438, 81901101, 61633018, 81571062, 81400890, 61431012), and the Primary Research & Development Plan of Shandong Province (grant no. 2017GGX10112).

6.6 Authors' contributions

KZ, YD, YF and YL analyzed the data and performed the measurements; YH, TH, JL, CS, PW, DW, QW, HW Y, HX Y, CY, BZ, XQ Z, YYZ, and XZ collected the data; KZ and YL were mainly responsible for preparing the manuscript; AA, DJ, BL, KX, YJZ, CY and TJ revised the manuscript; and YL supervised the project. All authors read and approved the final manuscript.

6.7 Acknowledgements

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and generous contributions from AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research provide funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. The ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

We sincerely thank Dr. Tianming Liu and Dr. Feng Shi for kindly helping to prepare this manuscript.

References

- 1.Hampel H, Burger K, Teipel SJ, Bokde AL, Zetterberg H, Blennow K. Core candidate neurochemical and imaging biomarkers of Alzheimer's disease. *Alzheimers Dement*. 2008;4(1):38–48.
- 2.Nakamura A, Kaneko N, Villemagne VL, Kato T, Doecke J, Dore V, et al. High performance plasma amyloid-beta biomarkers for Alzheimer's disease. *Nature*. 2018;554(7691):249–54.
- 3.Caldwell JZK, Berg JL, Cummings JL, Banks SJ, Alzheimer's Disease Neuroimaging I. Moderating effects of sex on the impact of diagnosis and amyloid positivity on verbal memory and hippocampal volume. *Alzheimers Res Ther*. 2017;9(1):72.
- 4.Rathore S, Habes M, Iftikhar MA, Shacklett A, Davatzikos C. A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. *Neuroimage*. 2017;155:530–48.
- 5.Xia M, Si T, Sun X, Ma Q, Liu B, Wang L, et al. Reproducibility of functional brain alterations in major depressive disorder: Evidence from a multisite resting-state functional MRI study with 1,434 individuals. *Neuroimage*. 2019;189:700–14.
- 6.Varoquaux G. Cross-validation failure: Small sample sizes lead to large error bars. *Neuroimage*. 2018;180(Pt A):68–77.
- 7.Teipel SJ, Wohler A, Metzger C, Grimmer T, Sorg C, Ewers M, et al. Multicenter stability of resting state fMRI in the detection of Alzheimer's disease and amnesic MCI. *Neuroimage Clin*. 2017;14:183–94.
- 8.Woo CW, Chang LJ, Lindquist MA, Wager TD. Building better biomarkers: brain models in translational neuroimaging. *Nat Neurosci*. 2017;20(3):365–77.
- 9.Shi F, Liu B, Zhou Y, Yu C, Jiang T. Hippocampal volume and asymmetry in mild cognitive impairment and Alzheimer's disease: Meta-analyses of MRI studies. *Hippocampus*. 2009;19(11):1055–64.
- 10.Young AL, Marinescu RV, Oxtoby NP, Bocchetta M, Yong K, Firth NC, et al. Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with Subtype and Stage Inference. *Nat Commun*. 2018;9(1):4273.
- 11.Sorensen L, Igel C, Liv Hansen N, Osler M, Lauritzen M, Rostrup E, et al. Early detection of Alzheimer's disease using MRI hippocampal texture. *Hum Brain Mapp*. 2016;37(3):1148–61.
- 12.Parmar C, Leijenaar RT, Grossmann P, Rios Velazquez E, Bussink J, Rietveld D, et al. Radiomic feature clusters and prognostic signatures specific for Lung and Head & Neck cancer. *Sci Rep*. 2015;5:11044.
- 13.Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology*. 2016;278(2):563–77.

- 14.Chaddad A, Desrosiers C, Niazi T. Deep Radiomic Analysis of MRI Related to Alzheimer's Disease. IEEE Access. 2018;6:58213–21.
- 15.Kassner A, Thornhill RE. Texture analysis: a review of neurologic MR imaging applications. AJNR Am J Neuroradiol. 2010;31(5):809–16.
- 16.Sorensen L, Igel C, Pai A, Balas I, Anker C, Lillholm M, et al. Differential diagnosis of mild cognitive impairment and Alzheimer's disease using structural MRI cortical thickness, hippocampal shape, hippocampal texture, and volumetry. Neuroimage Clin. 2017;13:470–82.
- 17.Chincarini A, Bosco P, Calvini P, Gemme G, Esposito M, Olivieri C, et al. Local MRI analysis approach in the diagnosis of early and prodromal Alzheimer's disease. Neuroimage. 2011;58(2):469–80.
- 18.de Oliveira MS, Balthazar ML, D'Abreu A, Yasuda CL, Damasceno BP, Cendes F, et al. MR imaging texture analysis of the corpus callosum and thalamus in amnesic mild cognitive impairment and mild Alzheimer disease. AJNR Am J Neuroradiol. 2011;32(1):60–6.
- 19.Feng F, Wang P, Zhao K, Zhou B, Yao H, Meng Q, et al. Radiomic Features of Hippocampal Subregions in Alzheimer's Disease and Amnesic Mild Cognitive Impairment. Front Aging Neurosci. 2018;10:290.
- 20.Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet. 2013;45(12):1452–8.
- 21.Torkamani A, Wineinger NE, Topol EJ. The personal and clinical utility of polygenic risk scores. Nat Rev Genet. 2018;19(9):581–90.
- 22.Jansen IE, Savage JE, Watanabe K, Bryois J, Williams DM, Steinberg S, et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. Nat Genet. 2019;51(3):404–13.
- 23.Mormino EC, Sperling RA, Holmes AJ, Buckner RL, De Jager PL, Smoller JW, et al. Polygenic risk of Alzheimer disease is associated with early- and late-life processes. Neurology. 2016;87(5):481–8.
- 24.Axelrud LK, Santoro ML, Pine DS, Talarico F, Gadelha A, Manfro GG, et al. Polygenic Risk Score for Alzheimer's Disease: Implications for Memory Performance and Hippocampal Volumes in Early Life. Am J Psychiatry. 2018;175(6):555–63.
- 25.Sepulcre J, Grothe MJ, d'Oleire Uquillas F, Ortiz-Teran L, Diez I, Yang HS, et al. Neurogenetic contributions to amyloid beta and tau spreading in the human cortex. Nat Med. 2018;24(12):1910–8.
- 26.Kim HJ, Park S, Cho H, Jang YK, San Lee J, Jang H, et al. Assessment of Extent and Role of Tau in Subcortical Vascular Cognitive Impairment Using 18F-AV1451 Positron Emission Tomography Imaging. JAMA Neurol. 2018;75(8):999–1007.

- 27.Hao Y, Wang T, Zhang X, Duan Y, Yu C, Jiang T, et al. Local label learning (LLL) for subcortical structure segmentation: application to hippocampus segmentation. *Hum Brain Mapp.* 2014;35(6):2674–97.
- 28.Li H, Habes M, Wolk DA, Fan Y, Alzheimer’s Disease Neuroimaging I, the Australian Imaging B, et al. A deep learning model for early prediction of Alzheimer’s disease dementia based on hippocampal magnetic resonance imaging data. *Alzheimers Dement.* 2019.
- 29.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.
- 30.Li J, Jin D, Li A, Liu B, Song C, Wang P, et al. ASAF: altered spontaneous activity fingerprinting in Alzheimer’s disease based on multisite fMRI. *Science Bulletin.* 2019;64(14):998–1010.
- 31.Liptak T. On the combination of independent tests. *Magyar Tud Akad Mat Kutato Int Kozl.* 1958;3:171–97.
- 32.Li T, Wang Q, Zhang J, Rolls ET, Yang W, Palaniyappan L, et al. Brain-Wide Analysis of Functional Connectivity in First-Episode and Chronic Stages of Schizophrenia. *Schizophr Bull.* 2017;43(2):436–48.
- 33.Zhang J, Cheng W, Liu Z, Zhang K, Lei X, Yao Y, et al. Neural, electrophysiological and anatomical basis of brain-network variability and its characteristic changes in mental disorders. *Brain.* 2016;139(Pt 8):2307–21.
- 34.Abraham A, Milham MP, Di Martino A, Craddock RC, Samaras D, Thirion B, et al. Deriving reproducible biomarkers from multi-site resting-state data: An Autism-based example. *Neuroimage.* 2017;147:736–45.
- 35.Rozycki M, Satterthwaite TD, Koutsouleris N, Erus G, Doshi J, Wolf DH, et al. Multisite Machine Learning Analysis Provides a Robust Structural Imaging Signature of Schizophrenia Detectable Across Diverse Patient Populations and Within Individuals. *Schizophr Bull.* 2018;44(5):1035–44.
- 36.Bennett DA, Schneider JA, Wilson RS, Bienias JL, Arnold SE. Neurofibrillary tangles mediate the association of amyloid load with clinical Alzheimer disease and level of cognitive function. *Arch Neurol.* 2004;61(3):378–84.
- 37.Schneider JA, Arvanitakis Z, Bang W, Bennett DA. Mixed brain pathologies account for most dementia cases in community-dwelling older persons. *Neurology.* 2007;69(24):2197–204.
- 38.Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007;81(3):559–75.
- 39.Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer’s disease: recommendations from the National Institute on Aging-

Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7(3):270–9.

40.Catani M, Dell'acqua F, Thiebaut de Schotten M. A revised limbic system model for memory, emotion and behaviour. *Neurosci Biobehav Rev*. 2013;37(8):1724–37.

41.Achterberg HC, van der Lijn F, den Heijer T, Vernooij MW, Ikram MA, Niessen WJ, et al. Hippocampal shape is predictive for the development of dementia in a normal, elderly population. *Hum Brain Mapp*. 2014;35(5):2359–71.

42.Gerardin E, Chetelat G, Chupin M, Cuingnet R, Desgranges B, Kim HS, et al. Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging. *Neuroimage*. 2009;47(4):1476–86.

43.Shen KK, Fripp J, Meriaudeau F, Chetelat G, Salvado O, Bourgeat P, et al. Detecting global and local hippocampal shape changes in Alzheimer's disease using statistical shape models. *Neuroimage*. 2012;59(3):2155–66.

44.Adler DH, Wisse LEM, Ittyerah R, Pluta JB, Ding SL, Xie L, et al. Characterizing the human hippocampus in aging and Alzheimer's disease using a computational atlas derived from ex vivo MRI and histology. *Proc Natl Acad Sci U S A*. 2018;115(16):4252–7.

45.Iglesias JE, Augustinack JC, Nguyen K, Player CM, Player A, Wright M, et al. A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: Application to adaptive segmentation of in vivo MRI. *Neuroimage*. 2015;115:117–37.

46.Csernansky JG, Wang L, Swank J, Miller JP, Gado M, McKeel D, et al. Preclinical detection of Alzheimer's disease: hippocampal shape and volume predict dementia onset in the elderly. *Neuroimage*. 2005;25(3):783–92.

47.Manning EN, Macdonald KE, Leung KK, Young J, Pepple T, Lehmann M, et al. Differential hippocampal shapes in posterior cortical atrophy patients: A comparison with control and typical AD subjects. *Hum Brain Mapp*. 2015;36(12):5123–36.

48.Hwang EJ, Kim HG, Kim D, Rhee HY, Ryu CW, Liu T, et al. Texture analyses of quantitative susceptibility maps to differentiate Alzheimer's disease from cognitive normal and mild cognitive impairment. *Med Phys*. 2016;43(8):4718.

49.Seshadri S, Fitzpatrick AL, Ikram MA, DeStefano AL, Gudnason V, Boada M, et al. Genome-wide analysis of genetic loci associated with Alzheimer disease. *JAMA*. 2010;303(18):1832–40.

50.Medland SE, Jahanshad N, Neale BM, Thompson PM. Whole-genome analyses of whole-brain data: working within an expanded search space. *Nat Neurosci*. 2014;17(6):791–800.

51. Logue MW, Panizzon MS, Elman JA, Gillespie NA, Hatton SN, Gustavson DE, et al. Use of an Alzheimer's disease polygenic risk score to identify mild cognitive impairment in adults in their 50s. *Mol Psychiatry*. 2019;24(3):421–30.
52. Foley SF, Tansey KE, Caseras X, Lancaster T, Bracht T, Parker G, et al. Multimodal Brain Imaging Reveals Structural Differences in Alzheimer's Disease Polygenic Risk Carriers: A Study in Healthy Young Adults. *Biol Psychiatry*. 2017;81(2):154–61.
53. Li J, Zhang X, Li A, Liu S, Qin W, Yu C, et al. Polygenic risk for Alzheimer's disease influences precuneal volume in two independent general populations. *Neurobiol Aging*. 2018;64:116–22.
54. Qiang W, Yau WM, Lu JX, Collinge J, Tycko R. Structural variation in amyloid-beta fibrils from Alzheimer's disease clinical subtypes. *Nature*. 2017;541(7636):217–21.
55. Ding Y, Sohn JH, Kawczynski MG, Trivedi H, Harnish R, Jenkins NW, et al. A Deep Learning Model to Predict a Diagnosis of Alzheimer Disease by Using (18)F-FDG PET of the Brain. *Radiology*. 2019;290(2):456–64.
56. Noble S, Scheinost D, Finn ES, Shen X, Papademetris X, McEwen SC, et al. Multisite reliability of MR-based functional connectivity. *Neuroimage*. 2017;146:959–70.
57. Eskildsen SF, Coupe P, Garcia-Lorenzo D, Fonov V, Pruessner JC, Collins DL, et al. Prediction of Alzheimer's disease in subjects with mild cognitive impairment from the ADNI cohort using patterns of cortical thinning. *Neuroimage*. 2013;65:511–21.
58. Meda SA, Narayanan B, Liu J, Perrone-Bizzozero NI, Stevens MC, Calhoun VD, et al. A large scale multivariate parallel ICA method reveals novel imaging-genetic relationships for Alzheimer's disease in the ADNI cohort. *Neuroimage*. 2012;60(3):1608–21.
59. 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(12):749–62.
60. Zhou HH, Singh V, Johnson SC, Wahba G, Alzheimer's Disease Neuroimaging I. Statistical tests and identifiability conditions for pooling and analyzing multisite datasets. *Proc Natl Acad Sci U S A*. 2018;115(7):1481–6.
61. van der Burgh HK, Schmidt R, Westeneng HJ, de Reus MA, van den Berg LH, van den Heuvel MP. Deep learning predictions of survival based on MRI in amyotrophic lateral sclerosis. *Neuroimage Clin*. 2017;13:361–9.
62. Li Z, Duan H, Zhao K, Ding Y. Stability of MRI Radiomics Features of Hippocampus: An Integrated Analysis of Test-Retest and Inter-Observer Variability. *IEEE Access*. 2019;7:97106–16.

- ## Tables

	Groups	N	Age	p	Gender (M/F)	p	MMSE	P
PL_S	NC	45	68.2±6.9		22/23		28.6±1.4	
	MCI	33	70.6±8.2	0.10	14/19	0.82	26.6±2.6	<0.001
	AD	37	71.8±8.4		16/21		17.5±5.7	
PL_G	NC	32	70.0±7.0		20/12		28.8±1.1	
	MCI	28	75.1±8.3	0.06	15/13	0.06	27.0±1.8	<0.001
	AD	40	71.6±9.4		14/26		19.3±4.5	
HH_Z	NC	24	65.5±6.2		9/15		28.8±1.2	
	MCI	33	65.4±8.3	0.53	10/23	0.24	25.9±2.5	<0.001
	AD	36	67.3±8.2		18/18		15.8±5.6	
QL_W	NC	42	65.5±6.8		12/30		28.4±1.8	
	MCI	16	66.1±7.4	0.20	8/8/	0.24	24.8±3.5	<0.001
	AD	63	68.0±7.1		26/37		19.0±3.5	
XW_H	NC	66	66.5±6.8		26/40		28.2±2.2	
	MCI	93	67.8±10.0	0.30	47/46	0.13	24.2±3.6	<0.001
	AD	47	69.1±8.5		16/31		16.7±6.4	
XW_Z	NC	22	65.5±8.2		8/14		28.3±1.5	
	MCI	20	70.7±8.5	0.06	11/9	0.48	21.3±5.5	<0.001
	AD	38	65.6±8.0		17/19		9.1±6.6	
ADNI	NC	356	73.9±5.9		171/185		29.1±1.2	
	MCI	550	73.3±7.5	0.01	324/226	0.004	27.6±1.8	<0.001
	AD	322	74.8±7.4		183/139		23.9±2.4	

A chi-squared test was used for the gender comparisons. Analysis of variance (ANOVA) was used for the age and MMSE comparisons. MMSE=Mini-Mental State Examination.

Abbreviations: Xuanwu Hospital Capital Medical University (XW_H and XW_Z), Chinese PLA General Hospital (PL_S and PL_G), Qilu Hospital of Shandong University (QL_W) and Tianjin Huanhu Hospital (HH_Z)

Table 2. Accuracy, sensitivity and specificity of the cross-validations.

Methods			ACC	SEN	SPE	AUC
CV1	In-house leave-center-out cross-validation		0.88	0.88	0.88	0.95
CV2	In-house 90% training, 10% testing for 1000 times		0.88±0.04	0.90±0.06	0.87±0.05	0.95±0.02
Training set		Testing set				
CV3	In-house	ADNI	0.79	0.87	0.72	0.89
CV4	ADNI	In-house	0.84	0.76	0.93	0.92

For the in-house data, N=715; for the ADNI data, N=1228. CV=cross-validation (details can be found in supplementary material S02)

Supplementary Files

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

- [Supplementary20191106final.docx](#)
- [Methodsformula.docx](#)