Science Bulletin 65 (2020) 1103–1113

Contents lists available at ScienceDirect

Science Bulletin

journal homepage: www.elsevier.com/locate/scib

Science Bulletin
www.scibull.com

Article

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

Kun Zhao [a,b,c], Yanhui Ding [c], Ying Han [d,q,r,s], Yong Fan [e], Aaron F. Alexander-Bloch [f], Tong Han [g], Dan Jin [a,h], Bing Liu [a,h,i], Jie Lu [j], Chengyuan Song [k], Pan Wang [l,m], Dawei Wang [n], Qing Wang [n], Kaibin Xu [a], Hongwei Yang [j], Hongxiang Yao [o], Yuanjie Zheng [c], Chunshui Yu [p], Bo Zhou [m], Xinqing Zhang [d], Yuying Zhou [l], Tianzi Jiang [a,h,i], Xi Zhang [m,*], Yong Liu [a,h,i,*], Alzheimer's Disease Neuroimaging Initiative [1], Multi-Center Alzheimer Disease Imaging Consortium

[a] Brainnetome Center & National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[b] School of Biological Science and Medical Engineering, Beihang University, Beijing 100191, China
[c] School of Information Science and Engineering, Shandong Normal University, Ji'nan 250358, China
[d] Department of Neurology, Xuanwu Hospital of Capital Medical University, Beijing 100053, China
[e] Department of Radiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA
[f] Department of Psychiatry, Yale University School of Medicine, New Haven, CT 06511, USA
[g] Department of Radiology, Tianjin Huanhu Hospital, Tianjin 300350, China
[h] School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China
[i] Center for Excellence in Brain Science and Intelligence Technology, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[j] Department of Radiology, Xuanwu Hospital of Capital Medical University, Beijing 100053, China
[k] Department of Neurology, Qilu Hospital of Shandong University, Ji'nan 250012, China
[l] Department of Neurology, Tianjin Huanhu Hospital, Tianjin University, Tianjin 300350, China
[m] Department of Neurology, The Secondary Medical Center, National Clinical Research Center for Geriatric Disease, Chinese PLA General Hospital, Beijing 100853, China
[n] Department of Radiology, Qilu Hospital of Shandong University, Ji'nan 250012, China
[o] Department of Radiology, The Secondary Medical Center, National Clinical Research Center for Geriatric Disease, Chinese PLA General Hospital, Beijing 100853, China
[p] Department of Radiology, Tianjin Medical University General Hospital, Tianjin 300052, China
[q] Center of Alzheimer's Disease, Beijing Institute for Brain Disorders, Beijing 100069, China
[r] Beijing Institute of Geriatrics, Beijing 100053, China
[s] National Clinical Research Center for Geriatric Disorders, Beijing 100053, China

ARTICLE INFO

Article history:
Received 6 December 2019
Received in revised form 31 January 2020
Accepted 17 March 2020
Available online 3 April 2020

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

ABSTRACT

Hippocampal morphological change is one of the main hallmarks of Alzheimer's disease (AD). However, whether hippocampal radiomic features are robust as predictors of progression from mild cognitive impairment (MCI) to AD dementia and whether these features provide any neurobiological foundation remains unclear. The primary aim of this study was to verify whether hippocampal radiomic features can serve as robust magnetic resonance imaging (MRI) markers for AD. Multivariate classifier-based support vector machine (SVM) analysis provided individual-level predictions for distinguishing AD patients (n = 261) from normal controls (NCs; n = 231) with an accuracy of 88.21% and intersite cross-validation. Further analyses of a large, independent the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset (n = 1228) reinforced these findings. In 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) Aβ, CSF Tau), and longitudinal changes in cognition ability; more importantly, the radiomic features had a consistently altered pattern with changes in the MMSE scores over 5 years of follow-up. These comprehensive results suggest that hippocampal radiomic features can serve as robust biomarkers for clinical application in AD/MCI, and further provide evidence for predicting whether an MCI subject would convert to AD based on the radiomics of the

* Corresponding authors.
E-mail addresses: zhangxi@301.hospital.com.cn (X. Zhang), yliu@nlpr.ia.ac.cn (Y. Liu).

[1] Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of the ADNI and/or provided data but did not participate in the analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

https://doi.org/10.1016/j.scib.2020.04.003
2095-9273/

hippocampus. The results of this study are expected to have a substantial impact on the early diagnosis of AD/MCI.

1. Introduction

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 and is important not only within the scientific community but also for the general public [5–9].

Hippocampal atrophy, or shape change, is one of the main hallmarks of AD [10]. 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 [11]. The hippocampus undergoes microstructural changes before severe atrophy, and pathological changes such as neurofibrillary tangles (NFTs) and amyloid-β (Aβ) are not directly detectable at the current resolution of clinical MRI [12]. 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 [13–15]. 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 [16] and is a useful way to extract detailed information from brain images, increase the precision of diagnosis, and assess disease prognosis [12,15,17–20]. Generalization is very important for clinical research; robust biomarkers should be stable over different samples of subjects or populations of subjects [5,7]. Thus, the generalization of radiomics features as biomarkers for AD should be explored further.

The apolipoprotein E (APOE) gene is known as a major genetic risk factor for AD [21]. Polygenic risk scores (PGRSs), comprehensive indicators that combine multiple risk alleles, provide a quantitative measure of genetic disease risk [22]. Previous large-scale genome-wide association studies found an association between the PGRSs for AD and the structural indices of certain brain regions [23–25]. During the progression of AD, Aβ plaques are considered to occur in the early stage, while Tau accumulation is considered to be the main factor underlying later dysfunction [26,27]. However, it is difficult to assess these risk factors with MRI; furthermore, despite the large number of classification studies, the association between imaging biomarkers and these neurobiological measures is not yet clear [28,29]. A strong association between imaging biomarkers and clinical information (such as APOE, PGRSs, Aβ and Tau) would be of great importance in the understanding of AD pathology, which would be implying that imaging biomarkers play an important role in determining the rate of disease progression in AD.

The development of valid biomarkers is crucial for optimizing individualized care in clinical application. Thus, identifying reproducible and generalizable markers is essential for the research community. In keeping with this goal, 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 showing initial promise as potential markers of AD traits or prodromal disease states. For this purpose, we investigated the neurobiological basis of the identified hippocampal radiomic features by relating these features to other variables, including APOE, PGRS, CSF Aβ, 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 Aβ and Tau pathology. Table 1 lists the demographic and clinical information, and additional details about the subjects and the image acquisition 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 a revised segmentation method with multi-atlas based local label learning (LLL) (https://www.nitrc.org/projects/locallabel), an automatic segmentation method that includes the N3 bias correction for imaging inhomogeneity. The generalization performance of the segmentation method had been confirmed by previous studies [30–32]. For each side, we 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 [20,33] and are listed in the Supplementary material S02.

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 difference at each site and then performed mega-analytic tests to integrate the multisite results [34]. 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

Table 1
Demographic and genetic characteristics of the target sample.

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

A *chi*-squared test was used for the gender comparisons. ANOVA was used for the age and MMSE comparisons. MMSE, Mini-Mental State Examination; XW_H and XW_Z, Xuanwu Hospital of Capital Medical University; PL_S and PL_G, Chinese PLA General Hospital; QL_W, Qilu Hospital of Shandong University; HH_Z, Tianjin Huanhu Hospital.

the Liptak-Stouffer *z*-score method [35] to integrate the results [36,37]. 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:

$$z = \frac{\sum_{i=1}^{k} w_i z_i}{\sqrt{\sum_{i=1}^{k} w_i^2}}, \quad (1)$$

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 intrasite cross-validation methods [38]. We next tested the generalizability of the classification with radiomic features in the ADNI data. Two additional independent cross-validation steps were performed: (1) the in-house data were used as the training set, and the ADNI data were applied as the testing set (CV3), and (2) conversely, the ADNI data was used as the training set, and the in-house data was used as the testing set (CV4) (Table 2, Supplementary material S03–S05). To further explore the effectiveness of the radiomics features, we also evaluated the performance of the classification with the XGBoost (https://github.com/dmlc/xgboost/), LightGBM (https://lightgbm.readthedocs.io), and linear discriminant analysis (LDA) with principal component analysis (PCA) models for feature dimension reduction with the above four cross-validation loops.

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) [39]. 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, PGRSs, Aβ, 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 mega-analysis and classification analysis (Fig. 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 Aβ+ and Aβ− in the MCI subgroups (also between Tau+ and Tau− and between APOE ε4+ and APOE ε4−). Aβ and Tau occasionally coexist, which greatly increases the risk for AD [26,40,41]. Thus, we also evaluated the differences in the identified features in three MCI subgroups (Aβ+&Tau+, Aβ−&Tau+ or Aβ+&Tau−, and Aβ−&Tau−). To explore the relationship between Aβ/Tau and the radiomic features, Pearson's correlations between the radiomic

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

	Methods		Model	ACC	SEN	SPE	AUC
CV1	In-house intersite cross-validation		SVM	0.88	0.88	0.88	0.95
			XGBoost	0.87	0.86	0.87	0.93
			LightGBM	0.87	0.85	0.88	0.93
			PCA+LDA	0.74	0.84	0.64	0.79
CV2	10 subjects from each center were randomly selected as the testing set 1000 times		SVM	0.88 ± 0.04	0.90 ± 0.06	0.87 ± 0.05	0.95 ± 0.02
			XGBoost	0.88 ± 0.04	0.88 ± 0.06	0.87 ± 0.06	0.94 ± 0.03
			LightGBM	0.87 ± 0.04	0.87 ± 0.06	0.87 ± 0.07	0.94 ± 0.03
			PCA+LDA	0.72 ± 0.06	0.72 ± 0.08	0.71 ± 0.07	0.78 ± 0.06
	Training set	Testing set					
CV3	In-house	ADNI	SVM	0.79	0.87	0.72	0.89
			XGBoost	0.76	0.91	0.63	0.89
			LightGBM	0.77	0.89	0.66	0.89
			PCA+LDA	0.75	0.87	0.61	0.81
CV4	ADNI	In-house	SVM	0.84	0.76	0.93	0.92
			XGBoost	0.82	0.75	0.89	0.89
			LightGBM	0.82	0.74	0.90	0.91
			PCA+LDA	0.69	0.58	0.78	0.75

For the in-house data, $n = 715$; for the ADNI data, $n = 1228$. (Details can be found in Supplementary material S03).
CV, cross-validation; VM, Support Vector Machine; XGBoost, eXtreme Gradient Boosting; LightGBM, Light Gradient Boosting Machine.

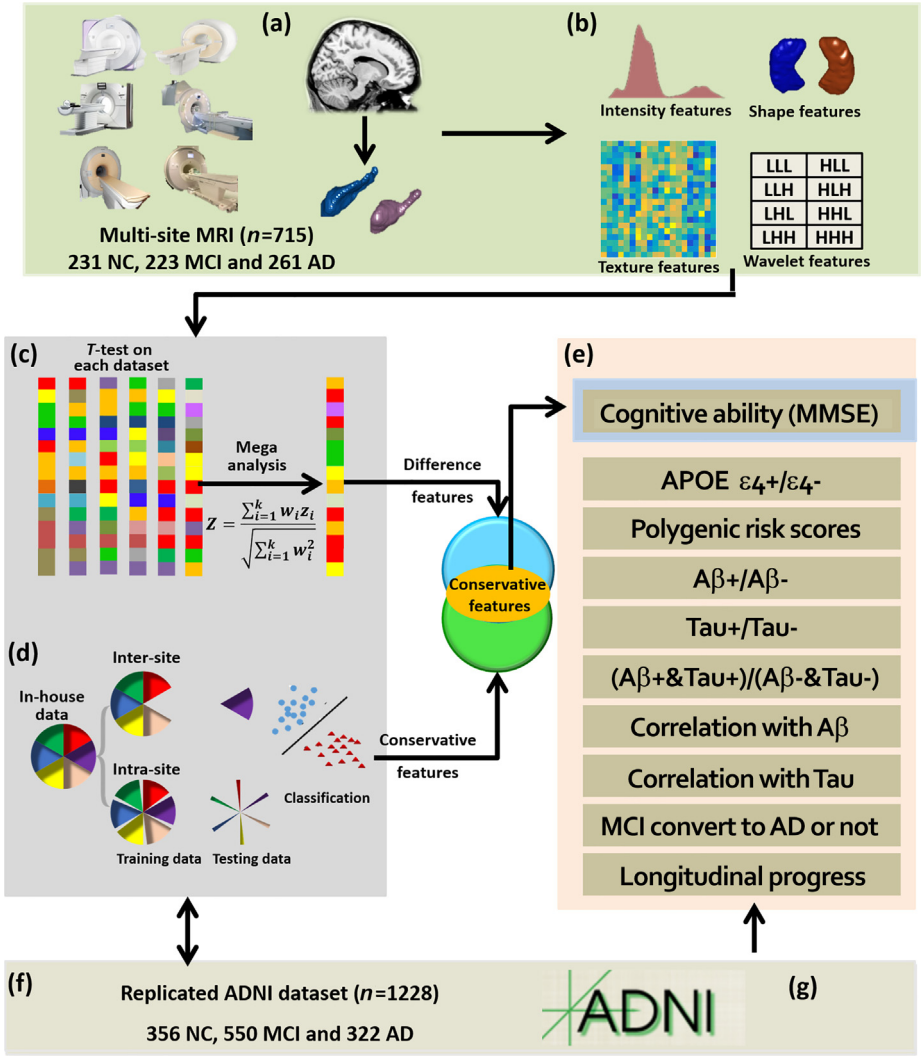

Fig. 1. (Color online) Schematic of the data analysis pipeline. (a) Automatic segmentation of the hippocampus in the individual space with the LLL method. (b) Computation of radiomic features. (c) *T*-test between NC and AD and mega-analysis to identify differences at the multicenter level. (d) Classification analysis between AD and NC using intersite and intrasite cross-validation to test the generalizability of the results. (e) Correlation between MMSE and radiomic features (and the decision value of the subject to the classification plane). (f) Validation of the ADNI dataset for using ADNI as training data with in-house data as testing data as well as using in-house as training data with ADNI data as testing data to test the generalizability of the results. (g) Relationship between radiomic features and clinical information (e.g., cognitive ability, APOE genetics, PGRSs, Aβ, Tau, and longitudinal changes in the cognitive abilities in MCI).

features and Aβ/Tau were also evaluated in the MCI groups (Supplementary material S06).

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 [42] and recent summary statistics for AD [23] 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 S07).

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 ± 24.57 mo nths) and stable MCI (SMCI; n = 134, interscan interval: 38.61 ± 29. 22 months) subjects (Supplementary material S08) 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 S08–S09).

2.6. 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.

3. Results

3.1. Demographic characteristics and neuropsychological assessment of the groups

In total, the information from 715 subjects, including 231 NCs, 261 CE patients and 223 MCI patients analyzed on 6 scanners, was studied as in-house data. The mean ages (ANOVA) and gender ratios between groups were not significantly different among centers (all $P > 0.05$). The cognitive ability scores (MMSE) were significantly different ($P < 0.001$), with AD patients having the lowest scores, the NC groups having the highest scores, and the MCI groups being in the middle (Table 1). Furthermore, a total of 1228 subjects from the ADNI database were included to validate and test the generalizability of the radiomics analysis. This dataset included 356 NCs (age: 73.90 ± 5.87; gender (M/F): 171/185; MMSE: 29.06 ± 1.16), 550 MCI patients (age: 73.25 ± 7.49; gender (M/F): 324/226; MMSE: 27.61 ± 1.78) and 322 CE patients (age: 74.78 ± 7.40; gender (M/F): 183/139; MMSE: 23.89 ± 2.43). Of these 1228 subjects, 708 (182 NCs, 328 MCI patients and 198 CE patients) had available genetic information and Aβ and Tau pathology. Detailed information can be found in the Supplementary material S06.

3.2. Hippocampal radiomic features differ 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 S10). 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) (Fig. 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 (Fig. 2d–f).

3.3. Classification performance

The intersite cross-validation showed an AUC = 0.95 (ACC = 0.88) with the default parameters (c = 2, g = 1/n, n: number of features) obtained by the SVM-RFE with 57 selected features (when selecting 50–250 features, ACC = 0.87 ± 0.01). 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 intrasite cross-validation, the mean AUC = 0.95 ± 0.02 (Fig. 3a, b, 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) (Fig. 3c, Table 2); when selecting 50–250 features, ACC = 0.78 ± 0.01. Conversely, when using the in-house data as the testing set, we obtained an AUC = 0.92 (ACC = 0.84) (Fig. 3d, Table 2); when selecting 50–250 features, ACC = 0.82 ± 0.005. Moreover, the classifier output in the test data was significantly negatively correlated with the MMSE score in the AD and MCI groups (Fig. 3e1, f1) and was also significantly negatively correlated with the MMSE score separately in the AD (red lines in Fig. 3e2, f2) and MCI groups (blue lines in Fig. 3e2, f2).

Furthermore, the intersite cross-validation yielded an AUC = 0.93 (ACC = 0.87) with the XGBoost model, an AUC = 0.93 (ACC = 0.87) with the LightGBM model, and an AUC = 0.79 (ACC = 0.74) with the LDA model (Table 2). The intrasite cross-validation yielded an AUC = 0.94 ± 0.03 (ACC = 0.88 ± 0.04) with the XGBoost model, an AUC = 0.94 ± 0.03 (ACC = 0.87 ± 0.04) with the LightGBM model, and AUC = 0.78 ± 0.06 (ACC = 0.72 ± 0.06) with the LDA model (Table 2). More importantly, when using the in-house data as the training data and the ADNI data as the testing data, we achieved an AUC = 0.89 (ACC = 0.76) with the XGBoost model, an AUC = 0.89 (ACC = 0.77) with the LightGBM model, and an AUC = 0.81 (ACC = 0.75) with the LDA model. Conversely, when using the in-house data as the testing set, we obtained an AUC = 0.89 (ACC = 0.82) with the XGBoost model, an AUC = 0.91 (ACC = 0.82) with the LightGBM model and an AUC = 0.75 (ACC = 0.69) with the LDA model.

3.4. Radiomic features and cognitive ability, APOE genotype, Aβ, Tau, and longitudinal cognitive changes in the MCI subjects

For the intersite 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 ($P < 0.05$) (Fig. 4a). More importantly, almost all of these results (>95%) were replicated in the ADNI datasets (Fig. 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 ε4+ subjects (n = 177) and the APOE ε4− subjects (n = 151) ($P < 0.05$) (Fig. 5, Supplementary material S06). Twenty-one of the 33 features were significantly associated with the PGRSs after accounting for the regression group and the influence of race (Fig. 5, Supplementary material S07). For Aβ, we

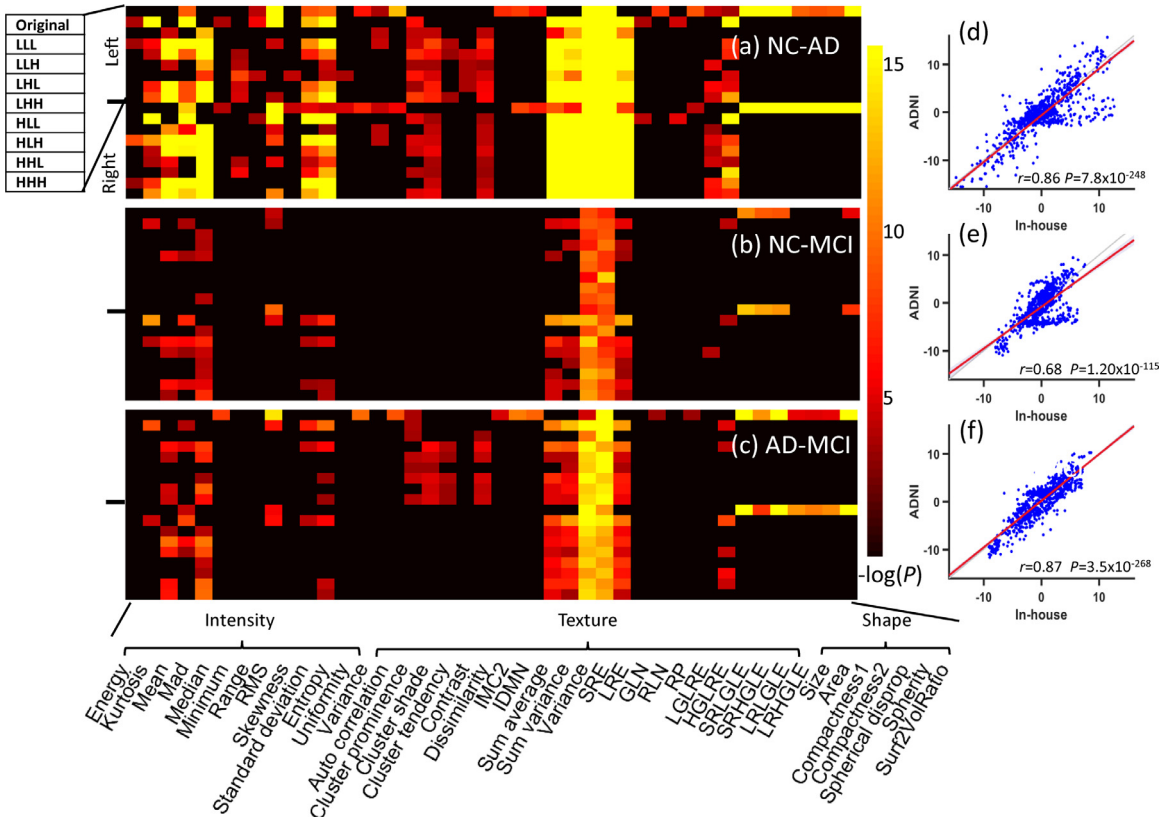

Fig. 2. (a) Group differences of the used radiomic features between AD and NC. (b) Group differences of the used radiomic features between NC and MCI. (c) Group differences of the used radiomic features between AD and MCI. For (a–c), each matrix has 18 rows, of which the first 9 rows show the left hippocampus, and the next 9 rows show the right hippocampus; for each hemisphere, the first row represents the features in the raw space, and the following 8 rows represent the results for different frequency domains of the wavelet transform (LLL, LLH, LHL, HLL, HHL, LHH, HLH, HHH). The color bar represents −log(*P*), and *P* is computed by a mega-analysis of the 6 sites. (d) Correlation between the in-house statistical scores and the ADNI statistical scores of the used radiomic features in NC and AD. (e) Correlation between the in-house statistical scores and the ADNI statistical scores of the used radiomic features in NC and MCI. (f) Correlation between the in-house statistical scores and the ADNI statistical scores of the used radiomic features in AD and MCI.

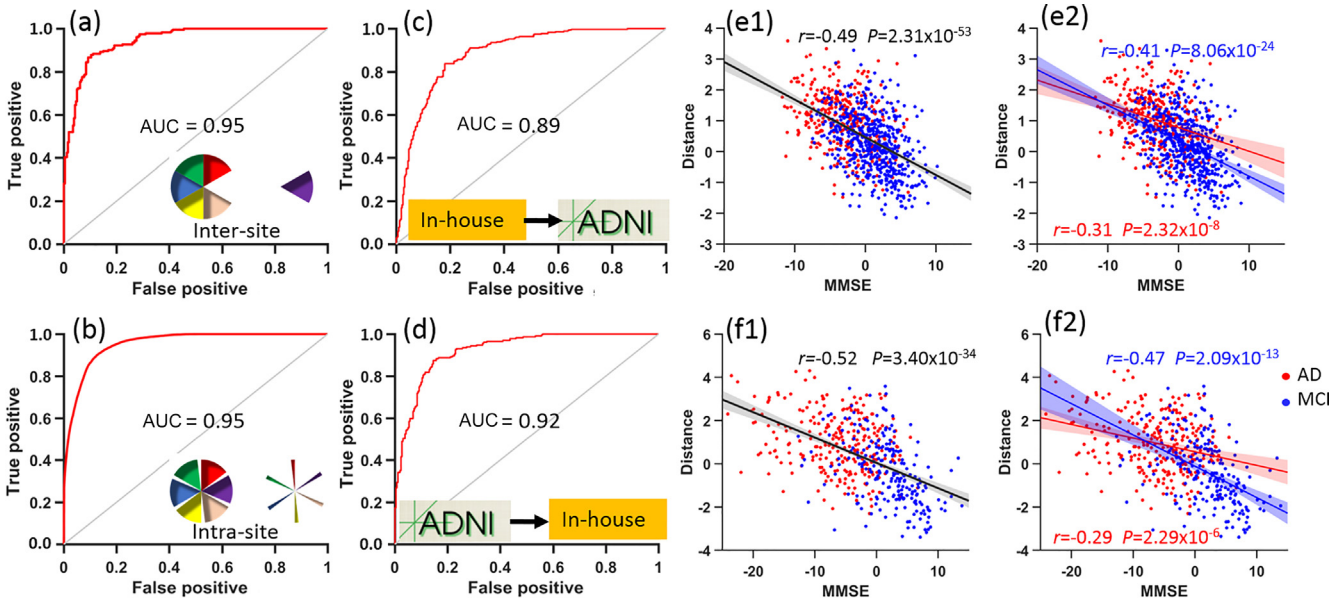

Fig. 3. (a) The ROC curve for classification between AD and NC with intersite cross-validation. (b) The ROC curve for classification between AD and NC intrasite cross-validation. (c) The ROC curve for classification using in-house data as training data and ADNI data as test data. (d) The ROC curve for classification using ADNI data as training data and in-house data as testing data. (e) Correlation of classifier outputs with the MMSE score of the ADNI data (e1, the MCI group and AD group; e2, only in the AD group or only in the MCI group. (f) Correlation of classifier outputs with the MMSE score for the in-house data (f1, MCI group and AD group; f2, AD group alone (red) or MCI group alone (blue)).

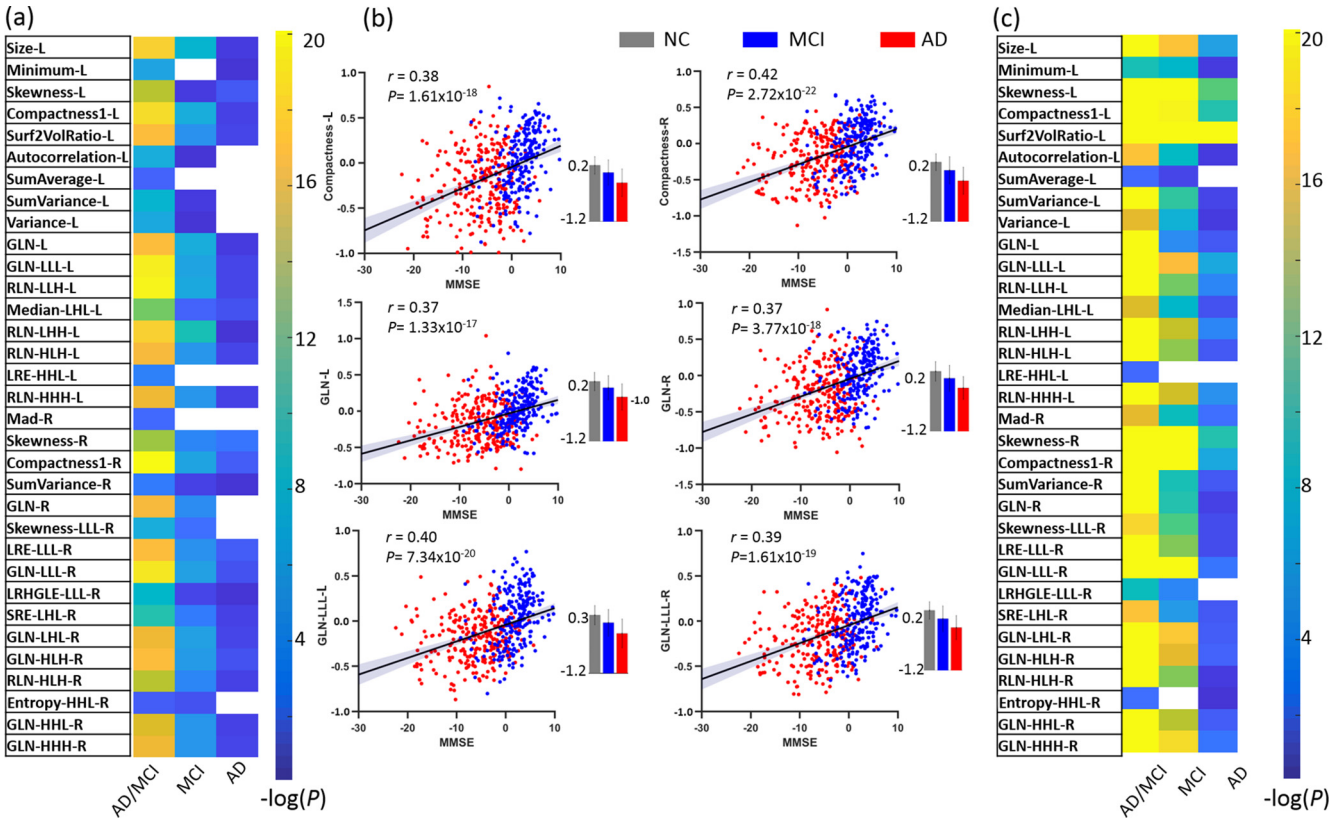

Fig. 4. (a) Correlations between the MMSE score and a subset of radiomic features. The values of the color bar reflect −log(*P*), and the 3 columns are the combined AD and MCI group, only the MCI group and only the AD group. (b) Scatterplots illustrate several significant correlations, and the small bar plots represent the mean (std) values of the NC (gray), MCI (blue) and AD (red) groups. (c) Correlation with the MMSE score and the radiomic features for the ADNI data. GLN: gray level nonuniformity, RLN: run length nonuniformity, LRE: long run emphasis, SRE: short run emphasis, Surf2VolRatio: ratio of surface area to volume, LRHGLE: long run high gray level emphasis. L: left hippocampus features, R: right hippocampus features. Additional information can be found in the Supplementary materials.

defined the Aβ+ group as Aβ < 675 (mean-0.5 × std, n = 86) and the Aβ− group as Aβ > 1017 (mean + 0.5 × std, n = 129). Nineteen of the 33 features were significantly different between the Aβ− and Aβ+ subjects, and 20 features showed significant correlations with the Aβ scores in the MCI subjects (n = 328, P < 0.05) (Fig. 5a, Supplementary material S06). 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) (Fig. 5a, Supplementary material S06). When comparing the subjects who were Aβ+&Tau+ (n = 43) to those who were Aβ−&Tau− (n = 52), 21 of the 33 features were significantly different. The subjects who were Aβ+/Tau− or Aβ−/Tau+ were intermediate between these two (Fig. 5b, Supplementary material S06).

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) (Fig. 5a, Supplementary material S08). 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 (Figs. 5c and Supplementary material S09). 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) (Figs. 5d and Supplementary material S08).

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, the enrichment analyses of the association between radiomic features and the genotype, Tau, Aβ and longitudinal cognitive variances highlight the solid neurobiological substrates underlying the progression of AD. This is of great significance for the early clinical diagnosis or prognostic follow-up in AD.

The development of valid biomarkers is crucial for optimizing individualized care in AD. The multisite MRI findings of altered hippocampal radiomic features in AD confirm and extend previous neuroimaging findings [12,15,20]. 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 [12], which may indicate atrophy or gray matter loss in the hippocampus [43,44]. 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 [17,45–47], and these phenomena indicate that atrophy of the hippocampus does not occur collaterally [48,49]. 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

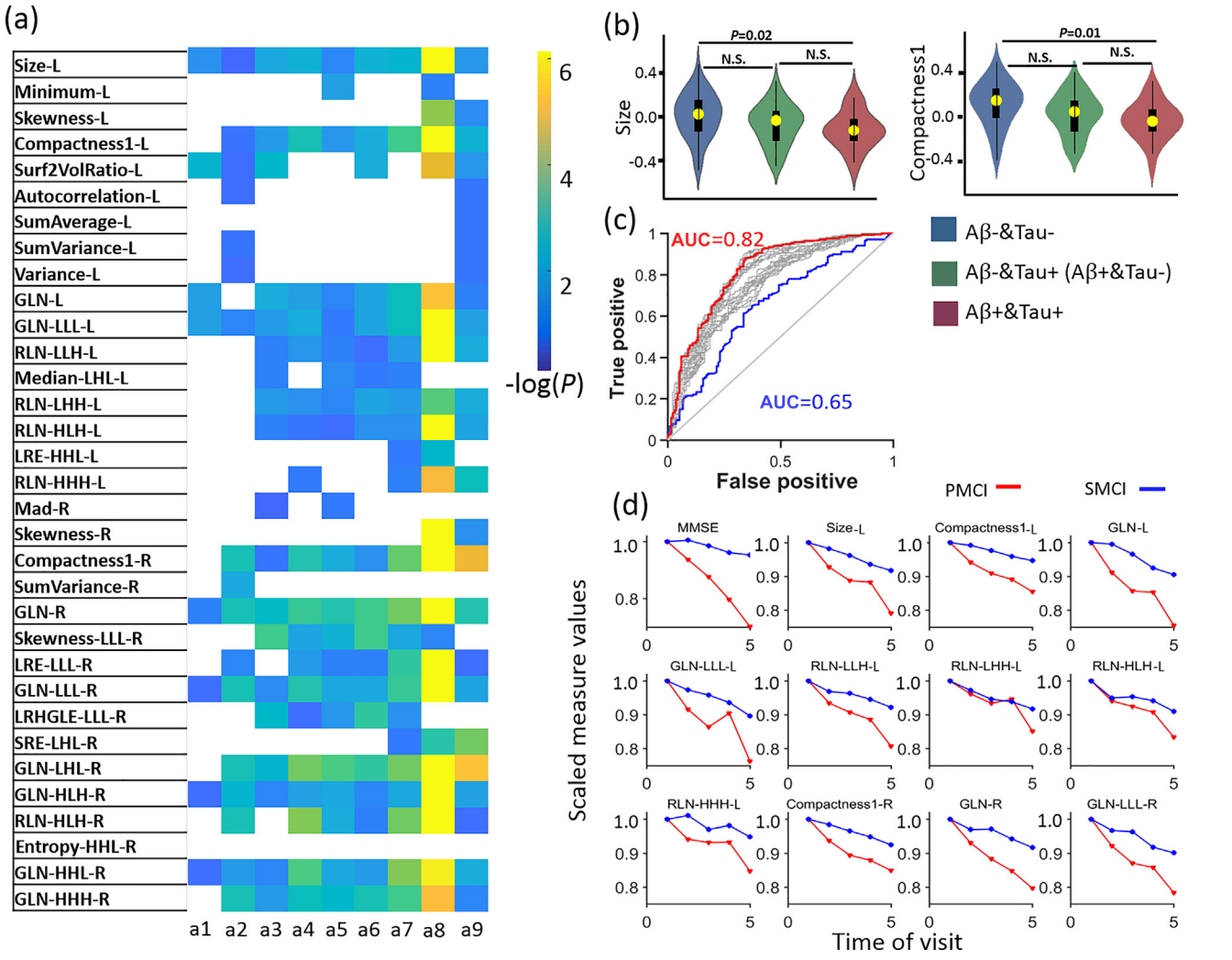

Fig. 5. (a) Negative logarithm of the *P*-value for either a *t*-test of the group difference or correlation with radiomic features for the following variables: a1, APOE ε4+ group vs. APOE ε4− group in the MCI subjects; a2, correlation with the polygenic risk scores; a3, high Aβ (Aβ−) vs. low Aβ(Aβ+) in the MCI subjects; a4, high Tau (Tau+) vs. low Tau (Tau−) in the MCI subjects; a5: (Aβ+&Tau+) vs. (Aβ−&Tau−); a6: correlation between the radiomic features and Aβ in the MCI subjects, a7: correlation between the radiomic features and Tau in the MCI subjects; a8: correlation between changes in cognitive ability and changes in features in the PMCI group; a9: correlation between changes in cognitive ability and changes in features in the SMCI group. (b) Examples of feature distributions in the subgroups Aβ+&Tau+, Aβ+&Tau− or Aβ−&Tau+ and Aβ−&Tau−. (c) The ROC curve for prediction of whether MCI conversion to AD increases from AUC = 0.65 (using age, gender and APOE status) to 0.80 (adding the subset of consistent radiomic features one by one until the AUC does not increase) using logistical regression (see Supplementary material S09 for detailed information). (d) Trend of MMSE changes and the changes in the identified features over 5 years in PMCI (red) and SMCI (blue). The *x*-axis indicates the visit time (1–5) and the *y*-axis indicates the scaled measure values ((score at visit time)/(score at baseline)). L: left hippocampus features, R: right hippocampus features, PMCI: progressive MCI, SMCI: stable MCI.

Aβ plaques, cannot be detected by MRI, these microstructural changes might lead to altered textural patterns detectable via texture analysis [12,50–52]. In support of this, the 27 identified features showed significant associations with the positron emission tomography (PET) amyloid value in the hippocampus (Supplementary material S11). A subset of radiomic features was also significantly associated with high-risk genetic status, Aβ 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" [14].

Imaging biomarkers are the cornerstone of modern radiology. 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 [53,54]. As expected, this study shows significant radiomic differences between APOE ε4+ and APOE ε4− 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 [24,25,55–57]. 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. Aβ plaques and Tau NFTs are the hallmark lesions of AD, and the strong association between radiomic features and the progression of Aβ and/or Tau in the MCI subjects indicates the strong neurobiological substrate of these features [2,12,17,58]. It is still a significant challenge to evaluate and predict the progress of MCI [32,59]. 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.

It is very important for the research community to elucidate reproducible and replicative biomarkers for various diseases. Translation MRI studies aimed at identifying robust disease biomarkers require large numbers of samples, which can be difficult to obtain from a single site [60], 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 [5,7]. It should be noted that the ADNI cohort, which has a large sample (more than 1500), was collected from more than 50 centers [61,62], and the mean sample size is relatively small (Supplementary material S01). 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 [4,38]. 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 [39]. 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 [63].

5. 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. 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 [64]. 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 [65]. 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 [30]. It should be noted that different subregions (for example, the dentate gyrus, cornu ammonis and subiculum) of the hippocampus are associated with distinct functions; the present radiomics protocol (especially with 3D wavelet transformation) might not work well for these relatively small regions. Therefore, new measures and the accurate and reliable automation segmentation of these subfields with high resolution imaging will also facilitate our research in the future. Finally, the use of radiomic features of the hippocampus alone are not enough to fully describe AD progression; combining markers from other brain regions will help identify robust and reproducible biomarkers for clinical applications [4,7,66–68].

6. Conclusion

This work represents a large and interdisciplinary effort to develop and validate AD neuroimaging biomarkers. Utilizing large multisite neuroimaging datasets and radiomics profiles, the present study developed novel biomarkers that, to the best of our knowledge, are the first personalized, reproducible and scientifically interpretable biomarkers for AD. 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.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

This work was partially supported by the National Key Research and Development Program of China (2016YFC1305904), the National Natural Science Foundation of China (81871438, 81901101, 61633018, 81571062, 81400890, 81871398), the Strategic Priority Research Program (B) of the Chinese Academy of Sciences (XDB32020200), the Beijing Municipal Science & Technology Commission (Z171100000117001, Z171100000117002), the Primary Research & Development Plan of Shandong Province (2017GGX10112), and the Open Project Program of the National Laboratory of Pattern Recognition (NLPR) (201900021).

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.

The detailed descriptions of radiomics features was defined by Hugo J.W.L. Aerts and colleagues.

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

Author contributions

Kun Zhao, Yanhui Ding, Yong Fan and Yong Liu analyzed the data and performed the measurements; Ying Han, Tong Han, Jie Lu, Chengyuan Song, Pan Wang, Dawei Wang, Qing Wang, Hongwei Yang, Hongxiang Yao, Chunshui Yu, Bo Zhou, Xinqing Zhang, Yuying Zhou, and Xi Zhang collected the data; Kun Zhao and Yong Liu were principally responsible for preparing the manuscript; Aaron F. Alexander-Bloch, Dan Jin, Bing Liu, Kaibin Xu, Yuanjie Zheng, Chunshui Yu and Tianzi Jiang revised the manuscript; Yong Liu supervised the project.

Appendix A. Supplementary materials

Supplementary materials to this article can be found online at https://doi.org/10.1016/j.scib.2020.04.003.

References

[1] Hampel H, Burger K, Teipel SJ, et al. Core candidate neurochemical and imaging biomarkers of Alzheimer's disease. Alzheimers Dement 2008;4:38–48.
[2] Nakamura A, Kaneko N, Villemagne VL, et al. High performance plasma amyloid-beta biomarkers for Alzheimer's disease. Nature 2018;554:249–54.
[3] Caldwell JZK, Berg JL, Cummings JL, et al. Moderating effects of sex on the impact of diagnosis and amyloid positivity on verbal memory and hippocampal volume. Alzheimers Res Ther 2017;9:72.
[4] Rathore S, Habes M, Iftikhar MA, et al. 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] Varoquaux G. Cross-validation failure: small sample sizes lead to large error bars. Neuroimage 2018;180:68–77.
[6] Teipel SJ, Wohlert A, Metzger C, et al. Multicenter stability of resting state fMRI in the detection of Alzheimer's disease and amnestic MCI. Neuroimage Clin 2017;14:183–94.
[7] Woo CW, Chang LJ, Lindquist MA, et al. Building better biomarkers: brain models in translational neuroimaging. Nat Neurosci 2017;20:365–77.
[8] Poldrack RA. The costs of reproducibility. Neuron 2019;101:11–4.
[9] Xing XX, Zuo XN. The anatomy of reliability: a must read for future human brain mapping. Sci Bull 2018;63:1606–7.
[10] Shi F, Liu B, Zhou Y, et al. Hippocampal volume and asymmetry in mild cognitive impairment and Alzheimer's disease: meta-analyses of MRI studies. Hippocampus 2009;19:1055–64.
[11] Young AL, Marinescu RV, Oxtoby NP, et al. Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with subtype and stage inference. Nat Commun 2018;9:4273.
[12] Sorensen 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.
[13] Parmar C, Leijenaar RT, Grossmann P, et al. Radiomic feature clusters and prognostic signatures specific for lung and head & neck cancer. Sci Rep 2015;5:11044.
[14] Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology 2016;278:563–77.
[15] Chaddad A, Desrosiers C, Niazi T. Deep radiomic analysis of MRI related to Alzheimer's disease. IEEE Access 2018;6:58213–21.
[16] Kassner A, Thornhill RE. Texture analysis: a review of neurologic MR imaging applications. AJNR Am J Neuroradiol 2010;31:809–16.
[17] Sorensen L, Igel C, Pai A, 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.
[18] Chincarini A, Bosco P, Calvini P, et al. Local MRI analysis approach in the diagnosis of early and prodromal Alzheimer's disease. Neuroimage 2011;58:469–80.
[19] 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.
[20] 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.
[21] Lambert JC, Ibrahim-Verbaas CA, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet 2013;45:1452–8.
[22] Torkamani A, Wineinger NE, Topol EJ. The personal and clinical utility of polygenic risk scores. Nat Rev Genet 2018;19:581–90.
[23] Jansen IE, Savage JE, Watanabe K, et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. Nat Genet 2019;51:404–13.
[24] Mormino EC, Sperling RA, Holmes AJ, et al. Polygenic risk of Alzheimer disease is associated with early- and late-life processes. Neurology 2016;87:481–8.
[25] Axelrud LK, Santoro ML, Pine DS, et al. Polygenic risk score for Alzheimer's disease: implications for memory performance and hippocampal volumes in early life. Am J Psychiatry 2018;175:555–63.
[26] Sepulcre J, Grothe MJ, d'Oleire Uquillas F, et al. Neurogenetic contributions to amyloid beta and tau spreading in the human cortex. Nat Med 2018;24:1910–8.
[27] Kim HJ, Park S, Cho 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;75:999–1007.
[28] Lian C, Liu M, Zhang J, et al. Hierarchical fully convolutional network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI. IEEE Trans Pattern Anal Mach Intell 2020;42:880–93.
[29] Davatzikos C. Machine learning in neuroimaging: progress and challenges. Neuroimage 2019;197:652–6.
[30] Li ZR, Duan HC, Zhao K, et al. Stability of MRI radiomics features of hippocampus: an integrated analysis of test-retest and inter-observer variability. IEEE Access 2019;7:97106–16.
[31] Hao Y, Wang T, Zhang X, et al. Local label learning (LLL) for subcortical structure segmentation: application to hippocampus segmentation. Hum Brain Mapp 2014;35:2674–97.
[32] Li H, Habes M, Wolk DA, et al. A deep learning model for early prediction of Alzheimer's disease dementia based on hippocampal magnetic resonance imaging data. Alzheimers Dement 2019;15:1059–70.
[33] Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5:4006.
[34] Li J, Jin D, Li A, et al. ASAF: altered spontaneous activity fingerprinting in Alzheimer's disease based on multisite fMRI. Sci Bull 2019;64:998–1010.
[35] Liptak T. On the combination of independent tests. Magyar Tud Akad Mat Kutato Int Kozl 1958;3:171–97.
[36] Li T, Wang Q, Zhang J, et al. Brain-wide analysis of functional connectivity in first-episode and chronic stages of schizophrenia. Schizophr Bull 2017;43:436–48.
[37] Zhang J, Cheng W, Liu Z, et al. Neural, electrophysiological and anatomical basis of brain-network variability and its characteristic changes in mental disorders. Brain 2016;139:2307–21.
[38] Abraham A, Milham MP, Di Martino A, et al. Deriving reproducible biomarkers from multi-site resting-state data: an Autism-based example. Neuroimage 2017;147:736–45.
[39] Rozycki M, Satterthwaite TD, Koutsouleris N, 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;44:1035–44.
[40] Bennett DA, Schneider JA, Wilson RS, et al. Neurofibrillary tangles mediate the association of amyloid load with clinical Alzheimer disease and level of cognitive function. Arch Neurol 2004;61:378–84.
[41] Schneider JA, Arvanitakis Z, Bang W, et al. Mixed brain pathologies account for most dementia cases in community-dwelling older persons. Neurology 2007;69:2197–204.
[42] Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–75.
[43] Albert MS, DeKosky ST, Dickson D, 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;7:270–9.
[44] Catani M, Dell'acqua F, Thiebaut de Schotten M. A revised limbic system model for memory, emotion and behaviour. Neurosci Biobehav Rev 2013;37:1724–37.
[45] Achterberg HC, van der Lijn F, den Heijer T, et al. Hippocampal shape is predictive for the development of dementia in a normal, elderly population. Hum Brain Mapp 2014;35:2359–71.
[46] Gerardin E, Chetelat G, Chupin M, et al. Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging. Neuroimage 2009;47:1476–86.
[47] Shen KK, Fripp J, Meriaudeau F, et al. Detecting global and local hippocampal shape changes in Alzheimer's disease using statistical shape models. Neuroimage 2012;59:2155–66.
[48] Adler DH, Wisse LEM, Ittyerah R, 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 USA 2018;115:4252–7.
[49] Iglesias JE, Augustinack JC, Nguyen K, 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.
[50] Csernansky JG, Wang L, Swank J, et al. Preclinical detection of Alzheimer's disease: hippocampal shape and volume predict dementia onset in the elderly. Neuroimage 2005;25:783–92.
[51] Manning EN, Macdonald KE, Leung KK, et al. Differential hippocampal shapes in posterior cortical atrophy patients: a comparison with control and typical AD subjects. Hum Brain Mapp 2015;36:5123–36.
[52] Hwang EJ, Kim HG, Kim D, et al. Texture analyses of quantitative susceptibility maps to differentiate Alzheimer's disease from cognitive normal and mild cognitive impairment. Med Phys 2016;43:4718.
[53] Seshadri S, Fitzpatrick AL, Ikram MA, et al. Genome-wide analysis of genetic loci associated with Alzheimer disease. JAMA 2010;303:1832–40.
[54] Medland SE, Jahanshad N, Neale BM, et al. Whole-genome analyses of whole-brain data: working within an expanded search space. Nat Neurosci 2014;17:791–800.
[55] Logue MW, Panizzon MS, Elman JA, et al. Use of an Alzheimer's disease polygenic risk score to identify mild cognitive impairment in adults in their 50s. Mol Psychiatry 2019;24:421–30.
[56] Foley SF, Tansey KE, Caseras X, et al. Multimodal brain imaging reveals structural differences in Alzheimer's disease polygenic risk carriers: a study in healthy young adults. Biol Psychiatry 2017;81:154–61.
[57] Li J, Zhang X, Li A, et al. Polygenic risk for Alzheimer's disease influences precuneal volume in two independent general populations. Neurobiol Aging 2018;64:116–22.
[58] Qiang W, Yau WM, Lu JX, et al. Structural variation in amyloid-beta fibrils from Alzheimer's disease clinical subtypes. Nature 2017;541:217–21.
[59] Ding Y, Sohn JH, Kawczynski MG, et al. A deep learning model to predict a diagnosis of Alzheimer disease by using (18)F-FDG PET of the brain. Radiology 2019;290:456–64.

[60] Noble S, Scheinost D, Finn ES, et al. Multisite reliability of MR-based functional connectivity. Neuroimage 2017;146:959–70.
[61] Eskildsen SF, Coupe P, Garcia-Lorenzo D, 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.
[62] Meda SA, Narayanan B, Liu J, et al. A large scale multivariate parallel ICA method reveals novel imaging-genetic relationships for Alzheimer's disease in the ADNI cohort. Neuroimage 2012;60:1608–21.
[63] Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017;14:749–62.
[64] Zhou HH, Singh V, Johnson SC, et al. Statistical tests and identifiability conditions for pooling and analyzing multisite datasets. Proc Natl Acad Sci USA 2018;115:1481–6.
[65] van der Burgh HK, Schmidt R, Westeneng HJ, et al. Deep learning predictions of survival based on MRI in amyotrophic lateral sclerosis. Neuroimage Clin 2017;13:361–9.
[66] Jack Jr CR, Barkhof F, Bernstein MA, et al. Steps to standardization and validation of hippocampal volumetry as a biomarker in clinical trials and diagnostic criterion for Alzheimer's disease. Alzheimers Dement 2011;7:e4.
[67] Dubois B, Feldman HH, Jacova C, et al. Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. Lancet Neurol 2014;13:614–29.
[68] Schott JM, Crutch SJ, Carrasquillo MM, et al. Genetic risk factors for the posterior cortical atrophy variant of Alzheimer's disease. Alzheimers Dement 2016;12:862–71.

Kun Zhao is currently a Ph.D. student at the School of Biological Science and Medical Engineering, Beihang University. His research mainly focuses on studying the Alzheimer's disease by structural magnetic resonance imaging.

Xi Zhang is a professor at the Department of Neurology, the Secondary Medical Center, National Clinical Research Center for Geriatric Disease, Chinese PLA General Hospital. His research aims at prevention, diagnosis and intervention research of Alzheimer's disease, especially the early diagnosis of Alzheimer's disease by multi-modal neuroimaging techniques.

Yong Liu is a professor in Brainnetome Center, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences (CASIA). He obtained his Ph.D. degree from CASIA in 2008. Since then, he joined CASIA as an assistant/associate research professor. He was a visiting scholar from April 2011 to March 2012 in Brain Mapping Unit in University of Cambridge. His present research is focused on studying the early diagnosis of Alzheimer's disease with multi-center multimodal brain imaging.