A neuroimaging biomarker for Individual Brain-Related Abnormalities In Neurodegeneration (IBRAIN): a cross-sectional study

Kun Zhao,[a] Pindong Chen,[b] Aaron Alexander-Bloch,[c,d] Yongbin Wei,[a] Martin Dyrba,[e] Fan Yang,[f] Xiaopeng Kang,[b] Dawei Wang,[g] Dongsheng Fan,[h,i,j] Shan Ye,[h,i,j] Yi Tang,[k] Hongxiang Yao,[l] Bo Zhou,[m] Jie Lu,[n] Chunshui Yu,[o] Pan Wang,[p] Zhengluan Liao,[q] Yan Chen,[q] Longjian Huang,[r] Xi Zhang,[m] Ying Han,[k,s,t] Shuyu Li,[u,**] and Yong Liu[a,b,*]

[a]School of Artificial Intelligence, Beijing University of Posts and Telecommunications, Beijing, China
[b]School of Artificial Intelligence, University of Chinese Academy of Sciences & Brainnetome Centre, Chinese Academy of Sciences, Beijing, China
[c]Department of Psychiatry, University of Pennsylvania, Philadelphia, USA
[d]Department of Child and Adolescent Psychiatry and Behavioral Science, Children's Hospital of Philadelphia, Philadelphia, USA
[e]German Centre for Neurodegenerative Diseases (DZNE), Rostock, Germany
[f]CAS Key Laboratory of Molecular Imaging, Institute of Automation, Beijing, China
[g]Department of Radiology, Qilu Hospital of Shandong University, Ji'nan, China
[h]Department of Neurology, Peking University Third Hospital, Beijing, China
[i]Key Laboratory for Neuroscience, National Health Commission/Ministry of Education, Peking University, Beijing, China
[j]Beijing Key Laboratory of Biomarker and Translational Research in Neurodegenerative Diseases, Beijing, China
[k]Department of Neurology, Xuanwu Hospital of Capital Medical University, Beijing, China
[l]Department of Radiology, The Second Medical Centre, National Clinical Research Centre for Geriatric Diseases, Chinese PLA General Hospital, Beijing, China
[m]Department of Neurology, The Second Medical Centre, National Clinical Research Centre for Geriatric Diseases, Chinese PLA General Hospital, Beijing, China
[n]Department of Radiology, Xuanwu Hospital of Capital Medical University, Beijing, China
[o]Department of Radiology, Tianjin Medical University General Hospital, Tianjin, China
[p]Department of Neurology, Tianjin Huanhu Hospital, Tianjin, China
[q]Department of Psychiatry, People's Hospital of Hangzhou Medical College, Zhejiang Provincial People's Hospital, Hangzhou, China
[r]Affiliated Hospital of Youjiang Medical University for Nationalities, Baise, China
[s]National Clinical Research Centre for Geriatric Disorders, Beijing, China
[t]Centre of Alzheimer's Disease, Beijing Institute for Brain Disorders, Beijing, China
[u]State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China

eClinicalMedicine
2023;65: 102276
Published Online xxx
https://doi.org/10.1016/j.eclinm.2023.102276

Summary

Background Alzheimer's disease (AD) is a prevalent neurodegenerative disorder that poses a worldwide public health challenge. A neuroimaging biomarker would significantly improve early diagnosis and intervention, ultimately enhancing the quality of life for affected individuals and reducing the burden on healthcare systems.

Methods Cross-sectional and longitudinal data (10,099 participants with 13,380 scans) from 12 independent datasets were used in the present study (this study was performed between September 1, 2021 and February 15, 2023). The Individual Brain-Related Abnormalities In Neurodegeneration (IBRAIN) score was developed via integrated regional- and network-based measures under an ensemble machine learning model based on structural MRI data. We systematically assessed whether IBRAIN could be a neuroimaging biomarker for AD.

Findings IBRAIN accurately differentiated individuals with AD from NCs ($AUC = 0.92$) and other neurodegenerative diseases, including Frontotemporal dementia (FTD), Parkinson's disease (PD), Vascular dementia (VaD) and Amyotrophic Lateral Sclerosis (ALS) ($AUC = 0.92$). IBRAIN was significantly correlated to clinical measures and gene expression, enriched in immune process and protein metabolism. The IBRAIN score exhibited a significant ability to reveal the distinct progression of prodromal AD (i.e., Mild cognitive impairment, MCI) (Hazard Ratio (HR) = 6.52 [95% CI: 4.42~9.62], $p < 1 \times 10^{-16}$), which offers similar powerful performance with Cerebrospinal Fluid (CSF) Aβ ($HR = 3.78$ [95% CI: 2.63~5.43], $p = 2.13 \times 10^{-14}$) and CSF Tau ($HR = 3.77$ [95% CI: 2.64~5.39], $p = 9.53 \times 10^{-15}$)

*Corresponding author. School of Artificial Intelligence, Beijing University of Posts and Telecommunications, Beijing 100876, China.
**Corresponding author.
E-mail addresses: yongliu@bupt.edu.cn (Y. Liu), shuyuli@bnu.edu.cn (S. Li).

based on the COX and Log-rank test. Notably, the IBRAIN shows comparable sensitivity (*beta* = –0.70, $p < 1 \times 10^{-16}$) in capturing longitudinal changes in individuals with conversion to AD than CSF Aβ (*beta* = –0.26, $p = 4.40 \times 10^{-9}$) and CSF Tau (*beta* = 0.12, $p = 1.02 \times 10^{-5}$).

Interpretation Our findings suggested that IBRAIN is a biologically relevant, specific, and sensitive neuroimaging biomarker that can serve as a clinical measure to uncover prodromal AD progression. It has strong potential for application in future clinical practice and treatment trials.

Funding Science and Technology Innovation 2030 Major Projects, the National Natural Science Foundation of China, Beijing Natural Science Funds, the Fundamental Research Funds for the Central University, and the Startup Funds for Talents at Beijing Normal University.

Keywords: Alzheimer's disease; Neuroimaging biomarker; Multisite; Specific; Longitudinal progression

Research in context

Evidence before this study

The available evidence highlighted the limitations of existing neuroimaging biomarkers, particularly their translation from research to clinical practice. Previous studies have explored deep learning and machine learning techniques to develop neuroimaging biomarkers for AD. Still, most studies have focused on single-site cohorts with limited sample sizes, potentially limiting generalisability. We searched PubMed, Web of Science, and Scopus for eligible studies from January 1, 2010 to March 30, 2023, mainly using the search terms ("classification" OR "diagnostic" OR "predict*") AND ("MRI" OR "Magnetic Resonance Imaging") AND ("Alzheimer*").

Added value of this study

Our study contributes to the existing evidence on neuroimaging biomarkers for AD. By developing the Individual Brain-Related Abnormalities In Neurodegeneration (IBRAIN) score, we address the limitations of previous biomarkers and offer several advancements. First, our study includes a large multisite cohort comprising over 10,000 participants, enhancing the generalisability of our findings. Second, IBRAIN integrates regional- and network-based measures using structural MRI data, providing a comprehensive and biologically relevant approach. Third, we systematically compare IBRAIN with CSF biomarkers, demonstrating its comparable performance in assessing clinical outcomes and longitudinal changes.

Implications of all the available evidence

Our findings have several implications for clinical practice, policy, and future research. The development of IBRAIN as a reliable and accessible neuroimaging biomarker holds promise for improving early detection, personalised prognosis, and monitoring of AD progression. IBRAIN's performance surpasses or matches that of CSF biomarkers, suggesting its potential as a less invasive alternative. The practicality and accessibility of IBRAIN make it a valuable tool for neurologists, radiologists, and researchers focused on AD biomarker development. Further research should explore the integration of IBRAIN into clinical workflows, evaluate its utility in diverse populations, and investigate its role in treatment trials and precision medicine approaches.

Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disorder representing a significant public health concern.[1] Quantifiable, individualised, and generalised biomarkers are critical for detecting prodromal AD, delineating its progression, and assessing the effectiveness of personalised prognosis.[2] The measurement of quantified Aβ using Positron Emission Tomography (PET) and Cerebrospinal Fluid (CSF) is a commonly employed method for diagnosing AD within professional medical organisations.[3] However, it is worth noting that this technique is not widely adopted in primary care settings, primarily due to its cost, invasiveness, and technical complexity,[4] which restricted the early diagnosis of AD within the high-risk population. Thus, developing a more generalised and readily available biomarker is urgently needed.

Structural Magnetic Resonance Imaging (sMRI) has become one of the most common measures for capturing the abnormal pattern of the brain in clinical practice.[5] Furthermore, advanced deep learning and machine learning techniques for constructing AD neuroimaging biomarkers based on sMRI provide novel insights for developing commonly used diagnosis assistants.[6,7] However, translating such biomarkers from bench to bedside still poses significant challenges, i.e., generalisation, biological explanation, cross-classical diagnostic categories, and identifying their subtle pathophysiological differences.[8] For these challenges, previous studies have emphasised the significance of validating biomarkers

using large datasets encompassing multiple research sites rather than relying solely on single datasets.[9,10] It is noteworthy that the most commonly employed framework for validation in previous studies remains k-fold or leave-n-out cross-validation,[10] despite a limited number of studies that have endeavoured to explore the applicability of inter-site cross-validation,[11,12] specifically within the inner ADNI dataset.[13–17] Therefore, extensive independent validation is needed to develop the MRI biomarker. Additionally, the early detection of AD is a multifaceted task that cannot be simplified as a two-class classification problem. Therefore, it is crucial to validate a specific biomarker that encompasses the cross-classification of diagnostic categories.[8] Moreover, validating a biomarker in clinical cohorts is a critical step often overlooked in clinical translation.[18] Many studies focus primarily on developing and evaluating biomarkers using research datasets, neglecting the essential aspect of proper validation in real-world clinical settings.[10,19]

To translate research advancements into clinical practice, a simple and robust model-based neuroimaging biomarker that can comprehensively capture the abnormal pattern of AD is needed due to its practicality and ease of interpretation compared to those complex and unexplainable models.[20] Furthermore, AD is a progressively degenerative condition characterised by a cross-scale complex phenotype, from subtle morphological changes[11] to macroscopic[21] and then to global brain damage.[22] Concentrating solely on a single-scale MRI measure is insufficient to fully capture the comprehensive abnormal pattern of AD. This limitation hinders the successful clinical translation of biomarkers. Therefore, the integration of cross-scale neuroimaging measures may be necessary to obtain a comprehensive assessment of the Brain-Related Abnormalities In Neurodegeneration (IBRAIN) for AD.

We hypothesised that an IBRAIN, derived from integrated different scales of brain structural measures, including subtle morphological changes (regional radiomics features, R2F) macroscopic (regional GM), brain covariance changes (regional radiomics similarity network, R2SN; and R2SN mean connectivity strength, RMCS), could serve as a reliable neuroimaging biomarker for AD. We aimed to test this hypothesis by evaluating: 1) the discriminative accuracy of IBRAIN in classifying AD from normal controls (NCs) and the samples with other neurodegenerative diseases; 2) the association between IBRAIN score and clinical symptoms and gene expression; and 3) the ability of IBRAIN to reveal distinct clinical outcomes or progression in NC and mild cognitive impaired (MCI) stages and to test its sensitivity in individuals with conversion to AD (Fig. 1).

Methods

Participants and clinical assessments

The present study involved the analysis of 12 independent datasets, encompassing a total of 10,099 participants with 13,380 scans of T1-weighted images (this study was performed between September 1, 2021 and February 15, 2023). The initial discovery dataset was derived from the Alzheimer's Disease Neuroimaging Initiative (ADNI; http://adni.loni.usc.edu), including 888 baseline participants (605 NCs and 283 AD, Supplementary Table S1). The data selection strategy employed in this study can be found in Supplementary Materials S01 (Supplementary Figure S1).

Additional testing datasets included, i.e., longitudinal participants and a complementary set of ADNI discovery datasets (entitled ADNI-L, as a primary testing dataset, 1877 participants with 4519 scans, Supplementary Materials S01, Supplementary Table S2), and other AD-NC datasets (a total of 4851 NCs and 2332 AD) (Table 1). In this study, a total of eight AD datasets were included, comprising 390 participants (M/F: 230/160) from the EDSD dataset, 419 participants (M/F: 357/62) from the AIBL dataset, 885 participants (M/F: 718/167) from OASIS dataset, 774 participants (M/F: 742/32) from ARWIBO dataset, 69 participants (708 scans) from MIRIAD dataset (M/F: 243/465), 3272 participants from NACC dataset (M/F: 2225/1047), 735 participants from MCADI dataset (M/F: 336/399). All participants comprising the MCADI dataset were approved by the Medical Ethics Committee of the local hospitals in China and signed written informed consent forms, with the remaining data sets being obtained from openly accessible sources.

It should be noted that the diagnosis of AD in these datasets was primarily based on clinical evaluation. Most of them followed NINCDS-AADRCA criteria.[23]

In addition, other neurodegenerative diseases cohorts contain individuals with Lewy body disease (LBD, N = 101), frontotemporal dementia (FTD, N = 184), Parkinson's disease (PD, N = 232), vascular dementia (VaD, N = 103), amyotrophic lateral sclerosis (ALS, N = 16), as well as those with other neurological, genetic, or infectious conditions (N = 35), depression (DEPR, N = 32) and cognitive impairment for other specified reasons (COR, N = 87) (Supplementary Materials S01, Supplementary Table S3), were also utilised for validation. Details can be found in Supplementary Materials S01.

Data processing and feature extraction

First, the T1-weighted image for each participant was first pre-processed by using the CAT12 toolbox (http://www.neuro.uni-jena.de/cat/, v12.8), and the regional GMV was quantified based on the Brainnetome atlas (246 brain regions).[24] Second, to evaluate the regional high-order morphological changes, the T1-weighted image was aligned to Montreal Neurological Institute (MNI) space by Advanced Normalisation Tools (ANTs) with "SyN" parameters for each participant. A total of 47 radiomics features (Detailed can be found in Supplementary Materials S02, Supplementary Tables S4

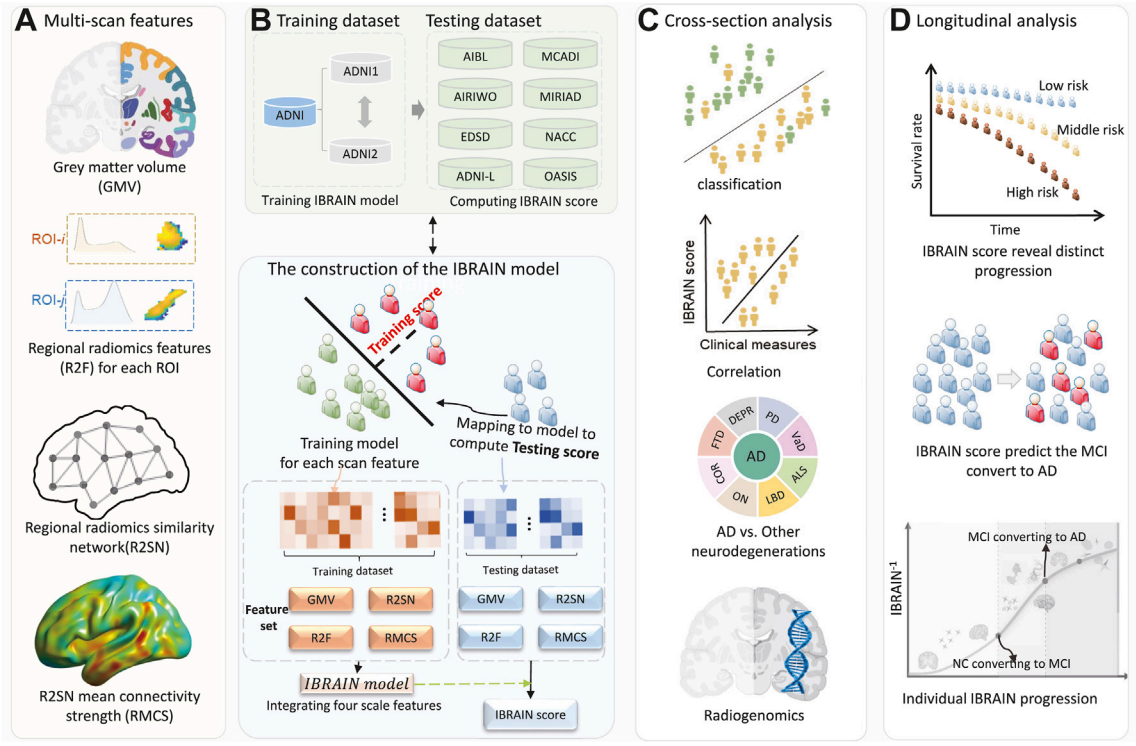

Fig. 1: **The pipeline of the present study.** (A) Multi-scan neuroimaging features were utilised in this study, including **grey** matter volume (GMV), regional radiomics features (R2F), regional radiomics similarity network (R2SN), and R2SN mean connectivity (RMCS). (B) Training and testing strategy employed in this study, including detailed progression for constructing the IBRAIN model. (C) Validation of the IBRAIN model in cross-sectional analysis, including its ability to distinguish AD and NC, its relationship to clinical measures and gene expression, and its ability to distinguish AD and non-AD disorders. (D) Validation of the IBRAIN in longitudinal analysis, including its ability to reveal the distinct progression of MCI, its ability to quantify MCI converting to AD within a certain period, and examining individual IBRAIN progression in MCI converting to AD.

and S5) for each brain region of the Brainnetome atlas were extracted.[11,14] Third, the individual R2SN was generated by mapping the individual's radiomics features (Supplementary Materials S03, Supplementary Table S6) into a radiomics similarity matrix of pair-wise interregional Pearson's correlations[25] (Detailed can be found in Supplementary Materials S03). Finally, the regional connectivity was also evaluated by computing the R2SN mean connectivity strength (RMCS) (Fig. 1A). The detailed brain regions' names in the Brainnetome atlas can be found in Supplementary Materials S03 (Supplementary Table S7).

Site	Group	Age (years)	Sex (M/F)
EDSD	NC (230)	68.76 ± 6.14	108/122
	AD (160)	72.52 ± 8.05	68/92
AIBL	NC (357)	67.93 ± 6.49	156/201
	AD (62)	69.31 ± 7.96	25/37
OASIS	NC (718)	66.46 ± 9.22	289/429
	AD (167)	74.78 ± 7.85	80/87
ARWIBO	NC (742)	51.87 ± 14.52	309/433
	AD (32)	73.50 ± 7.17	14/18
MIRIAD	NC (243)	69.86 ± 6.94	127/116
	AD (465)	69.56 ± 6.86	188/277
NACC	NC (2225)	65.58 ± 10.73	705/1520
	AD (1047)	71.72 ± 8.34	521/526
MCADI	NC (336)	64.71 ± 8.86	148/188
	AD (399)	69.23 ± 9.16	154/245
ADNI-L	NC (1402)	73.32 ± 6.11	662/740
	AD (915)	74.33 ± 7.57	523/392

Table 1: **The summary of the testing participants (scans) in the present study.**

Computing IBRAIN via ensemble machine learning

Initially, a classification model using a classical support vector machine (SVM) was developed for each feature set (GM, R2F, R2SN, and RMCS) using data from both AD and CN participants. To establish the optimal classification model for each feature set, internal 5-fold cross-validation was performed on the ADNI discovery dataset (N = 888), which resulted in the generation of training scores for each feature set. Subsequently, a linear regression model was re-trained to combine the training scores of the four feature sets via the discovery dataset. IBRAIN score was calculated for the testing set by integrating the four SVM prediction scores for each feature set based on the trained linear model (Fig. 1B).

Details can be found in Supplementary Materials S04 (Supplementary Figures S2 and S3). It is important to note that age, sex, and apoe4 status are known factors associated with neurodegeneration. To specifically capture the neurodegeneration-related changes, the IBRAIN was calculated by integrating various neuroimaging feature sets without incorporating age, sex, APOE4 status, or other demographic information. We aim to create a biomarker that directly reflects the underlying neurodegenerative processes by focusing solely on the neuroimaging features and excluding demographic variables.

Furthermore, to further explore whether the IBRAIN score was robust for different machine-specific sequence parameters, we included two traveling datasets (dataset 1: three healthy traveling participants with ten sites[26]; dataset 2: nine healthy participants with 12 sites[27]) to explore the repeatability of the IBRAIN score among different MRI machines (Supplementary Materials S01). The intraclass correlation coefficient (*ICC*) was used to estimate the reproducibility of IBRAIN. The *ICC* has a value between 0 and 1; $ICC = 0$ indicates no reproducibility, and $ICC = 1$ indicates absolute reproducibility.[28]

The distinguishability and the biological basis for IBRAIN

To evaluate the differentiability of IBRAIN in NC and AD, we examined the accuracy of AD and NC classification based on the IBRAIN score of eight independent datasets, with a total of 6253 NCs and 3247 ADs. In addition, we conducted a further comparison to assess the areas under the curve (*AUC*) for classifying individuals with AD and NC using IBRAIN in contrast to three representative deep learning models (ResNet,[29] 3DAN,[30] and 3DViT[31]) as well as regional GM volume (i.e., hippocampus, amygdala, and other altered brain regions). Details can be found in Supplementary Materials S05.

Subsequently, we conducted a correlation analysis to investigate the biological underpinnings of IBRAIN by examining its relationship with clinical and biological indices using the ADNI-L datasets ($N = 1877$ with 4519 scans, including 1402 NC, 2202 MCI, and 915 AD scans). It should be noted that age and sex were considered concomitant variables. Here, the biological indices encompassed CSF measures (amyloid-beta ($A\beta_{1-42}$), total Tau (Tau), and phosphorylation Tau_{181} ($p\text{-}tau_{181}$)) (https://adni.loni.usc.edu/summary-of-dates-of-csf-biomarker-analysis/). Those concentrations have been measured using the INNO-BIA AlzBio3 RUO test (Fujirebio, Ghent, Belgium). This test utilises a microbead-based multiplex immunoassay format performed on the Luminex platform. Additionally, global fluorodeoxyglucose (FDG) and global AV45 measures were incorporated as brain metabolism and amyloid deposition indicators, respectively. We also included the polygenic hazard score (PHS), which reflects an individual's genetic risk for developing AD. Regarding clinical indices, we included the mini-mental state examination (MMSE), a widely used screening tool for assessing cognitive function. The Alzheimer Disease Assessment Scale Cognitive score (ADAS-Cog) was utilised, specifically ADAS-Cog11, ADAS-Cog13, and ADAS-CogQ4, which assess various cognitive domains affected by AD pathology. The Rey Auditory Verbal Learning Test (AVLT) was included to evaluate immediate recall and learning abilities (Supplementary Materials S01, Supplementary Table S2).

Our study obtained gene expression profiles of peripheral blood samples from ADNI participants from Bristol-Myers Squibb (BMS) laboratories. The Affymetrix Human Genome U219 Array (www.affymetrix.com) was utilised for expression profiling, which includes 530,467 probes targeting 49,293 transcripts.[32] At last, 744 participants had genome-wide data with 18,635 genes among these participants. We then performed a Gene-set enrichment analysis to examine whether the genes significantly correlated with IBRAIN are associated with Gene Ontology terms via the Metascape platform (https://metascape.org/gp/index.html#/main/step1) (Fig. 1C). Metascape offers bioinformatics tools for gene annotation, pathway analysis, and functional enrichment analysis. It provides researchers with a platform to explore and interpret gene expression data by identifying enriched biological processes, molecular functions, cellular components, and pathways associated with a set of genes of interest.[33]

The specificity of IBRAIN: AD vs. eight other non-AD disorders

To investigate whether IBRAIN is specific for AD, we conducted the analysis for AD vs. non-AD participants classification. The IBRAIN score of the non-AD participants was also computed based on the AD-NC IBRAIN model trained in the ADNI baseline dataset. This step consists of 915 AD scans from the ADNI-L dataset and 790 participants with eight non-AD neurodegenerative diseases. Moreover, we further examined the discriminative performance of IBRAIN for AD and each non-AD disorder by implementing a bootstrap framework. Specifically, we randomly selected the same number of AD participants as each non-AD condition from the 915 AD scans 1000 times. We computed the distribution of all *AUC* values. This approach allowed us to evaluate the diagnostic performance of IBRAIN for AD and each non-AD disorder across multiple iterations, thereby increasing the robustness and generalisability of our findings (Fig. 1C).

Besides, to further investigate the performance of IBRAIN in accurately identifying AD within a population that includes individuals with mixed pathologies in real-life clinical practice. Two clinical cohorts were included in the present study (an OASIS-4 dataset and a

clinical cohort from the Affiliated Hospital of Youjiang Medical University for Nationalities of China). Details can be found in Supplementary Materials S01.

The association between IBRAIN and MCI longitudinal conversion

CSF Aβ and CSF total Tau are the most common upstream and downstream biomarkers for AD within the "AT(N)" framework. Therefore, in our study, we compared the performance of the association with AD progression between CSF Aβ/Tau and IBRAIN. We examined the capacity of IBRAIN to discern the likelihood of developing AD dementia among MCI individuals compared to the CSF Aβ and total Tau biomarkers. We subdivided the patients with MCI (662 scans had both IBRAIN, CSF Aβ, Tau, and their longitudinal information) into three subgroups with low- (first quartile), middle- (second and third quartile), and high-values (fourth quartile) via IBRAIN, CSF Aβ, and CSF total Tau, respectively. We employed Kaplan–Meier analysis to examine the longitudinal disease conversion within these subgroups. Initially, we computed the median survival time (*MST*) for each subgroup. Subsequently, we conducted the Log-rank test to assess the statistical significance of disease progression differences among the various subgroups. Furthermore, we calculated the hazard ratio (*HR*) with a 95% confidence interval (CI) for the high-risk group in comparison to the low-risk group, utilising the Cox test (Fig. 1D).

We computed the difference of the IBRAIN, CSF Aβ, and CSF total tau between stable MCI (sMCI) and progressive MCI (pMCI). We also endeavoured to develop a time-to-event prognostic Cox model to forecast the likelihood of individuals with MCI transitioning to AD based on integrating IBRAIN and clinical measures (Fig. 1D).

The longitudinal changes of IBRAIN from individuals with MCI converting to AD

The present study conducted a longitudinal analysis of a cohort consisting of 308 individuals diagnosed with MCI who subsequently progressed to AD over an average follow-up period of 2.51 ± 2.11 years. Among these participants, 211 had available CSF measures, and 141 had undergone more than two visits during the study period. A linear mixed model was employed to assess the progression of the IBRAIN score and CSF biomarker (CSF Aβ and CSF total Tau) over time while controlling for confounding factors such as age and sex. The objective was to examine the longitudinal changes in these biomarkers and evaluate their respective abilities to track disease progression in individuals with MCI converting to AD.

Role of the funding source

The funding provided support for data collection, but they did not participate in the study's design, data analysis, interpretation, or manuscript writing. Authors KZ and YL had access to the dataset and final responsibility for the decision to submit it for publication.

Results

IBRAIN exhibits diagnostic capabilities

First, the IBRAIN demonstrated a high *ICC* value between two traveling datasets (*ICC* = 0.98 in the first traveling dataset; *ICC* = 0.94 in the second traveling dataset). This finding further emphasises the transportability and stability of the IBRAIN measure, even in the presence of varying MRI machine parameters (Supplementary Materials S04, Supplementary Table S8). The construction and evaluation of the IBRAIN model were conducted using a discovery dataset of 888 participants from the ADNI baseline dataset (Fig. 2A). Subsequently, the performance of the model was assessed using a testing dataset comprising 9500 scans diagnosed with either AD or NC from eight independent datasets. AD and NC were classified using the IBRAIN framework, which yielded an impressive *AUC* of 0.92 across all testing datasets (Fig. 2B). Remarkably, the *AUC* exceeded 0.95 in the ADNI-L, EDSD, MIRIAD, and AIBL datasets, highlighting the robustness of the model in accurately distinguishing between AD and NC. The IBRAIN demonstrated high discriminative accuracy in distinguishing AD with Aβ+ from NC with Aβ–, with an *AUC* of 0.97. Furthermore, it also performed well in differentiating individuals with MCI with Aβ+ from NC with Aβ–, with an *AUC* of 0.83 (Supplementary Materials S05, Supplementary Figure S4).

The IBRAIN exhibited a comparable classification accuracy to advanced deep learning models while maintaining low complexity and high interpretability (Supplementary Materials S05, Supplementary Table S11). Furthermore, integrating multiple scan features for AD and NC classification exhibited superior performance compared to each feature set in isolation and regional GM volume (i.e., hippocampus, amygdala), thus demonstrating the effectiveness of the fusion strategy employed in this study (Supplementary Materials S05, Supplementary Table S12).

Furthermore, we quarter-categorised patients with AD of ADNI-L into five stages based on their cognitive ability assessed by the MMSE score. These stages included severe AD (MMSE score 1–19), late-middle AD (MMSE score 19–21), early-middle AD (MMSE score 21–23), and early AD (MMSE score 23–30). Our findings showed that the IBRAIN model performed excellently in differentiating severe AD (*ACC* = 99.58%), late-middle AD (*ACC* = 95.26%), early-middle AD (*ACC* = 96.72%), and early AD (*ACC* = 84.39%).

IBRAIN exhibits clinically relevant

Pearson correlation analysis revealed that 623 genes were significantly associated with IBRAIN (all $p < 0.01$).

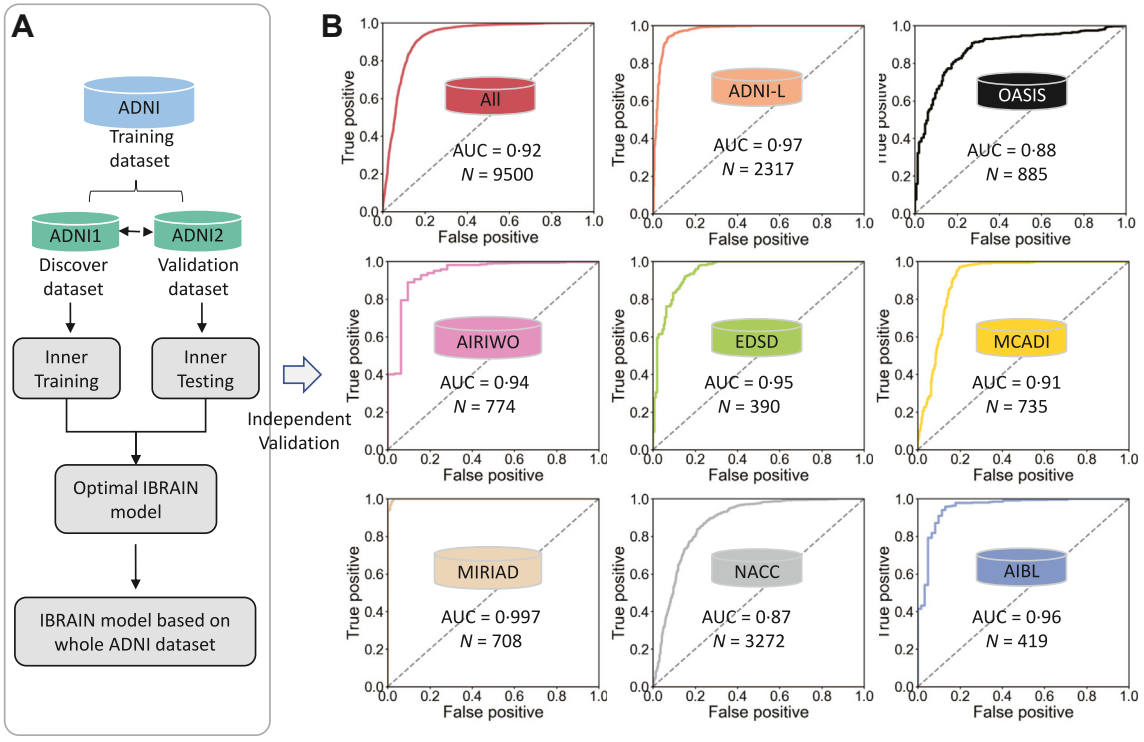

Fig. 2: **The results for distinguishing AD and NC.** (A) The IBRAIN model was developed using an inner cross-validation strategy applied to the ADNI dataset. The dataset was partitioned into inner training and testing sets, with the inner training set used for constructing the IBRAIN model using different parameter settings. The performance of the model was assessed on the inner testing set to determine the optimal configuration of the IBRAIN model for accurately distinguishing between individuals with AD and NC. The process of the training IBRAIN model. (B) The receiver operating character (ROC) curve and areas under the ROC curve (AUC) for all testing datasets and each independent dataset.

Furthermore, our gene enrichment analysis demonstrated that immunity-related terms, such as neutrophil degranulation (R-HSA-6798695, $p = 1.20 \times 10^{-26}$), leukocyte activation (GO:0045321, $p = 3.89 \times 10^{-14}$), immune response-regulating signalling pathway (GO:0002764, $p = 4.27 \times 10^{-14}$), and innate immune response (GO:0045087, $p = 1.35 \times 10^{-11}$), as well as protein metabolism-related terms, such as metabolism of lipids (R-HSA-556833, $p = 3.98 \times 10^{-8}$), and protein phosphorylation (GO:0006468, $p = 4.90 \times 10^{-7}$), were significantly associated with IBRAIN (Fig. 3A). Detailed results for gene enrichment analysis can be found in Supplementary Materials S06 (Supplementary Table S13). The Metascape enrichment network visualisation comprehensively represents the relationships and similarities between enriched terms. It depicts both intra-cluster and inter-cluster connections, highlighting up to ten terms per cluster. Each cluster is assigned a unique colourcode, allowing for easy identification and interpretation of the results. By visualising the functional associations and patterns among the enriched terms, the network offers valuable insights into the underlying biological processes associated with IBRAIN. It should be noted that the results mentioned above serve as unsubstantial evidence to support the hypothesis that IBRAIN may be related to brain immunity and protein metabolism. This uncertainty assertion is contingent upon the absence of a pathologically confirmed dataset.

Moreover, our analysis revealed significant differences in IBRAIN scores between the NC, MCI, and AD groups ($p < 1 \times 10^{-16}$). We also found significant associations between IBRAIN scores and various clinical measures, including CSF Aβ, CSF total Tau, CSF *p*-tau$_{181}$, FDG, AV45, PHS, MMSE, Alzheimer Disease ADAS-Cog, i.e., ADAS-Cog11, ADAS-Cog13, ADAS-CogQ4, AVLT-immediately, and AVLT-learning (all $p < 1 \times 10^{-12}$) (Supplementary Materials S06, Supplementary Figure S7). Furthermore, the correlation between cognitive ability and IBRAIN was higher than that between cognitive ability and FDG, CSF Aβ, CSF Tau, and CSF *p*-tau (Fig. 3B). CSF Aβ ($p = 9.94 \times 10^{-4}$) and CSF total Tau ($p = 5.80 \times 10^{-3}$) showed significant differences between IBRAIN negative (IBRAIN >0) and IBRAIN positive (IBRAIN <0) (Supplementary Materials S06, Supplementary Figure S7).

IBRAIN is a diagnostic specificity neuroimaging biomarker for AD

The IBRAIN score exhibited significant differences between individuals with AD ($N = 915$) and those without

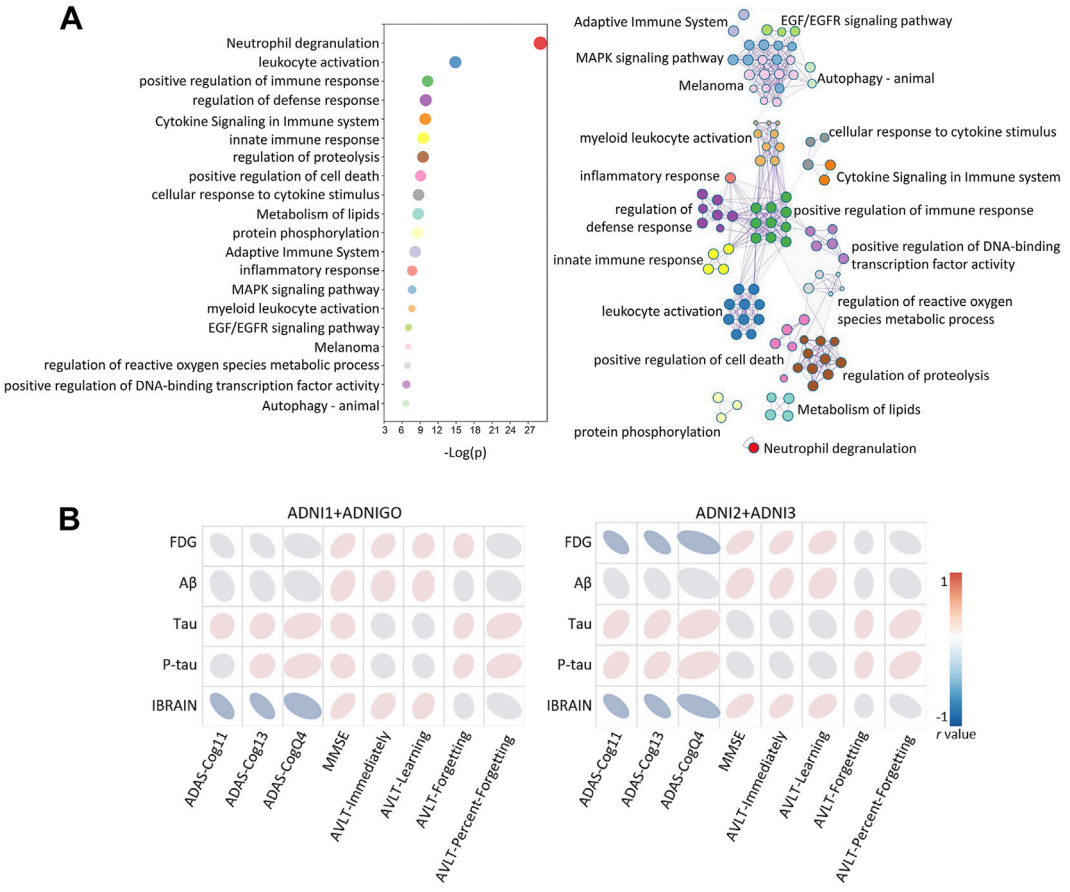

Fig. 3: **The biological basis of the IBRAIN.** (A) The most correlated terms and gene pathways associated with IBRAIN include immunity-related terms such as neutrophil degranulation, leukocyte activation, immune response-regulating signalling pathway, and innate immune response. Additionally, protein metabolism-related terms such as metabolism of lipids and protein phosphorylation are also significantly associated with IBRAIN. (B) The correlation between existing biomarkers and IBRAIN to cognitive ability in the ADNI1&GO dataset and ADNI2&3 dataset, respectively. Here, the biomarkers primarily comprise CSF Aβ, Tau181, Ptau181, and global FDG, while the cognitive ability is primarily assessed by MMSE, ADAS-Cog score, and AVLT score.

AD (N = 789) ($p < 1 \times 10^{-16}$). Among all non-AD disorders, the distribution of IBRAIN scores in FTD was found to be most similar to that observed in AD (Fig. 4A and B). The IBRAIN score also demonstrated diagnostic capabilities in distinguishing AD from non-AD conditions, with an AUC of 0.92 (Fig. 4C). In addition, using a bootstrap framework, we found that the AUC was higher than 0.90 in distinguishing AD from PD, VaD, ALS, ON, DEPR, and COR, respectively, but lower in differentiating AD from LBD (AUC = 0.85) and FTD (AUC = 0.83) (Fig. 4D). Herein, our results indicate that the IBRAIN score can potentially differentiate individuals with AD from those with non-AD disorders across classical diagnostic categories.

The results of our study demonstrate that the IBRAIN biomarker exhibits promising performance in recognising AD from individuals with mixed pathologies using the OASIS-4 dataset. Specifically, IBRAIN achieved an AUC of 0.78 for distinguishing AD (AD variants, non-neurodegenerative disorders, AD with coexisting vascular pathology, and Alzheimer's disease dementia) from other conditions. Additionally, IBRAIN correctly identified 11 out of 15 AD cases within the Guangxi hospital clinical cohort. However, it is essential to acknowledge that there were also some instances where patients with non-AD were misdiagnosed as AD (Supplementary Materials S05, Supplementary Figure S5).

IBRAIN score mirrors the conversion from MCI to AD

The present study employed Kaplan–Meier analysis to investigate differences in conversion timing among various subgroups stratified by IBRAIN, CSF $A\beta_{42}$, and CSF Tau within a subset of the ADNI-L. Our results underscore the efficacy of IBRAIN in effectively discerning distinct longitudinal clinical conversion

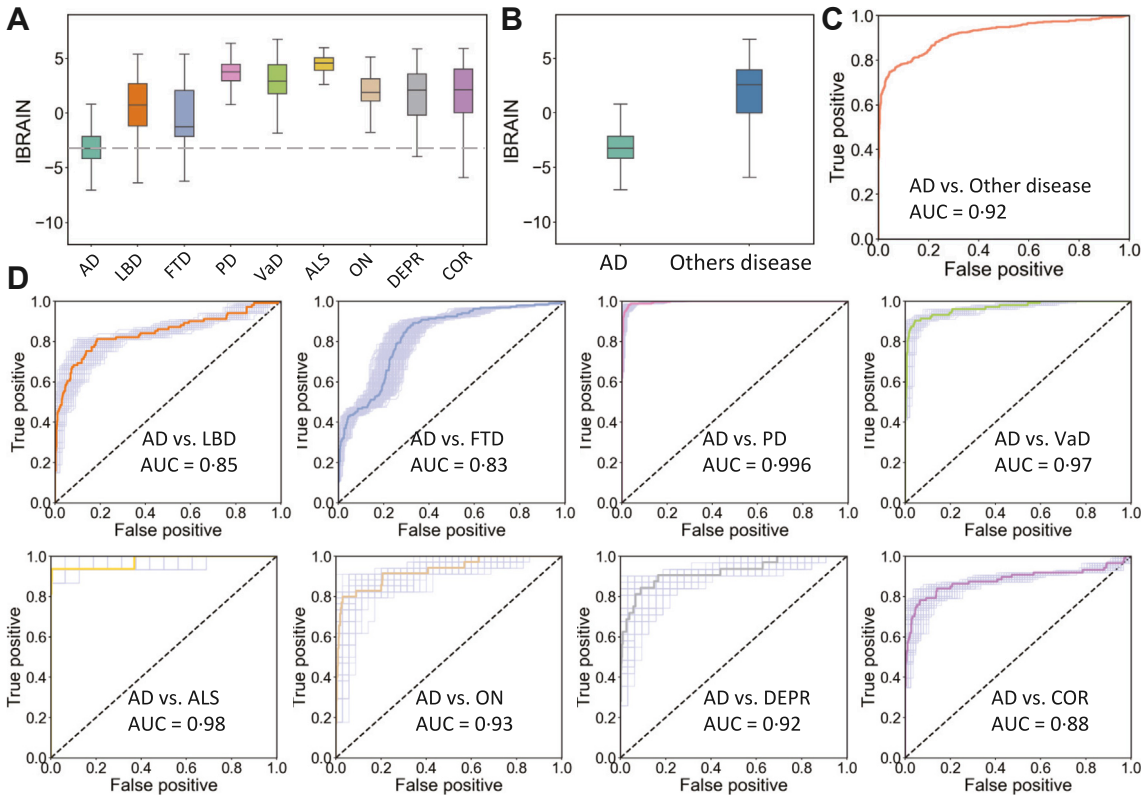

***Fig. 4:* The specific of the IBRAIN for AD.** (A) The distribution of IBRAIN levels in patients with AD and patients with Lewy Body disease (LBD), Frontotemporal Dementia (FTD), Parkinson's dementia (PD), Vascular dementia (VaD), Amyotrophic Lateral Sclerosis (ALS), other neurologic, genetic or infectious conditions (ON), Depression (DEPR), and Cognitive impairment for other specified reasons, i.e., written-in values (COR). (B) The distribution of IBRAIN in AD and all non-AD disorders. (C). The ROC curve for differentiation AD and patients with non-AD. (D) The ROC curve for differentiation of AD and each non-AD disorder via a bootstrap framework.

trajectories among individuals diagnosed with MCI. Notably, the high-risk subgroup, as identified by IBRAIN, displayed a notably abbreviated *MST* of 2 years (95% CI: 1.39~2.61). Conversely, the low-risk subgroup exhibited a substantially lengthier *MST* of 10 years (95% CI: 9.98~10.02). The *HR* for the high-risk group, relative to the low-risk group, was notably elevated at 6.52 (95% CI: 4.42~9.62), with statistical significance reflected in a *p*-value of 9.27×10^{-26}. Besides, IBRAIN exhibited comparable performance to that of CSF Aβ (*HR* = 3.78 [95% CI: 2.63~5.43], $p = 2.13 \times 10^{-14}$) and CSF Tau (*HR* = 3.77 [95% CI: 2.64~5.39], $p = 9.53 \times 10^{-15}$) (Fig. 5A).

Then, the results also suggested that IBRAIN exhibits comparable performance to CSF biomarker combinations, i.e., CSF $A\beta_{42}/A\beta_{40}$, CSF $Tau/A\beta_{40}$, and CSF $Ptau_{181}/A\beta_{40}$. This result was obtained based on the 306 patients with MCI in ADNI2 (Supplementary Materials S07, Supplementary Figure S10). Furthermore, similar results were obtained in all ADNI-L patients with MCI (Supplementary Materials S07, Supplementary Figure S8) or the NC converting to AD groups (Supplementary Materials S07, Supplementary Figure S9). IBRAIN better identifies distinct longitudinal clinical conversion with MCI than hippocampal and amygdala volumes (Supplementary Materials S07, Supplementary Table S15). The IBRAIN derived from the R2F showed the highest performance for individual feature sets in identifying distinct longitudinal clinical conversion with MCI (Supplementary Materials S07, Supplementary Table S17).

We observed significant differences in the levels of IBRAIN (*Cohen's d* = 1.20 [95% CI: 0.89~1.21], $p < 1 \times 10^{-16}$), CSF Aβ (*Cohen's d* = 0.88 [95% CI: 0.85~0.92], $p < 1 \times 10^{-16}$), and CSF Tau (*Cohen's d* = –0.74 [95% CI: –0.77 to –0.72], $p < 1 \times 10^{-16}$) between the sMCI and pMCI groups (Supplementary Materials S07, Supplementary Figure S11). Notably, IBRAIN demonstrated an *AUC* greater than 0.8 in predicting clinical conversion in patients with MCI (Supplementary Materials S07, Supplementary Figure S12). Besides, the IBRAIN was significantly correlated with AD progression in two independent datasets (Supplementary Materials S07, Supplementary Table S18).

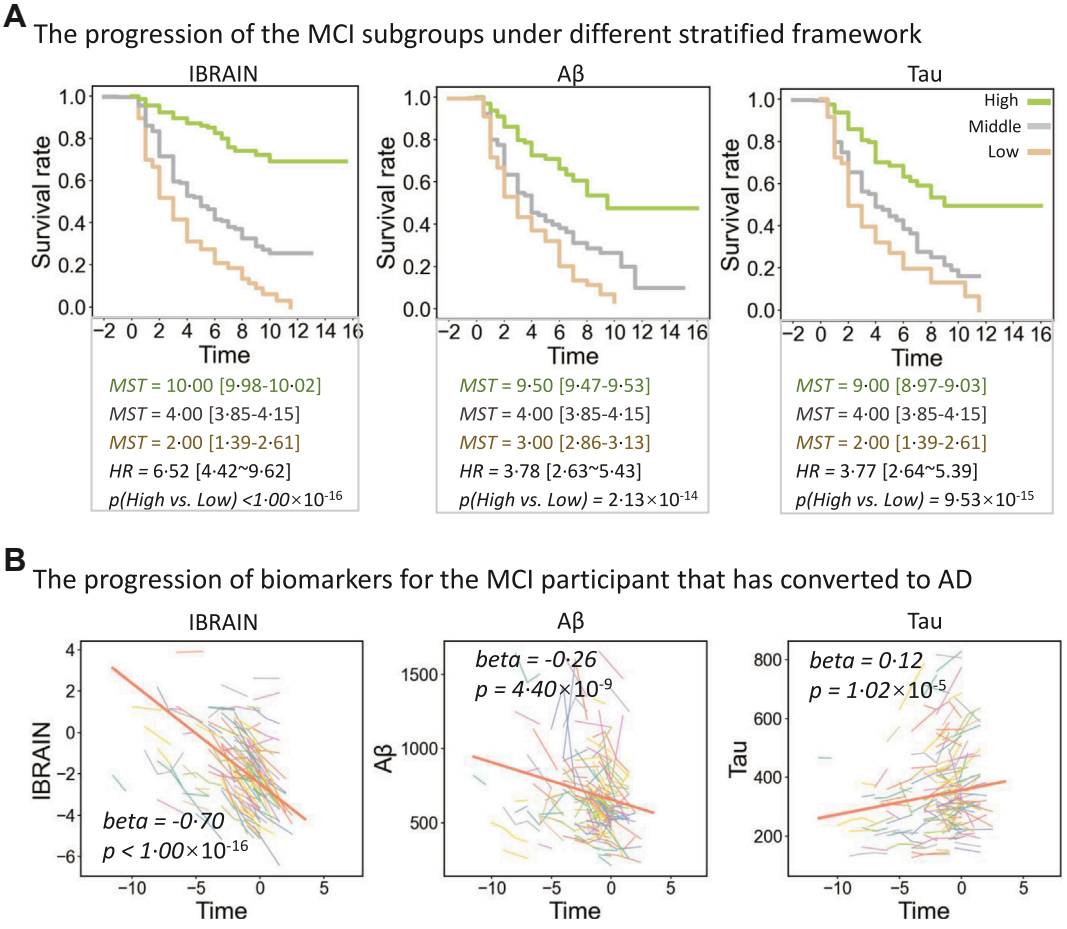

Fig. 5: **Longitudinal analysis of the IBRAIN in patients with MCI.** (A) The conversion of the MCI participants under the stratified framework via IBRAIN, CSF Aβ, and CSF Tau, respectively. Here, the patients with MCI were subdivided into three subgroups with low- (first quartile), middle- (second and third quartile), and high-values (fourth quartile) via IBRAIN, CSF Aβ, and CSF total Tau. (B) The individual progression of the IBRAIN, CSF Aβ, and CSF Tau in the patients with MCI that have undergone conversion to AD.

The longitudinal change of the biomarkers from individuals with MCI converting to AD

In this study, we investigated the utility of individual IBRAIN as a tracer in the conversion of MCI to AD. In addition, we compared the diagnostic performance of IBRAIN with the commonly used biomarkers CSF Aβ and CSF total Tau. Our findings indicate that all three biomarkers, IBRAIN ($beta = -0.70$, $p < 1 \times 10^{-16}$), CSF Aβ ($beta = -0.40$, $p = 4.40 \times 10^{-9}$), and CSF Tau ($beta = 0.12$, $p = 1.02 \times 10^{-5}$), were statistically significant in a linear mixed model (Fig. 5B). These results suggest that IBRAIN may serve as a valuable diagnostic tool for identifying individuals at risk of developing AD. IBRAIN performs better statistical parameters of the longitudinal progression in conversion with MCI to AD than hippocampal and amygdala volumes (Supplementary Materials S07, Supplementary Table S14). For individual feature sets, the IBRAIN derived from the R2F showed the highest statistical significance in the longitudinal progression in conversion with MCI to AD (Supplementary Materials S07, Supplementary Table S16).

Discussion

We have proposed an integrated machine learning framework that is simple and effective for developing a novel biomarker, IBRAIN, for AD diagnosis. The findings indicate that IBRAIN has the great potential to serve as a valid and reliable neuroimaging biomarker for detecting and monitoring AD progression in individuals with a high risk. Furthermore, the collective evidence suggests that IBRAIN possesses considerable promise for clinical implementation owing to its practicality and accessibility.

In much primary care in China and other countries that have not yet reached a highly developed level of medical care, brain MRI is routinely obtained for

individuals at risk for cognitive decline.[34] Still, the diagnostic and prognostic yield of MRI remains low.[8] In translating research advancements into the clinic, specific biomarkers are needed to evaluate comprehensive brain abnormal patterns.[35] The present study proposed an integrated machine-learning approach to investigate AD's local and global abnormal patterns. The integrated model demonstrated superior performance compared to isolated feature sets, highlighting the effectiveness of the fusion strategy. IBRAIN is computed utilising an integrated machine learning framework, demonstrating similar accuracy in classifying individuals with AD from NC compared to several deep learning models.[12] Notably, IBRAIN maintains a low complexity and offers high interpretability. Moreover, IBRAIN can capture multiscale brain abnormal patterns at the individual level, encompassing subtle regional morphological changes and macroscopic alterations and disruptions to the brain's connectome. Given that any progressions in these MRI scans would have a considerable effect on the worldwide diagnostic and prognostic assessment of AD,[36] as a general principle, applying a simplified model endows a heightened level of explicability, which renders it advantageous for comprehending the mechanisms underlying AD and facilitating the development of prognostic therapy. The IBRAIN model showed promise in identifying advanced cognitive impairments in the severe and middle stages of AD, further refinement is needed to enhance its sensitivity in detecting subtle cognitive changes in early AD. Future research efforts should focus on optimising and validating the IBRAIN model, particularly in the early stages of AD.

Compared to the other three feature sets, the limited weight assigned to RMCS in the IBRAIN model can be attributed to the potential lack of independence between RMCS and the other features. This phenomenon suggests that RMCS may capture overlapping information already captured by the other feature sets. However, despite its relatively lower weight, including RMCS in constructing the IBRAIN score remains crucial. RMCS allows us to capture and quantify specific changes in regional connectivity patterns uniquely associated with AD pathology. This integration of RMCS enhances the comprehensive nature of the IBRAIN model. It provides valuable insights into the brain's connectivity network disruptions in individuals with AD. Besides, it is essential to recognise that different feature sets may exhibit specific performance in different stages of AD. Our study observed that the grey matter volume feature set demonstrated superior performance in distinguishing AD from normal controls. This result suggests that GM set is particularly informative for classifying AD at a global level. Meanwhile, the regional radiomics features showed the highest statistical significance in the longitudinal conversion from MCI to AD, highlighting the importance of considering regional-level subtle alterations in the brain for assessing disease progression and identifying individuals at higher risk of conversion. IBRAIN, as a comprehensive biomarker, integrates multiple feature sets to offer a comprehensive assessment of AD progression. This integration enhances the classification accuracy compared to R2F and improves the sensitivity compared to GM. These findings demonstrate the effectiveness of the fusion strategy employed in this study.

An optimal biomarker should be linked to either an upstream or a downstream biomarker.[8] The current study demonstrates that the IBRAIN biomarker exhibits a significant correlation with the common upstream biomarker (CSF Aβ) and the common downstream biomarker (CSF total tau) for AD within the "AT(N)" framework.[37] Therefore, IBRAIN might mirror the changes of indiscernible pathological alterations that are indiscernible utilising traditional sMRI diagnostic methodologies. In a word, IBRAIN is a biologically relevant neuroimaging biomarker. More importantly, the longitudinal analysis revealed the remarkable sensitivity and potential applicability of IBRAIN in unveiling disease progression in high-risk individuals, surpassing even the capabilities of CSF biomarkers. Previously proposed MRI biomarkers, i.e., single grey matter volume and cortex thickness, also show low intra-individual variability over time,[38] limiting the ability to track disease progression and downstream pathology.[8] Hence, IBRAIN exhibits the potential to advance the current clinical diagnosis framework and treatment for AD.

Radiogenomics created a bridge between IBRAIN and its underlying biological mechanisms. A gene enrichment analysis has demonstrated a significant correlation between IBRAIN and immunity-related terms. Neuroinflammation is a complex and dynamic process involving activating immune cells and releasing pro-inflammatory molecules within the central nervous system. It is increasingly recognised as a significant factor in the pathogenesis and progression of AD.[39] Neuroinflammation in AD is characterised by the activation of microglia and astrocytes, increased production of pro-inflammatory cytokines, and the presence of immune cells in affected brain regions. This inflammatory response is believed to contribute to the progression of neurodegeneration and the accumulation of amyloid-beta plaques and neurofibrillary tangles, leading to neuronal damage and cognitive decline.[40] In our study, the gene enrichment analysis revealed significant associations between IBRAIN and immunity-related terms, such as neutrophil degranulation, leukocyte activation, immune response-regulating signalling pathway, and innate immune response. These findings suggest that neuroinflammation may play a role in the underlying mechanisms captured by IBRAIN. However, further investigation is needed to elucidate the precise mechanisms by which neuroinflammation influences

the IBRAIN biomarker and its relationship to AD pathology.

Interestingly, IBRAIN was also correlated with terms related to protein metabolism or phosphorylation. Abnormalities in these processes have been linked to the accumulation of misfolded proteins, such as Aβ, Tau, and *p*-tau, which also target neuronal damage and cognitive decline.[41] Our finding emphasises the significance of the immune-mediated mechanisms that precipitate protein irregularities in AD. Furthermore, the results imply that IBRAIN, an index derived from comprehensive multi-scan features, is a promising potential indicator for both immune and protein-associated processes. Based on this speculation, IBRAIN might exhibit potential for predicting the therapeutic outcomes of anti-inflammatory agents and small-molecule drugs targeting AD protein metabolism. It is imperative to establish the veracity of this hypothesis within a pathologically confirmed dataset, specifically by investigating the activity of glial cells.

In clinical practice, the early detection of AD is a complex task that cannot be taken as a two-class classification problem. There exist significant challenges in clinical settings, as there is substantial overlap among neurodegenerative diseases that share similar pathophysiology, such as tauopathy,[42] brain structure abnormalities,[43] and genetic risk.[44] Any biomarker employed for AD detection demonstrates high specificity that can accurately differentiate AD from samples with other neurodegenerative diseases, such as FTD and VaD.[8] Consequently, we still lack AD-specific imaging biomarkers that can be used to guide clinical decisions.[45,46] This observation suggests that the IBRAIN framework may be capable of detecting and characterising specific cerebral modifications associated with AD. These characteristic holds promise for the potential clinical implementation of the IBRAIN. Upon the successful conclusion of biomarker discovery and validation, which involves the analytical validation of the biomarker assay and the clinical validation of its accuracy, the clinical utility phase of biomarker development ensues. As such, prospective randomised clinical trials are warranted to assess the effects of integrating IBRAIN into clinical workflows.

Validating a biomarker in clinical cohorts is a crucial and often overlooked step in achieving clinical translation.[18] Many studies focus on developing and evaluating biomarkers using research datasets without adequate validation in real-world clinical settings.[10] We systematically assessed the feasibility of IBRAIN for application in clinical practice based on two independent clinical cohorts. The findings from the validation in clinical cohorts are instrumental in establishing the clinical utility and reliability of the IBRAIN biomarker. We have taken a significant step toward achieving clinical transformation by confirming its accuracy and robustness in real-world clinical settings. The results of the classification analysis for distinguishing AD from non-AD individuals in two clinical cohorts offer valuable insights from two perspectives. On the one hand, the findings demonstrate the promising potential of IBRAIN in accurately identifying AD within a population that includes individuals with mixed pathologies, as observed in the OASIS-4 dataset. On the other hand, it is essential to acknowledge the limitations associated with the reconstruction of low-resolution T1 images, which may have impacted the performance of IBRAIN. Therefore, addressing these limitations and refining the IBRAIN model to optimise its accuracy and performance, specifically in reconstructing low-resolution T1 images, is imperative. By addressing these challenges, we can enhance the reliability and applicability of IBRAIN as a neuroimaging biomarker for AD diagnosis.

Several recent studies have demonstrated the diagnostic performance[4,47–49] and predicted the longitudinal progression of AD in individuals at high risk of blood biomarkers, such as $p\text{-tau}^{181}$ and $p\text{-tau}^{217}$.[3,50–53] We acknowledge the potential of blood biomarkers as promising tools for AD diagnosis and prognosis. The accessibility and cost-effectiveness of blood biomarkers make them particularly attractive for large-scale screening and routine clinical use. Unlike blood biomarkers, IBRAIN can comprehensively assess the brain's abnormal patterns, capturing information at different scales and regions, allowing for a more holistic understanding of the disease and its progression. It is important to note that the choice between neuroimaging and blood biomarkers is not mutually exclusive. Both approaches have advantages and limitations, and a combined multimodal approach may offer even greater diagnostic and prognostic accuracy.

IBRAIN robustly distinguished disease from a healthy brain and was able to capture early pathological alterations indicative of prodromal AD. Furthermore, this novel methodology provides the capacity for a personalised imaging biomarker with high precision and sensitivity. It's worth noting that the implementation of mass screening with IBRAIN could pose several challenges, including cost considerations and resource limitations. Consequently, our primary vision for IBRAIN is initially focused on targeted screening, particularly in high-risk populations. This approach is viable within clinical practice primarily because IBRAIN exclusively relies on high-resolution sMRI data. This strategy aligns with contemporary healthcare practices, where resource allocation often prioritises individuals based on their level of risk and need. Another pivotal factor supporting the suitability of IBRAIN for targeted screening in high-risk populations is its foundation on high-resolution MRI data. This sets it apart from typical clinical cohorts that frequently employ lower-resolution data. Concurrently, we are actively promoting the development of a novel IBRAIN model based on general clinical cohorts with lower-resolution data. This

initiative aims to broaden the adoption and utilisation of IBRAIN across various applications. Furthermore, our current study underscores IBRAIN's significant correlation with common upstream biomarkers and common downstream biomarkers within the "AT(N)" framework for AD. This suggests that IBRAIN has the potential to capture subtle pathological alterations that may not be discernible through traditional sMRI diagnostic methods. Moreover, IBRAIN represents a neuroimaging biomarker that amalgamates diverse neuroimaging features across multiple scales, thus providing a more comprehensive means of characterising abnormal brain patterns in AD. Consequently, IBRAIN holds promise as a potential prognostic index for AD. The attainment of this objective within clinical practice, whether in radiology departments, neurology units at local hospitals, or high-end physical examination institutions, is eminently feasible due to IBRAIN's exclusive reliance on high-resolution sMRI data. As a result, IBRAIN exhibits substantial potential for future translational applications.

Notwithstanding the progress, the present study has several limitations. Firstly, a significant proportion of the study participants lacked CSF or PET measures, potentially leading to false positives in diagnosing AD. To address this limitation and enhance the reliability of our findings, we will collect and obtain access to datasets encompassing both MRI and CSF measures in the future. Secondly, the select nature of the populations should also be further considered, as the study cohorts appear to favour a larger number of non-Hispanic white people greatly. We will additionally strive to include more diverse cohorts to enhance the generalisability and applicability of our findings. Thirdly, it is important to emphasise a limitation regarding the application of IBRAIN in clinical practice. IBRAIN was developed using high-resolution MRI data, setting it apart from typical clinical cohorts that often employ lower-resolution data. As a result, its applicability may be constrained, and it is best suited for use in high-risk populations with access to high-resolution MRI data.

Contributors

Wrote the draft: Kun Zhao and Pindong Chen; Study design: Kun Zhao, Yong Liu, and Shuyu Li; Data collection: Dawei Wang, Dongsheng Fan, Shan Ye, Yi Tang, Hongxiang Yao, Bo Zhou, Jie Lu, Chunshui Yu, Pan Wang, Zhengluan Liao, Yan Chen, Xi Zhang, Longjian Huang, and Ying Han; Data processing: Kun Zhao, Fan Yang, and Xiaopeng Kang; Statistical analysis: Kun Zhao and Yong Liu; and Manuscript revision: Pindong Chen, Aaron Alexander-Bloch, Yongbin Wei, Martin Dyrba, Shuyu Li, and Yong Liu; Data access and verification: Kun Zhao, Pindong Chen, and Yong Liu. They have full access to all the data in the study. All authors read and approved the final version of the manuscript and accept responsibility for the decision to submit it for publication.

Data sharing statement

In the current study, all coding and IBRAIN model has been made publicly available through the GitHub repository maintained by YongLiuLab (https://github.com/YongLiuLab). All of these datasets were publicly available or had conditional acquisition from the corresponding author. Detailed information regarding these datasets and associated toolkits are provided in Supplementary Materials S08.

Declaration of interests

The authors report no biomedical financial interests or potential conflicts of interest.

Acknowledgements

This work was partially supported by the Science and Technology Innovation 2030 Major Projects (No. 2022ZD0211600), the National Natural Science Foundation of China (Nos. 82172018, 81972160, 62333002, 82001350, 82202264), the Beijing Natural Science Funds for Distinguished Young Scholars (No. JQ20036), the Fundamental Research Funds for the Central Universities, China (No. 2021XD-A03), Beijing Municipal Natural Science Foundation (No. 7232341), the Startup Funds for Talents at Beijing Normal University.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.eclinm.2023.102276.

References

1 Gauthier S, Rosa-Neto P, Morais JA, Webster C. *World Alzheimer report 2021: journey through the diagnosis of dementia*. 2021.
2 Adkins-Jackson PB, Belsky DW. Alzheimer's disease risk biomarkers: progress and challenges. *Lancet Healthy Longev*. 2022; 3(9):e575–e576. https://doi.org/10.1016/S2666-7568(22)00191-X.
3 Moscoso A, Grothe MJ, Ashton NJ, et al. Longitudinal associations of blood phosphorylated Tau181 and neurofilament light chain with neurodegeneration in Alzheimer disease. *JAMA Neurol*. 2021;78(4):396–406. https://doi.org/10.1001/jamaneurol.2020.4986.
4 Thijssen EH, La Joie R, Strom A, et al. Plasma phosphorylated tau 217 and phosphorylated tau 181 as biomarkers in Alzheimer's disease and frontotemporal lobar degeneration: a retrospective diagnostic performance study. *Lancet Neurol*. 2021;20(9):739–752. https://doi.org/10.1016/S1474-4422(21)00214-3.
5 Plachti A, Kharabian S, Eickhoff SB, et al. Hippocampus co-atrophy pattern in dementia deviates from covariance patterns across the lifespan. *Brain*. 2020;143(9):2788–2802. https://doi.org/10.1093/brain/awaa222.
6 Rathore S, Abdulkadir A, Davatzikos C. Analysis of MRI data in diagnostic neuroradiology. *Annu Rev Biomed Data Sci*. 2020; 3(1):365–390. https://doi.org/10.1146/annurev-biodatasci-022620-015538.
7 Martin SA, Townend FJ, Barkhof F, Cole JH. Interpretable machine learning for dementia: a systematic review. *Alzheimers Dement*. 2023;19(5):2135–2149. https://doi.org/10.1002/alz.12948.
8 Hansson O. Biomarkers for neurodegenerative diseases. *Nat Med*. 2021;27(6):954–963. https://doi.org/10.1038/s41591-021-01382-x.
9 Woo CW, Chang LJ, Lindquist MA, Wager TD. Building better biomarkers: brain models in translational neuroimaging. *Nat Neurosci*. 2017;20(3):365–377. https://doi.org/10.1038/nn.4478.
10 Wu J, Zhao K, Li Z, et al. A systematic analysis of diagnostic performance for Alzheimer's disease using structural MRI. *Psychoradiology*. 2022;2(1):1–9. https://doi.org/10.1093/psyrad/kkac001.
11 Zhao K, Ding Y, Han Y, et al. Independent and reproducible hippocampal radiomic biomarkers for multisite Alzheimer's disease: diagnosis, longitudinal progress and biological basis. *Sci Bull*. 2020;65(13):1103–1113. https://doi.org/10.1016/j.scib.2020.04.003.
12 Zhao Q, Huang G, Xu P, et al. IDA-Net: inheritable deformable attention network of structural MRI for Alzheimer's disease diagnosis. *Biomed Signal Process Control*. 2023;84:104787. https://doi.org/10.1016/j.bspc.2023.104787.
13 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(8):1059–1070. https://doi.org/10.1016/j.jalz.2019.02.007.
14 Zhao K, Zheng Q, Dyrba M, et al. Regional radiomics similarity networks reveal distinct subtypes and abnormality patterns in mild cognitive impairment. *Adv Sci*. 2022;9(12):e2104538. https://doi.org/10.1002/advs.202104538.
15 Lian C, Liu M, Wang L, Shen D. Multi-task weakly-supervised attention network for dementia status estimation with structural

MRI. *IEEE Trans Neural Netw Learn Syst*. 2022;33(8):4056–4068. https://doi.org/10.1109/TNNLS.2021.3055772.
16 Lian C, Liu M, Zhang J, Shen D. Hierarchical fully convolutional network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI. *IEEE Trans Pattern Anal Mach Intell*. 2020;42(4):880–893. https://doi.org/10.1109/TPAMI.2018.2889096.
17 Pan Y, Liu M, Xia Y, Shen D. Disease-image-specific learning for diagnosis-oriented neuroimage synthesis with incomplete multi-modality data. *IEEE Trans Pattern Anal Mach Intell*. 2022;44(10):6839–6853. https://doi.org/10.1109/TPAMI.2021.3091214.
18 Perlis RH. Translating biomarkers to clinical practice. *Mol Psychiatry*. 2011;16(11):1076–1087. https://doi.org/10.1038/mp.2011.63.
19 Planche V, Bouteloup V, Pellegrin I, et al. Validity and performance of blood biomarkers for Alzheimer disease to predict dementia risk in a large clinic-based cohort. *Neurology*. 2023;100(5):e473–e484. https://doi.org/10.1212/WNL.0000000000201479.
20 Marek S, Tervo-Clemmens B, Calabro FJ, et al. Reproducible brain-wide association studies require thousands of individuals. *Nature*. 2022;603(7902):654–660. https://doi.org/10.1038/s41586-022-04492-9.
21 Wang J, Knol MJ, Tiulpin A, et al. Gray matter age prediction as a biomarker for risk of dementia. *Proc Natl Acad Sci U S A*. 2019;116(42):21213–21218. https://doi.org/10.1073/pnas.1902376116.
22 Pichet Binette A, Gonneaud J, Vogel JW, et al. Morphometric network differences in ageing versus Alzheimer's disease dementia. *Brain*. 2020;143(2):635–649. https://doi.org/10.1093/brain/awz414.
23 Dubois B, Feldman HH, Jacova C, et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurol*. 2007;6(8):734–746. https://doi.org/10.1016/S1474-4422(07)70178-3.
24 Fan L, Li H, Zhuo J, et al. The human brainnetome atlas: a new brain atlas based on connectional architecture. *Cereb Cortex*. 2016;26(8):3508–3526. https://doi.org/10.1093/cercor/bhw157.
25 Zhao K, Zheng Q, Che T, et al. Regional radiomics similarity networks (R2SNs) in the human brain: reproducibility, small-world properties and a biological basis. *Netw Neurosci*. 2021;5(3):783–797. https://doi.org/10.1162/netn_a_00200.
26 Tong Q, He H, Gong T, et al. Multicenter dataset of multi-shell diffusion MRI in healthy traveling adults with identical settings. *Sci Data*. 2020;7(1):157. https://doi.org/10.1038/s41597-020-0493-8.
27 Tanaka SC, Yamashita A, Yahata N, et al. A multi-site, multi-disorder resting-state magnetic resonance image database. *Sci Data*. 2021;8(1):227. https://doi.org/10.1038/s41597-021-01004-8.
28 Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155–163. https://doi.org/10.1016/j.jcm.2016.02.012.
29 He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *Presented at: 2016 IEEE conference on computer vision and pattern recognition (CVPR); 2016*. 2016. https://doi.org/10.1109/cvpr.2016.90.
30 Jin D, Zhou B, Han Y, et al. Generalizable, reproducible, and neuroscientifically interpretable imaging biomarkers for Alzheimer's disease. *Adv Sci*. 2020;7(14):2000675. https://doi.org/10.1002/advs.202000675.
31 Dosovitskiy A, Beyer L, Kolesnikov A, et al. *An image is worth 16x16 words: transformers for image recognition at scale. arXiv pre-print server*. 2020. https://doi.org/10.48550/arXiv.2010.11929.
32 Liu C, Chyr J, Zhao W, et al. Genome-wide association and mechanistic studies indicate that immune response contributes to Alzheimer's disease development. *Front Genet*. 2018;9:410. https://doi.org/10.3389/fgene.2018.00410.
33 Zhou Y, Zhou B, Pache L, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. 2019;10(1):1523. https://doi.org/10.1038/s41467-019-09234-6.
34 He L, Yu H, Shi L, et al. Equity assessment of the distribution of CT and MRI scanners in China: a panel data analysis. *Int J Equity Health*. 2018;17(1):157. https://doi.org/10.1186/s12939-018-0869-y.
35 Califf RM. Biomarker definitions and their applications. *Exp Biol Med*. 2018;243(3):213–221. https://doi.org/10.1177/1535370217750088.
36 GBD 2016 Dementia Collaborators. Global, regional, and national burden of Alzheimer's disease and other dementias, 1990-2016: a systematic analysis for the global burden of disease study 2016. *Lancet Neurol*. 2019;18(1):88–106. https://doi.org/10.1016/S1474-4422(18)30403-4.
37 Dubois B, Villain N, Frisoni GB, et al. Clinical diagnosis of Alzheimer's disease: recommendations of the International Working Group. *Lancet Neurol*. 2021;20(6):484–496. https://doi.org/10.1016/S1474-4422(21)00066-1.
38 Jack CR Jr, Holtzman DM. Biomarker modeling of Alzheimer's disease. *Neuron*. 2013;80(6):1347–1358. https://doi.org/10.1016/j.neuron.2013.12.003.
39 Heneka MT, Carson MJ, El Khoury J, et al. Neuroinflammation in Alzheimer's disease. *Lancet Neurol*. 2015;14(4):388–405. https://doi.org/10.1016/S1474-4422(15)70016-5.
40 Leng F, Edison P. Neuroinflammation and microglial activation in Alzheimer disease: where do we go from here? *Nat Rev Neurol*. 2021;17(3):157–172. https://doi.org/10.1038/s41582-020-00435-y.
41 Brinkmalm G, Zetterberg H. The phosphorylation cascade hypothesis of Alzheimer's disease. *Nat Aging*. 2021;1(6):498–499. https://doi.org/10.1038/s43587-021-00077-9.
42 Zhao J, Fu Y, Yamazaki Y, et al. APOE4 exacerbates synapse loss and neurodegeneration in Alzheimer's disease patient iPSC-derived cerebral organoids. *Nat Commun*. 2020;11(1):5540. https://doi.org/10.1038/s41467-020-19264-0.
43 Lombardi J, Mayer B, Semler E, et al. Quantifying progression in primary progressive aphasia with structural neuroimaging. *Alzheimers Dement*. 2021;17(10):1595–1609. https://doi.org/10.1002/alz.12323.
44 Harms M, Benitez BA, Cairns N, et al. C9orf72 hexanucleotide repeat expansions in clinical Alzheimer disease. *JAMA Neurol*. 2013;70(6):736–741. https://doi.org/10.1001/2013.jamaneurol.537.
45 Jones D, Lowe V, Graff-Radford J, et al. A computational model of neurodegeneration in Alzheimer's disease. *Nat Commun*. 2022;13(1):1643. https://doi.org/10.1038/s41467-022-29047-4.
46 Qiu S, Miller MI, Joshi PS, et al. Multimodal deep learning for Alzheimer's disease dementia assessment. *Nat Commun*. 2022;13(1):3404. https://doi.org/10.1038/s41467-022-31037-5.
47 Hansson O, Edelmayer RM, Boxer AL, et al. The Alzheimer's Association appropriate use recommendations for blood biomarkers in Alzheimer's disease. *Alzheimers Dement*. 2022;18(12):2669–2686. https://doi.org/10.1002/alz.12756.
48 West T, Kirmess KM, Meyer MR, et al. A blood-based diagnostic test incorporating plasma Abeta42/40 ratio, ApoE proteotype, and age accurately identifies brain amyloid status: findings from a multi cohort validity analysis. *Mol Neurodegen*. 2021;16(1):30. https://doi.org/10.1186/s13024-021-00451-6.
49 Janelidze S, Bali D, Ashton NJ, et al. Head-to-head comparison of 10 plasma phospho-tau assays in prodromal Alzheimer's disease. *Brain*. 2023;146(4):1592–1601. https://doi.org/10.1093/brain/awac333.
50 Yakoub Y, Ashton NJ, Strikwerda-Brown C, et al. Longitudinal blood biomarker trajectories in preclinical Alzheimer's disease. *Alzheimers Dement*. 2023. https://doi.org/10.1002/alz.13318.
51 Cullen NC, Leuzy A, Palmqvist S, et al. Individualized prognosis of cognitive decline and dementia in mild cognitive impairment based on plasma biomarker combinations. *Nat Aging*. 2021;1(1):114–123. https://doi.org/10.1038/s43587-020-00003-5.
52 Mattsson-Carlgren N, Salvado G, Ashton NJ, et al. Prediction of longitudinal cognitive decline in preclinical Alzheimer disease using plasma biomarkers. *JAMA Neurol*. 2023;80(4):360–369. https://doi.org/10.1001/jamaneurol.2022.5272.
53 Leuzy A, Smith R, Cullen NC, et al. Biomarker-based prediction of longitudinal tau positron emission tomography in Alzheimer disease. *JAMA Neurol*. 2022;79(2):149–158. https://doi.org/10.1001/jamaneurol.2021.4654.