



## OPEN ACCESS

## EDITED BY

Wencai Liu,  
Shanghai Jiao Tong University, China

## REVIEWED BY

Kai Xu,  
Shandong First Medical University, China  
Lizhihan Yu,  
Texas A and M University, United States

## \*CORRESPONDENCE

Dantao Peng  
✉ pengdantao2000@163.com

RECEIVED 19 October 2024

ACCEPTED 05 December 2024

PUBLISHED 19 December 2024

## CITATION

Zhen W, Wang Y, Zhen H, Zhang W, Shao W, Sun Y, Qiao Y, Jia S, Zhou Z, Wang Y, Chen L, Zhang J and Peng D (2024) Association between Alzheimer's disease pathologic products and age and a pathologic product-based diagnostic model for Alzheimer's disease. *Front. Aging Neurosci.* 16:1513930. doi: 10.3389/fnagi.2024.1513930

## COPYRIGHT

© 2024 Zhen, Wang, Zhen, Zhang, Shao, Sun, Qiao, Jia, Zhou, Wang, Chen, Zhang and Peng. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Association between Alzheimer's disease pathologic products and age and a pathologic product-based diagnostic model for Alzheimer's disease

Weizhe Zhen<sup>1</sup>, Yu Wang<sup>2</sup>, Hongjun Zhen<sup>3</sup>, Weihe Zhang<sup>2</sup>, Wen Shao<sup>2</sup>, Yu Sun<sup>2</sup>, Yanan Qiao<sup>2</sup>, Shuhong Jia<sup>2</sup>, Zhi Zhou<sup>2</sup>, Yuye Wang<sup>2</sup>, Leian Chen<sup>2</sup>, Jiali Zhang<sup>1</sup> and Dantao Peng<sup>1,2\*</sup>

<sup>1</sup>Graduate School, Beijing University of Chinese Medicine, Beijing, China, <sup>2</sup>Department of Neurology, China-Japan Friendship Hospital, Beijing, China, <sup>3</sup>Department of Orthopedics, Handan Chinese Medicine Hospital, Handan, Hebei, China

**Background:** Alzheimer's disease (AD) has a major negative impact on people's quality of life, life, and health. More research is needed to determine the relationship between age and the pathologic products associated with AD. Meanwhile, the construction of an early diagnostic model of AD, which is mainly characterized by pathological products, is very important for the diagnosis and treatment of AD.

**Method:** We collected clinical study data from September 2005 to August 2024 from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. Using correlation analysis method like cor function, we analyzed the pathology products (t-Tau, p-Tau, and A $\beta$  proteins), age, gender, and Minimum Mental State Examination (MMSE) scores in the ADNI data. Next, we investigated the relationship between pathologic products and age in the AD and non-AD groups using linear regression. Ultimately, we used these features to build a diagnostic model for AD.

**Results:** A total of 1,255 individuals were included in the study (mean [SD] age, 73.27 [7.26] years; 691male [55.1%]; 564 female [44.9%]). The results of the correlation analysis showed that the correlations between pathologic products and age were, in descending order, Tau (Corr=0.75), p-Tau (Corr=0.71), and A $\beta$  (Corr=0.54). In the AD group, t-Tau protein showed a tendency to decrease with age, but it was not statistically significant. p-Tau protein levels similarly decreased with age and its decrease was statistically significant. In contrast to Tau protein, in the AD group, A $\beta$  levels increased progressively with age. In the non-AD group, the trend of pathologic product levels with age was consistently opposite to that of the AD group. We finally screened the optimal AD diagnostic model (AUC=0.959) based on the results of correlation analysis and by using the Xgboost algorithm and SVM algorithm.

**Conclusion:** In a novel finding, we observed that Tau protein and A $\beta$  had opposite trends with age in both the AD and non-AD groups. The linear regression curves of the AD and non-AD groups had completely opposite trends. Through a machine learning approach, we constructed an AD diagnostic model with excellent performance based on the selected features.

## KEYWORDS

Alzheimer's disease, pathologic product, machine learning, age, diagnostic model, linear regression

## 1 Introduction

Alzheimer's disease (AD) is a neurodegenerative disease that seriously jeopardizes human health and affects patients' quality of life (2023, 2024). It is the number one cause of dementia and mainly affects the middle-aged and elderly population (2022; Scheltens et al., 2021). Before developing Alzheimer's disease, patients will go through the stages of subjective memory complaints (SMC), mild cognitive decline and so on. How to diagnose Alzheimer's disease more accurately and distinguish it from the preclinical stage of Alzheimer's disease as well as normal people has been a hot topic of research. Previous researchers have attempted to differentiate AD using biomarkers, imaging, and some behavioral-based kinesiology tests, among others, with biomarker research undoubtedly receiving the most attention (Bai et al., 2021; Winchester et al., 2023; Küçükali et al., 2023; Yang et al., 2020). A large number of biomarkers have been detected in blood, cerebrospinal fluid tests, etc., which have a good ability to differentiate between patients with AD (Izzo et al., 2021; Kumari et al., 2022). The researchers even spent a great deal of time studying longitudinal changes in these biomarkers, monitoring changes in their levels throughout the course of the onset of Alzheimer's disease (Jia et al., 2024; Yakoub et al., 2023). And with the continuous advancement of histologic research techniques, more and more biomarkers are being discovered in a higher throughput manner. Our team has previously uncovered a very large number of AD biomarkers through both blood and urine testing methods, using histology-related techniques, and has built an AD diagnostic model based on them (Wang et al., 2023a,b).

$\beta$ -amyloid (A $\beta$ ) and Tau proteins are the focus of biomarker research as recognized markers of AD pathology. Many of both protein families have been found to be closely associated with the onset and progression of AD (Ferrari-Souza et al., 2022; Ashton et al., 2021; Horie et al., 2023). Previous studies have focused on the particular significance of these two pathologic products in the molecular mechanisms underlying the developmental process of Alzheimer's disease (Zhang H. et al., 2021; Busche and Hyman, 2020). In contrast, the association between these two pathologic products and age in the preclinical and onset stages of AD has lacked elucidation in large-sample studies (Stern et al., 2023).

In terms of research on diagnostic models for AD, there are many studies that use biomarkers as features, and not a few of them incorporate A $\beta$  and Tau protein (Ferreiro et al., 2023). However, in previous studies, A $\beta$  and Tau protein were hardly used as core features, and the inclusion of other biomarkers mixed the significance of the two in modeling. At the same time, previous studies also suffered from the shortcomings of using algorithms mainly focusing on regression algorithms, a single type of algorithm and a lack of sample size (Hammond et al., 2020; Gao et al., 2023). We used classification algorithms with excellent performance in this study to construct and train an AD diagnostic model using A $\beta$  and Tau protein as the core features.

## 2 Methods

### 2.1 Design

The Alzheimer's Disease Neuroimaging Initiative (ADNI) database<sup>1</sup> provided the data used in this investigation, which were gathered between September 2005 and August 2024. Established in 2003, the ADNI program is a research endeavor with the goal of examining the course of Alzheimer's disease and its preclinical phases through the use of MRI, PET, biomarkers, and clinical and neuropsychological examinations. All participating institutions' Institutional Review Boards have given their approval for the ADNI trial. All participants, or their authorized representatives, have given written informed permission to ADNI in compliance with the Declaration of Helsinki.

### 2.2 Participants

After undergoing a battery of cognitive functioning tests, each individual was assigned to one of four groups: AD, mild cognitive impairment (MCI), SMC, or control (CN). The Mini-Mental State Examination (MMSE) scores for AD were 20–26, while for CN, SMC, and MCI, they were 24–30. For CN, MCI, and AD, the Clinical Dementia Rating (CDR) was 0.5, 0.5, and  $\geq 0.5$ . For the various ADNI cohorts, the enrollment processes and inclusion criteria were generally the same. Previous descriptions have been made of the specific enrollment processes and inclusion criteria for the various diagnostic categories of the ADNI cohort (Petersen et al., 2010). You could find the ADNI database protocol<sup>2</sup> with detailed inclusion and exclusion requirements. In order to evaluate cognitive function, we used the MMSE.

### 2.3 Biomarker collection and analysis

CSF Tau protein and A $\beta$  protein data from the ADNI database were used in our study. Methods of CSF collection and biomarker measurement have been previously reported (Hampel et al., 2010; Shaw et al., 2009). The ADNI database did not report outliers. However, biomarker assays have detection intervals. The upper limit of detection for A $\beta$  is 1,700 pg/mL and the lower limit is 200 pg/mL. The upper limit of detection for t-Tau is 1,300 pg/mL and the lower limit is 80 pg/mL. The upper limit of detection for p-Tau is 120 pg/mL and the lower limit is 8 pg/mL. The detections that exceeded the detection interval the most accounted for <15% of the total data, and only a very small number of the other detections exceeded the detection interval. For data exceeding the detection range of Tau or A $\beta$  proteins, we took half of the lower detection limit value to replace data below the lower detection limit, and used the upper detection limit value to replace data above the upper detection limit.

<sup>1</sup> <http://adni.loni.usc.edu>

<sup>2</sup> <https://adni.loni.usc.edu/help-faqs/adni-documentation/>

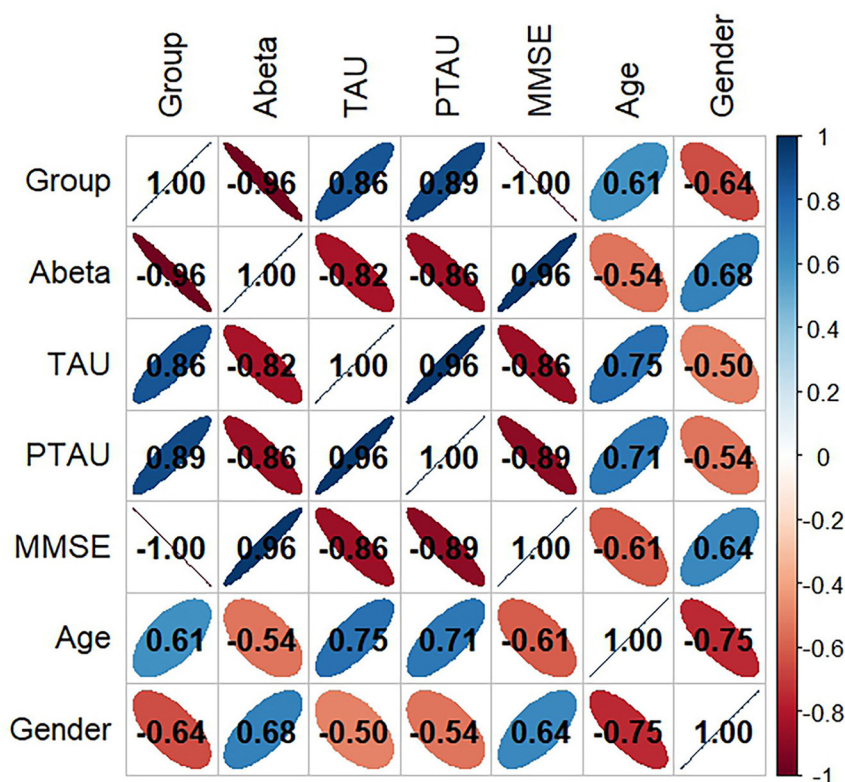


FIGURE 1

Variable correlation analysis of all features. A $\beta$ ,  $\beta$ -amyloid; t-Tau, total-Tau; p-Tau, phosphorylated tau; MMSE, Minimum Mental State Examination.

## 2.4 Model constructing and training

We employed two machine learning algorithms, support vector machine (SVM) and extreme gradient boosting (XGBoost), to build AD diagnostic models. Cross-validation is used to evaluate the performance of machine learning models as well as for hyperparameter tuning. Ten-fold cross-validation was applied to the training and validation sets in order to reduce overfitting and enhance the model's functionality. We tuned the hyperparameters based on the results of the ten-fold cross-validation to get the best performing model. The best model for this study was determined by looking at the Receiver Operating Characteristic (ROC) curve and selecting the model with the highest area under the curve (AUC).

## 2.5 Statistical analysis

R software (version 4.3.1) and IBM SPSS Statistics for Windows version 27.0 were used to conduct all statistical tests. We employed nonparametric tests to compare the Non-AD (include CN, SMC, EMCI, and LMCI) and AD groups for variables like total-Tau (t-Tau), phosphorylated tau (p-Tau), A $\beta$ , age, and MMSE that did not match the requirements of analysis of variance (ANOVA). The chi-square test was used for statistical analysis of counts like gender. The threshold for a difference to be deemed statistically significant was  $p < 0.05$ . Since the data are derived from public databases, the occurrence of missing values is often unavoidable.

In this study, the proportion of missing values to the total data has been well over 50%. In order to minimize the error in the study, we used direct culling of missing values in the data instead of using interpolation. In contrast, this is the optimal way to ensure data integrity and accuracy.

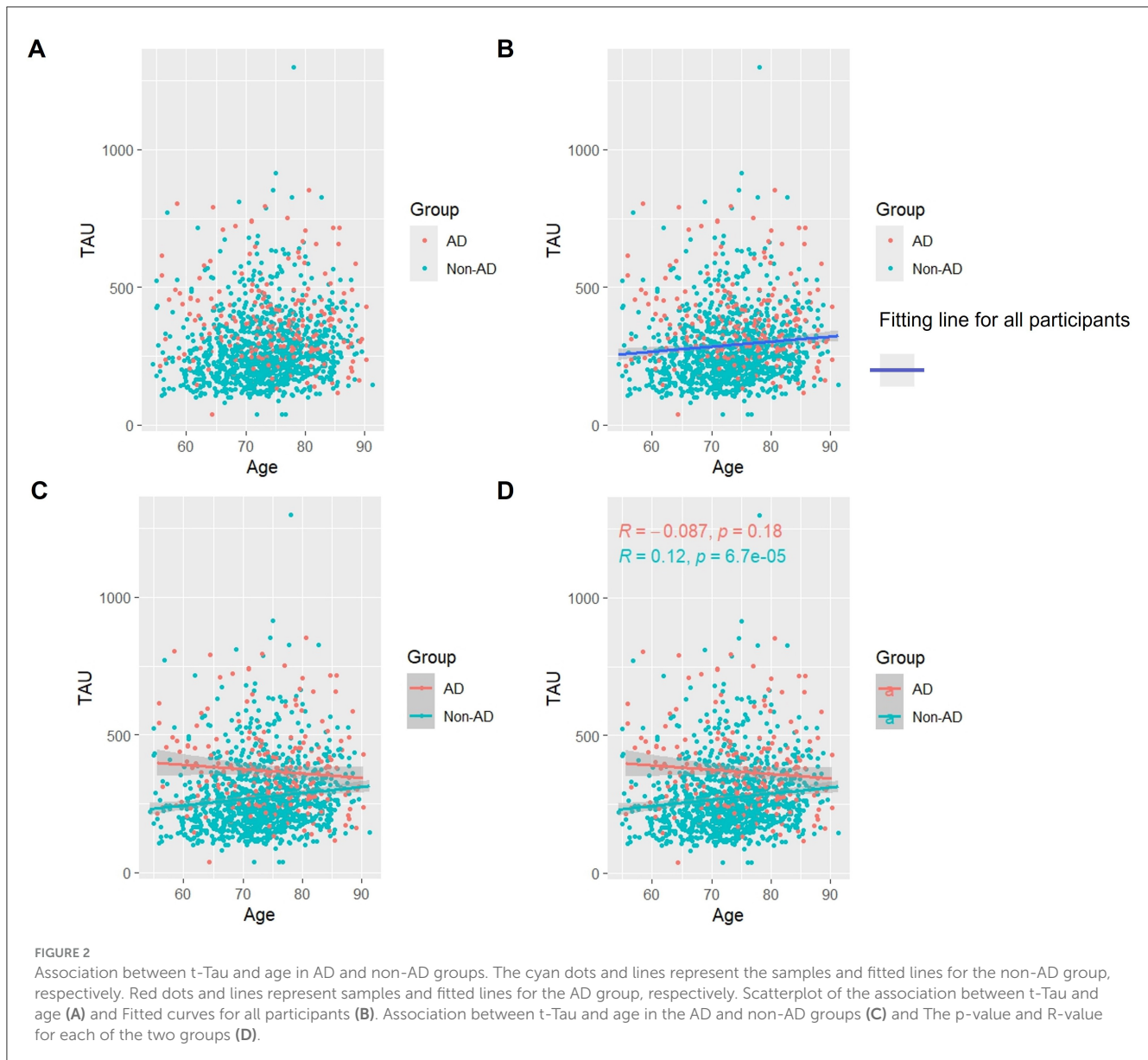
One popular technique for determining how closely variables correlate with one another is correlation analysis. To determine the correlation coefficient and compute the correlation of variables, we utilize the cor function found in R-4.3.1. The correlation between the variables is higher the closer the correlation coefficient's absolute value is to 1.

In several contexts, the relationship between diseased products and age was estimated using linear regression analysis. Initially, we examined the relationship between pathogenic products (include A $\beta$ , t-Tau protein, and p-Tau protein) and age in the AD and non-AD groups. We also investigated the relationship between p-Tau/t-Tau and A $\beta$ /t-Tau and age using linear regression in an effort to better understand the relationship between these pathogenic products and age.

## 3 Results

### 3.1 Demographic and clinical characteristics of patients

A total of 1,255 individuals were included in the study (mean [SD] age, 73.27 [7.26] years; 691 male [55.1%]; 564 female [44.9%]).



Six factors in all were examined: pathogenic products ( $A\beta$ , t-Tau, and p-Tau), age, sex, and MMSE scores. The average score for the MMSE test was 26.96 [3.18], the average score for  $A\beta$  was 966.19 [458.57], the average score for t-Tau was 290.57 [136.04], and the average score for p-Tau was 27.96 [14.91] for every subject. There were statistically significant differences in every attribute between the groups (Table 1).

## 3.2 Correlation analysis between variables

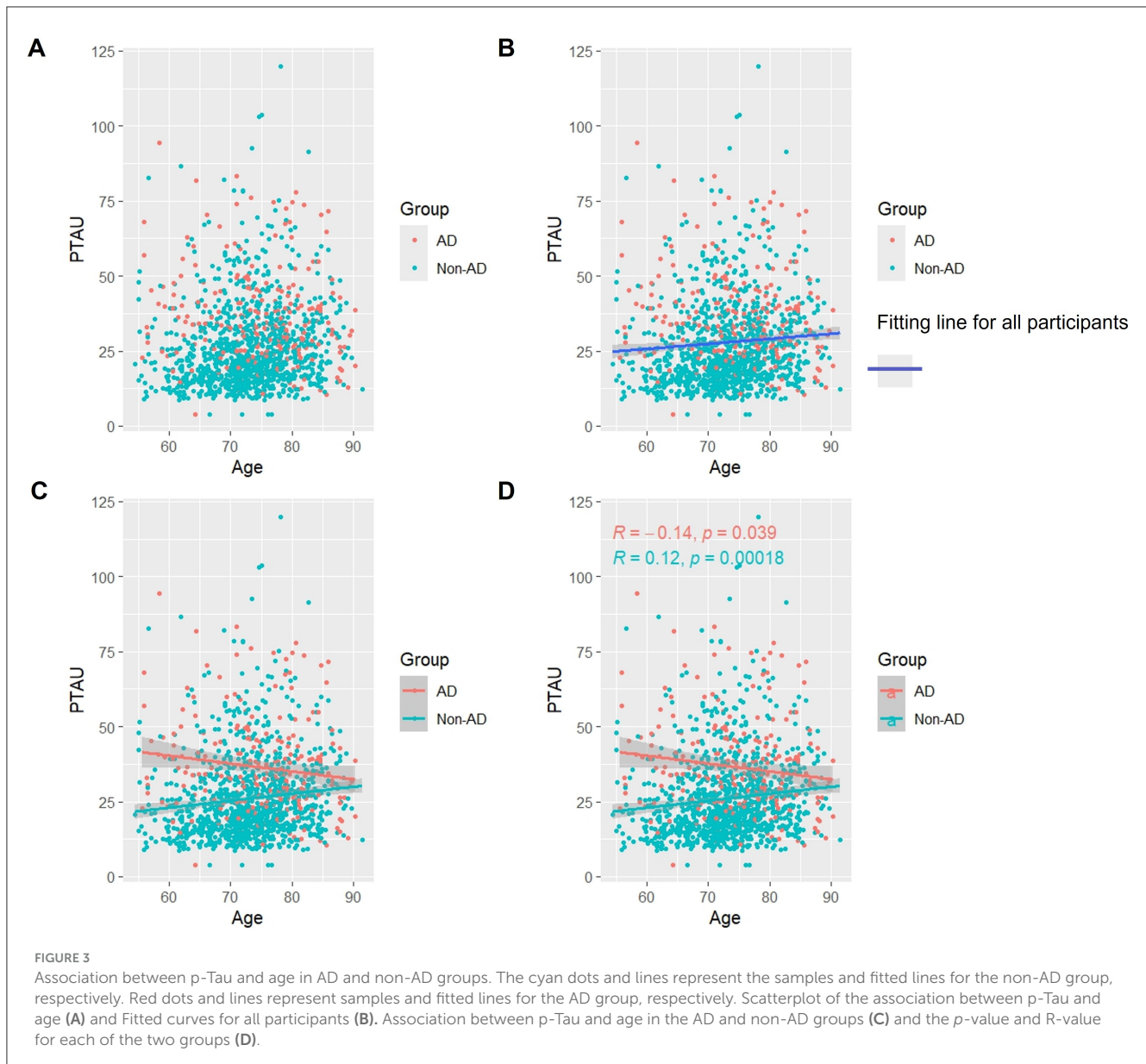
Figure 1 displays the findings of the six factors' correlation study with disease type. where the Corr values are between  $-1.0$  and  $1.0$ ; the higher the correlation between the variables, the closer the Corr value's absolute value is to 1. Conversely, a correlation is less the closer it is to 0. The pathogenic products and age had the following associations, in decreasing order of absolute Corr values: t-Tau (0.75), p-Tau (0.71), and  $A\beta$  (0.54).

## 3.3 Association of Tau proteins with age

As Figure 2 shows, when we studied all subjects, we found that overall the level of Tau protein increased with age (Figures 2A, B). Interestingly, in the AD group, Tau protein appeared to slowly decrease with age, although this decrease was not statistically significant ( $p > 0.05$ ). In the Non-AD group, the longitudinal rise in Tau protein was similarly associated with a lateral increase in age ( $p < 0.05$ ) (Figures 2C, D).

## 3.4 Association of p-Tau protein with age

As Figure 3 shows, when we studied all subjects, we found that, overall, the levels of p-Tau protein increased with age (Figures 3A, B). Like t-Tau protein, in the AD group, p-Tau protein appeared to slowly decrease with



age. However, the difference was that the decrease in p-Tau protein was statistically significant ( $p < 0.05$ ). In the non-AD group, the longitudinal rise in p-Tau protein was similarly associated with a lateral increase in age ( $p < 0.05$ ) (Figures 3C, D).

### 3.5 Association of p-Tau/t-Tau with age

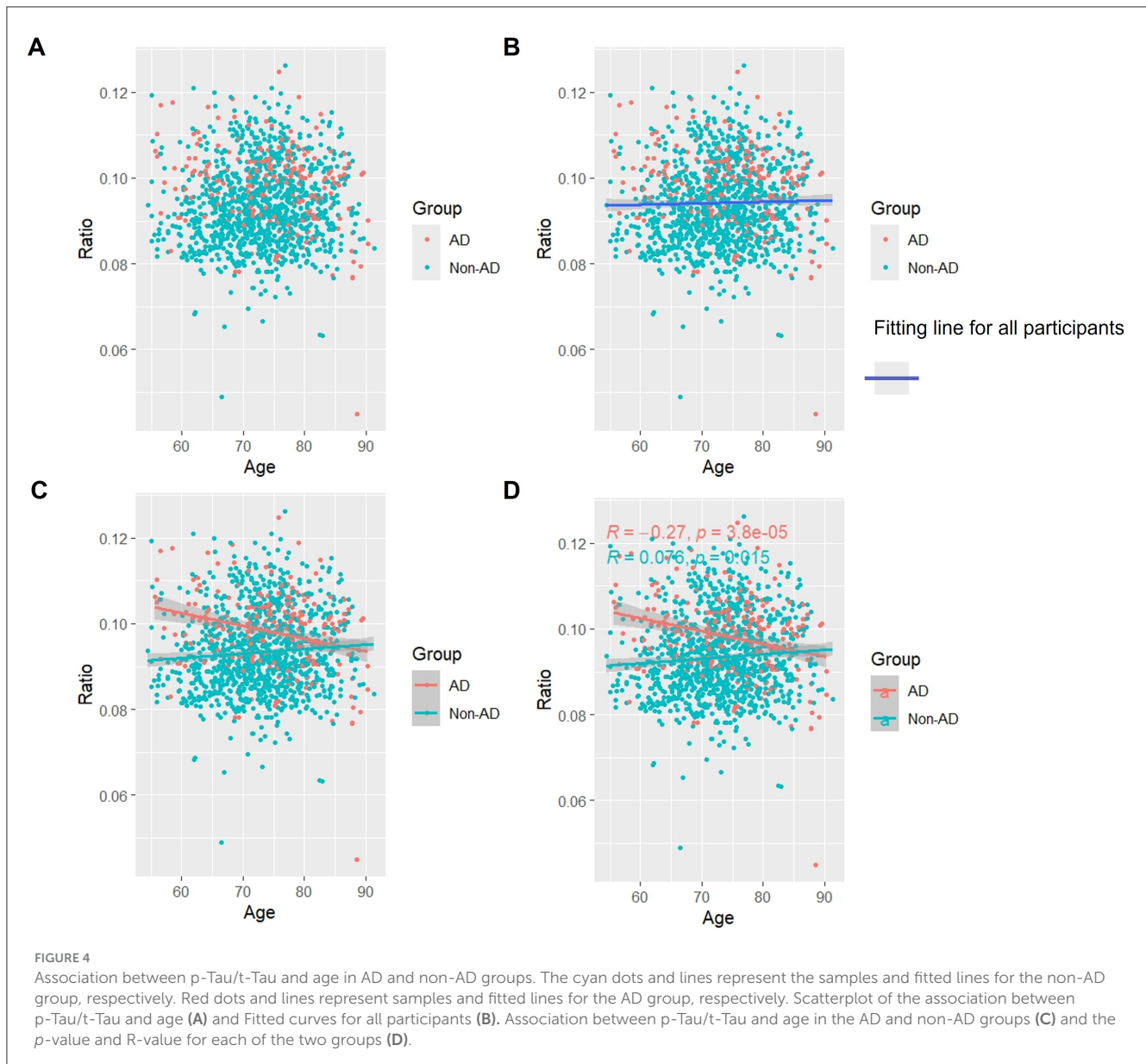
Overall, as seen in Figures 4A, B, there was a slight but steady tendency for p-Tau/t-Tau levels to rise with aging. p-Tau/t-Tau significantly decreased in the AD group as age increased, and this decline was statistically distinct ( $p < 0.05$ ). On the other hand, p-Tau/t-Tau increased with age ( $p < 0.05$ ) in the non-AD group (Figures 4C, D).

### 3.6 Association of A $\beta$ with age

Overall, as Figure 5 illustrates, A $\beta$  levels steadily declined with age, which was different from t-Tau and p-Tau (Figures 5A, B). A $\beta$  levels significantly increased with age in the AD group, which was likewise in contrast to t-Tau and p-Tau ( $p < 0.05$ ). Conversely, A $\beta$  levels in the non-AD group dropped with age ( $p < 0.05$ ) (Figures 5C, D).

### 3.7 Association of A $\beta$ /t-Tau with age

As shown in Figure 6, as a whole, A $\beta$ /t-Tau levels gradually decreased with age, which is consistent with the change of A $\beta$  with age (Figures 6A, B). In the AD group, A $\beta$ /t-Tau levels increased significantly with age ( $p < 0.05$ ). While in the non-AD group, A $\beta$ /t-Tau instead decreased with age ( $p < 0.05$ ) (Figures 6C, D).

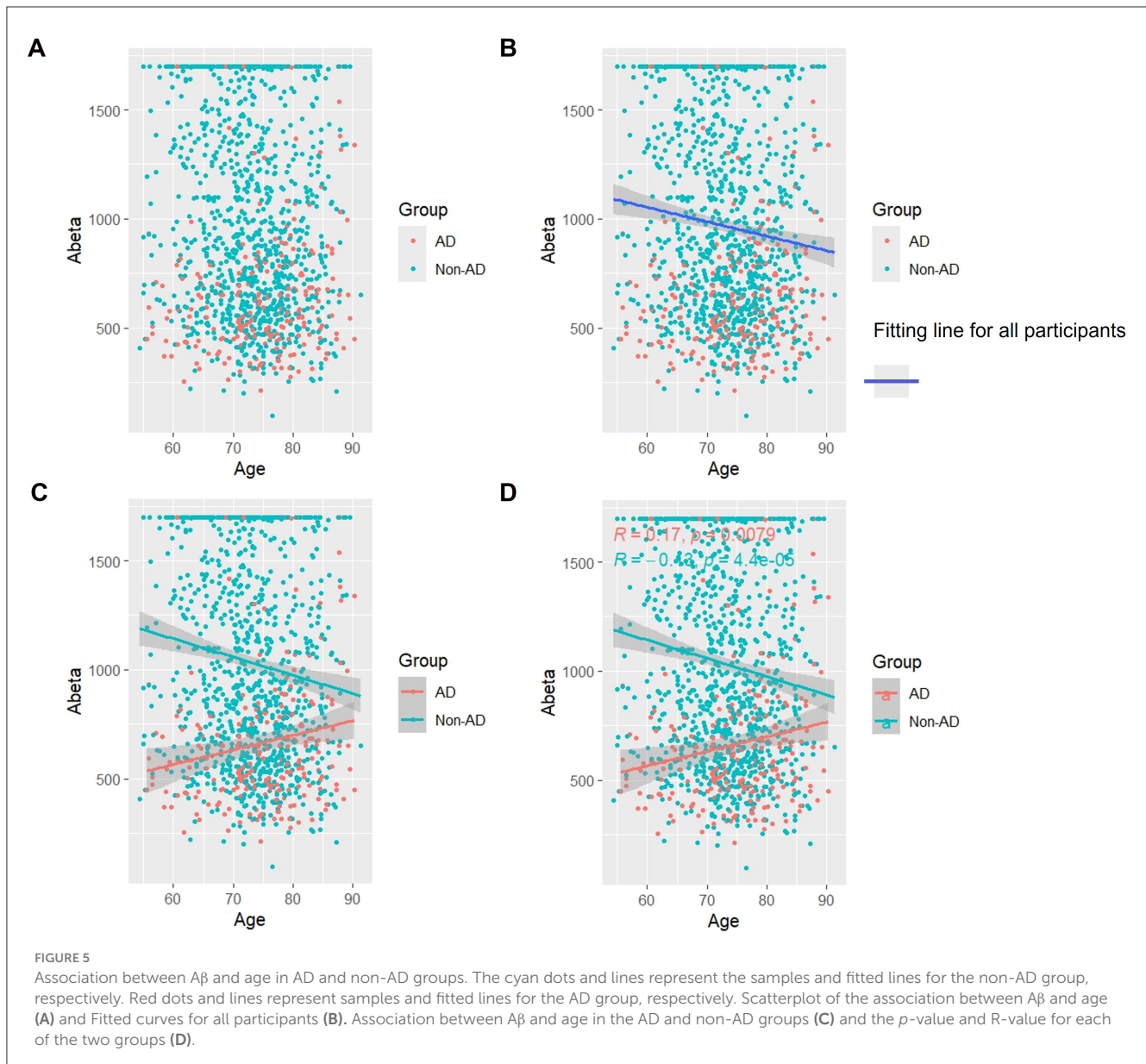


### 3.8 Construction and optimization of AD diagnostic models

We built two machine learning models for diagnosing AD using the XGBoost classifier and the SVM classifier, respectively, based on the six previously mentioned features. To avoid overfitting, we further enhanced the model performance via ten-fold cross-validation. Among them, the classifier model based on the XGBoost algorithm has superior performance (AUC of 0.959), accuracy of 0.69, specificity of 0.86, and sensitivity of 0.95. The classifier model based on the support vector machine (SVM) algorithm has an AUC of 0.924, accuracy of 0.90, sensitivity of 0.96, and Specificity of 0.66 (Figure 7).

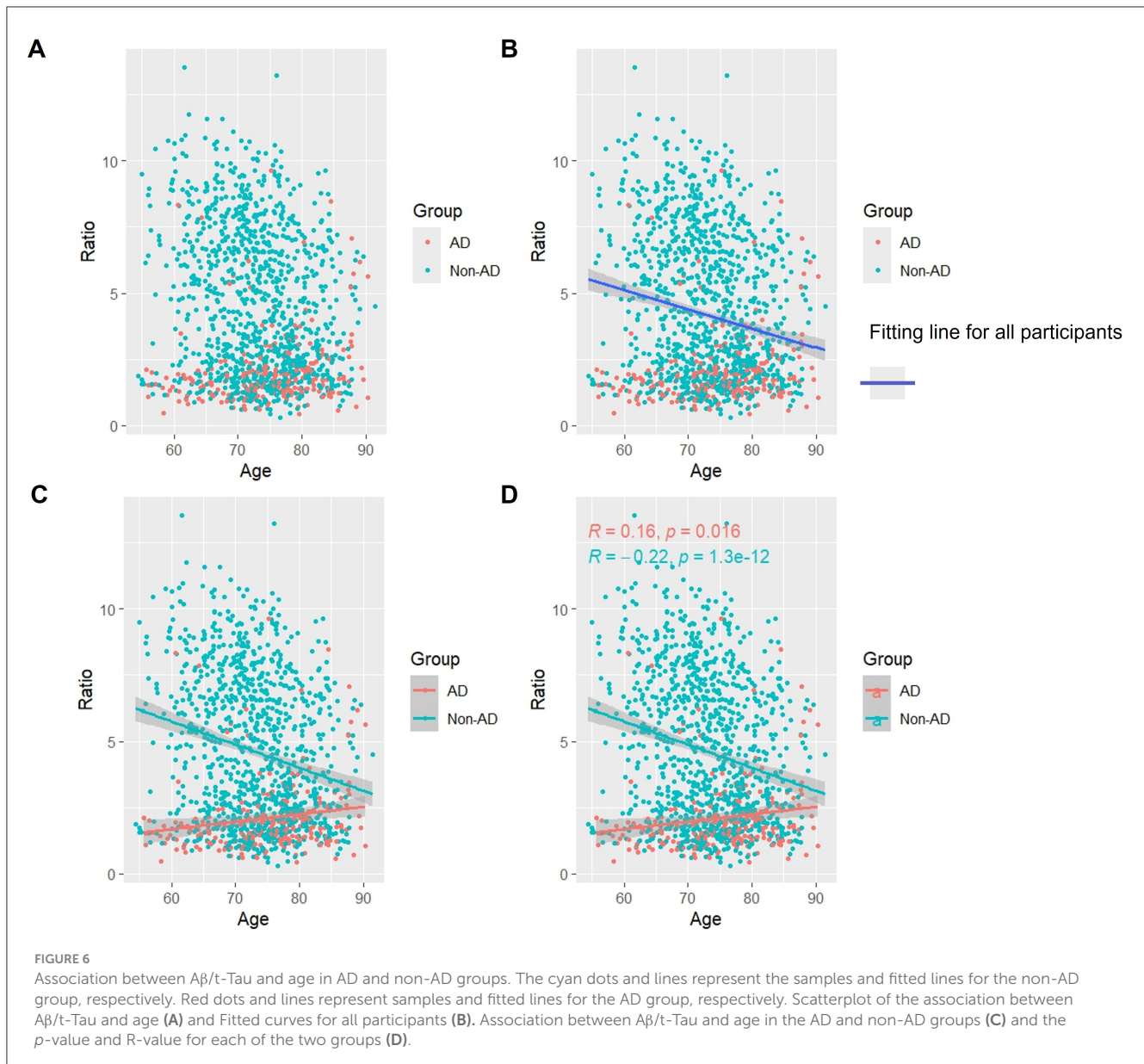
## 4 Discussion

Our study systematically analyzed the relationship between changes in intracranial t-Tau protein, p-Tau protein, and A $\beta$  protein levels and age in AD patients and Non-AD population through data mining and analysis of the ADNI database. The results of the study were very interesting. We found that for both t-Tau protein and A $\beta$  protein, the trends with age were diametrically opposed in AD patients and Non-AD populations. For both t-Tau protein and p-Tau protein, the levels of these pathogens progressively decreased the older the AD patient was. The difference, however, was that t-Tau protein showed a decreasing trend but was not statistically different compared to



the statistically significant decrease in p-Tau protein. As for A $\beta$  protein, the older the age of AD patients, the A $\beta$  level was continuously increasing. In contrast, A $\beta$  levels in the Non-AD population gradually declined with age. Previous studies on the correlation between these pathologic products and patient age are lacking and not clearly recognized or elaborated. However, longitudinal changes in these pathologic products over time at the individual level have been examined in previous studies, which is different from our observation of the association between pathologic products and age at the population level (Barthélemy et al., 2020). Additionally, it is not clear that some studies have focused on changes in the levels of pathologic products long before the onset of Alzheimer's disease, and have not examined changes after the onset of the disease (Jia et al., 2024). There are also studies that do not distinguish between studies of AD patients and Non-AD groups (Cogswell et al., 2024).

One of the main reasons we chose A $\beta$  and Tau proteins for our study is that it has a very important impact in the course and mechanisms of AD (Pang et al., 2022; Sadleir and Vassar, 2023). Abnormal aggregation of A $\beta$  and Tau proteins is an important pathogenesis and pathological hallmark of AD. Questions about how the two are produced and how their levels change during disease progression have been an important issue affecting our understanding of AD, as well as a focus and difficulty in research. Previous findings suggest that  $\delta$ -secretase cleaved Tau proteins may stimulate A $\beta$  production by upregulating STAT1-BACE1 signaling in AD patients (Zhang Z. et al., 2021). This is a rather important finding. It not only reveals the molecular regulatory process between the two pathologic products. More importantly, it suggests to us that the molecular regulatory process that exists between the two may allow the levels of these pathology products to be dynamically regulated during the course of the disease. This



is likely to be an important mechanism for the pathogenesis and disease progression of AD. Not only that, Aβ pathology may induce changes in soluble tau release and phosphorylation (Mattsson-Carlgren et al., 2020). These findings show that the relationship between Aβ and t-Tau and p-Tau is bi-directionally regulated and mutually restrained. Fluctuations in the levels of the two pathogens are influenced by each other. This also gives the ratio of the two pathologic products more value for clinical studies. In turn, abnormal aggregation of the two could ultimately drive the disease progression by leading to loss of synapses, affecting synaptic function and thus disrupting memory formation (Li et al., 2018). While the roles of Aβ and Tau in the pathogenesis of AD continue to be elucidated, researchers are monitoring changes in their levels, further exploring their potential as biomarkers of the disease, and even screening for other AD biomarkers with good predictive ability based on them (Chiu et al., 2021; Boza-Serrano et al., 2022). As the technology associated with

biomedical engineering continues to advance, more assays have been developed for the detection of Aβ and Tau. In addition to furthering our understanding of the pathologic processes of AD, we are discovering more pathologic processes associated with the aggregation of these pathogens. These are also one of the hotspots and directions for future research (Pichet Binette et al., 2021).

We constructed and trained a machine learning model using pathology products such as Aβ and Tau proteins as core features. In terms of the model's performance, its ability to distinguish between AD and non-AD is excellent. Previous studies have rarely focused only on the contribution of Tau and Aβ proteins to constructing AD diagnostic models, often incorporating some other biomarkers (Gaetani et al., 2021; Ficiarà et al., 2021; Franciotti et al., 2023; Khan et al., 2024). These studies have their own innovations and strengths, but inevitably have some shortcomings. The shortcomings mainly lie in the lack of performance of the diagnostic model, the small sample size, the excessive number of



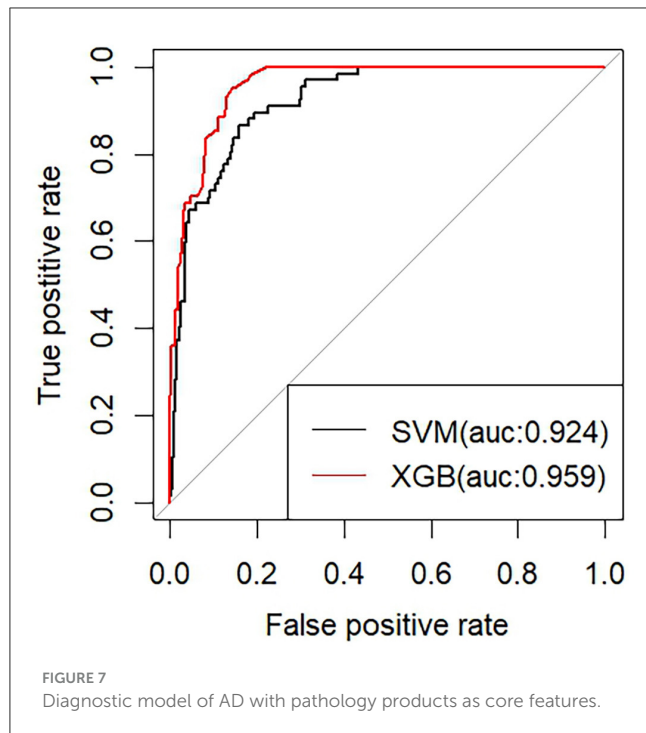


TABLE 1 Baseline demographics and clinical characteristics.

Characteristics	Total (1,255)	Non-AD (1,022)	AD (233)	P value
Age (years)	73.27	72.94	74.71	<0.01
<b>Gender</b>				
Male (%)	55.1	54.2	58.8	0.204
Female (%)	44.9	45.8	41.2	0.204
MMSE	26.96	27.88	22.94	<0.01
A $\beta$ (pg/mL)	966.19	1,034.97	664.48	<0.01
t-Tau (pg/mL)	290.57	272.79	368.51	<0.01
p-Tau (pg/mL)	27.96	26.01	36.51	<0.01

A $\beta$ ,  $\beta$ -amyloid; t-Tau, total-Tau; p-Tau, phosphorylated tau; MMSE, Minimum Mental State Examination.

incorporated features, and so on. The performance of the AD diagnostic model constructed in our study is relatively superior. The relative singularity of the incorporated features can highlight more the importance of the pathology products in the model construction. Our choice of algorithms for machine learning that performs well in dealing with classification problems is an important guarantee of the sophistication of our study (Li J. et al., 2022; Yi et al., 2023; Binder et al., 2022). The Xgboost algorithm has excellent performance in dealing with classification problems. It is characterized by its ability to handle large volumes of data while maintaining accuracy and predictive performance over other classification algorithms (Yue et al., 2022; Li Q. et al., 2022). Compared to XGBoost, the SVM algorithm is slightly less accurate and less predictive. However, in some special problems, SVM has advantages that other algorithms do not have, such as when

dealing with nonlinearly differentiable data and when dealing with high-dimensional data (Huang et al., 2018; Ding et al., 2022).

Undeniably, biomarker-based machine learning diagnostic models for AD are still the most dominant research (Shah et al., 2023; Kononikhin et al., 2022). The biomarkers involved in these studies include not only the pathology products that we generally recognize, but also some lipids, proteins and so on that are closely related to the pathogenesis of AD as screened by new research methods (Wang et al., 2022). The great progress in molecular biology research has also led to the expansion of the scope of clinical biomarkers (Krokidis et al., 2023). Their potential for clinical application will be enhanced if the cost of the assay can be reduced. In addition to the traditional work related to the construction of biomarker-based AD diagnostic models, more AD-related ancillary test results have been included in the study. Common neurological examinations in the clinic, such as electroencephalography, are used to construct AD diagnostic models, which also have good diagnostic performance (Parreño Torres et al., 2023). In addition to common examination means, more and more medical devices provide more dimensional examination results. As a new type of model, AD diagnostic model based on eye movement and language has a good prospect for clinical application due to its noninvasive and easy-to-operate characteristics (Jang et al., 2021). Overall, research efforts in AD machine learning diagnostic modeling have produced a large number of models with very good performance and potential for clinical applications. However, unfortunately, more comprehensive AD diagnostic models covering multidimensional markers have not yet been developed.

Inevitably, there are shortcomings in our study. In the section examining the question of the association between pathologic products and age, we had planned to compare all preclinical stages of AD separately. However, upon attempting this, we found that this would make the presentation and interpretation of the results extraordinarily difficult, and the selection of a control group of AD patients presented considerable difficulties. In terms of AD model construction and training, we were hampered by the fact that public databases do not categorize the A $\beta$  and Tau protein families in great detail. Other than that, we mainly chose the Xgboost algorithm and SVM algorithm, which have excellent performance in solving classification problems. We have also tried other algorithms and will try to use more algorithms to try to build AD diagnostic models in the future (Zhang et al., 2023). In addition to this, our model construction has not been externally validated due to lack of data from other database sources. However, because the data provided by the ADNI database is collected from multiple centers and has a large sample size, its accuracy and reliability for real-world application will be guaranteed. In the meantime, we are currently in the process of clinically collecting additional data to complete external validation.

## Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

## Ethics statement

All participating institutions' Institutional Review Boards have given their approval for the ADNI trial. All participants, or their authorized representatives, have given written informed permission to ADNI in compliance with the Declaration of Helsinki. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

## Author contributions

WZ: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. YuW: Data curation, Resources, Writing – review & editing. HZ: Methodology, Resources, Software, Writing – review & editing. WZha: Data curation, Resources, Writing – review & editing. WS: Data curation, Resources, Writing – review & editing. YS: Data curation, Resources, Writing – review & editing. YQ: Data curation, Resources, Writing – review & editing. SJ: Data curation, Resources, Writing – review & editing. ZZ: Data curation, Resources, Writing – review & editing. YuyW: Data curation, Resources, Writing – review & editing. LC: Data curation, Writing – review & editing. JZ: Data curation, Writing – review & editing. DP: Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – review & editing.

## References

- Alzheimer's Association Report (2022). 2022 Alzheimer's disease facts and figures. *Alzheimers. Dement.* 18, 700–789. doi: 10.1002/alz.12638
- Alzheimer's Association Report (2023). 2023 Alzheimer's disease facts and figures. *Alzheimers. Dement.* 19, 1598–1695. doi: 10.1002/alz.13016
- Alzheimer's Association Report (2024). 2024 Alzheimer's disease facts and figures. *Alzheimers. Dement.* 20, 3708–3821. doi: 10.1002/alz.13809
- Ashton, N. J., Pascoal, T. A., Karikari, T. K., Benedet, A. L., Lantero-Rodriguez, J., Brinkmalm, G., et al. (2021). Plasma p-tau231: a new biomarker for incipient Alzheimer's disease pathology. *Acta Neuropathol.* 141, 709–724. doi: 10.1007/s00401-021-02275-6
- Bai, B., Vanderwall, D., Li, Y., Wang, X., Poudel, S., Wang, H., et al. (2021). Proteomic landscape of Alzheimer's disease: novel insights into pathogenesis and biomarker discovery. *Mol. Neurodegener.* 16:55. doi: 10.1186/s13024-021-00474-z
- Barthélemy, N. R., Li, Y., Joseph-Mathurin, N., Gordon, B. A., Hassenstab, J., Benzinger, T. L. S., et al. (2020). A soluble phosphorylated tau signature links tau, amyloid and the evolution of stages of dominantly inherited Alzheimer's disease. *Nat. Med.* 26, 398–407. doi: 10.1038/s41591-020-0781-z
- Binder, J., Ursu, O., Bologna, C., Jiang, S., Maphis, N., Dadras, S., et al. (2022). Machine learning prediction and tau-based screening identifies potential Alzheimer's disease genes relevant to immunity. *Commun. Biol.* 5:125. doi: 10.1038/s42003-022-03068-7
- Boza-Serrano, A., Vrillon, A., Minta, K., Paulus, A., Camprubi-Ferrer, L., Garcia, M., et al. (2022). Galectin-3 is elevated in CSF and is associated with A $\beta$  deposits and tau aggregates in brain tissue in Alzheimer's disease. *Acta Neuropathol.* 144, 843–859. doi: 10.1007/s00401-022-02469-6
- Busche, M. A., and Hyman, B. T. (2020). Synergy between amyloid- $\beta$  and tau in Alzheimer's disease. *Nat. Neurosci.* 23, 1183–1193. doi: 10.1038/s41593-020-0687-6
- Chiu, M. J., Yang, S. Y., Chen, T. F., Lin, C. H., Yang, F. C., Chen, W. P., et al. (2021). Synergistic association between plasma A $\beta$ (1-42) and p-tau in Alzheimer's disease but not in Parkinson's disease or frontotemporal dementia. *ACS Chem. Neurosci.* 12, 1376–1383. doi: 10.1021/acscchemneuro.1c00010
- Cogswell, P. M., Lundt, E. S., Therneau, T. M., Wiste, H. J., Graff-Radford, J., Algeciras-Schimnich, A., et al. (2024). Modeling the temporal evolution of plasma p-tau in relation to amyloid beta and tau PET. *Alzheimers. Dement.* 20, 1225–1238. doi: 10.1002/alz.13539
- Ding, C., Bao, T. Y., and Huang, H. L. (2022). Quantum-inspired support vector machine. *IEEE Trans. Neural Netw. Learn. Syst.* 33, 7210–7222. doi: 10.1109/TNNLS.2021.3084467
- Ferrari-Souza, J. P., Ferreira, P. C. L., Bellaver, B., Tissot, C., Wang, Y. T., Leffa, D. T., et al. (2022). Astrocyte biomarker signatures of amyloid- $\beta$  and tau pathologies in Alzheimer's disease. *Mol. Psychiatry* 27, 4781–4789. doi: 10.1038/s41380-022-01716-2
- Ferreiro, A. L., Choi, J., Ryou, J., Newcomer, E. P., Thompson, R., Bollinger, R. M., et al. (2023). Gut microbiome composition may be an indicator of preclinical Alzheimer's disease. *Sci. Transl. Med.* 15:eabo2984. doi: 10.1126/scitranslmed.abo2984
- Ficiarà, E., Boschi, S., Ansari, S., D'agata, F., Abollino, O., Caroppo, P., et al. (2021). Machine learning profiling of Alzheimer's disease patients based on current cerebrospinal fluid markers and iron content in biofluids. *Front. Aging Neurosci.* 13:607858. doi: 10.3389/fnagi.2021.607858
- Franciotti, R., Nardini, D., Russo, M., Onofri, M., and Sensi, S. L. (2023). Comparison of machine learning-based approaches to predict the conversion to

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by National Key R&D Program of China (Grant No. 2022YFC2010103), Central health research project (Grant No. 2020ZD10), and 2023 Independent Research Projects for Graduate Students of Beijing University of Chinese Medicine (Grant No. ZJKT2023116).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Alzheimer's disease from mild cognitive impairment. *Neuroscience* 514, 143–152. doi: 10.1016/j.neuroscience.2023.01.029
- Gaetani, L., Bellomo, G., Parnetti, L., Blennow, K., Zetterberg, H., and Di Filippo, M. (2021). Neuroinflammation and Alzheimer's disease: a machine learning approach to CSF proteomics. *Cells* 10:1930. doi: 10.3390/cells10081930
- Gao, F., Lv, X., Dai, L., Wang, Q., Wang, P., Cheng, Z., et al. (2023). A combination model of AD biomarkers revealed by machine learning precisely predicts Alzheimer's dementia: China Aging and Neurodegenerative Initiative (CANDI) study. *Alzheimers Dement.* 19, 749–760. doi: 10.1002/alz.12700
- Hammond, T. C., Xing, X., Wang, C., Ma, D., Nho, K., Crane, P. K., et al. (2020).  $\beta$ -amyloid and tau drive early Alzheimer's disease decline while glucose hypometabolism drives late decline. *Commun Biol* 3:352. doi: 10.1038/s42003-020-1079-x
- Hampel, H., Blennow, K., Shaw, L. M., Hoessler, Y. C., Zetterberg, H., and Trojanowski, J. Q. (2010). Total and phosphorylated tau protein as biological markers of Alzheimer's disease. *Exp. Gerontol.* 45, 30–40. doi: 10.1016/j.exger.2009.10.010
- Horie, K., Salvadó, G., Barthélemy, N. R., Janelidze, S., Li, Y., He, Y., et al. (2023). CSF MTBR-tau243 is a specific biomarker of tau tangle pathology in Alzheimer's disease. *Nat. Med.* 29, 1954–1963. doi: 10.1038/s41591-023-02443-z
- Huang, S., Cai, N., Pacheco, P. P., Narrandes, S., Wang, Y., and Xu, W. (2018). Applications of support vector machine (SVM) learning in cancer genomics. *Cancer Genom. Proteom.* 15, 41–51. doi: 10.21873/cgp.20063
- Izzo, N. J., Yuede, C. M., Labarbera, K. M., Limegrover, C. S., Rehak, C., Yurko, R., et al. (2021). Preclinical and clinical biomarker studies of CT1812: a novel approach to Alzheimer's disease modification. *Alzheimers Dement.* 17, 1365–1382. doi: 10.1002/alz.12302
- Jang, H., Soroski, T., Rizzo, M., Barral, O., Harisinghani, A., Newton-Mason, S., et al. (2021). Classification of Alzheimer's disease leveraging multi-task machine learning analysis of speech and eye-movement data. *Front. Hum. Neurosci.* 15:716670. doi: 10.3389/fnhum.2021.716670
- Jia, J., Ning, Y., Chen, M., Wang, S., Yang, H., Li, F., et al. (2024). Biomarker Changes during 20 Years Preceding Alzheimer's Disease. *N. Engl. J. Med.* 390, 712–722. doi: 10.1056/NEJMoa2310168
- Khan, A., Zubair, S., Shuaib, M., Sheneamer, A., Alam, S., and Assiri, B. (2024). Development of a robust parallel and multi-composite machine learning model for improved diagnosis of Alzheimer's disease: correlation with dementia-associated drug usage and AT(N) protein biomarkers. *Front. Neurosci.* 18:1391465. doi: 10.3389/fnins.2024.1391465
- Kononikhin, A. S., Zakharova, N. V., Semenov, S. D., Bugrova, A. E., Brzhozovskiy, A. G., Indeykina, M. I., et al. (2022). Prognosis of Alzheimer's disease using quantitative mass spectrometry of human blood plasma proteins and machine learning. *Int. J. Mol. Sci.* 23:7907. doi: 10.3390/ijms23147907
- Krokidis, M. G., Vrahatis, A. G., Lazaros, K., Skolariki, K., Exarchos, T. P., and Vlamos, P. (2023). Machine learning analysis of Alzheimer's disease single-cell RNA-sequencing data across cortex and hippocampus regions. *Curr. Issues Mol. Biol.* 45, 8652–8669. doi: 10.3390/cimb45110544
- Küçükalı, F., Neumann, A., Van Dongen, J., De Pooter, T., Joris, G., De Rijk, P., et al. (2023). Whole-exome rare-variant analysis of Alzheimer's disease and related biomarker traits. *Alzheimers Dement.* 19, 2317–2331. doi: 10.1002/alz.12842
- Kumari, S., Singh, A., Singh, A. K., Yadav, Y., Bajpai, S., Kumar, P., et al. (2022). Circulatory GSK-3 $\beta$ : blood-based biomarker and therapeutic target for Alzheimer's disease. *J. Alzheimers Dis.* 85, 249–260. doi: 10.3233/JAD-215347
- Li, J., Zhang, Y., Lu, T., Liang, R., Wu, Z., Liu, M., et al. (2022). Identification of diagnostic genes for both Alzheimer's disease and Metabolic syndrome by the machine learning algorithm. *Front. Immunol.* 13:1037318. doi: 10.3389/fimmu.2022.1037318
- Li, K., Wei, Q., Liu, F. F., Hu, F., Xie, A. J., Zhu, L. Q., et al. (2018). Synaptic dysfunction in Alzheimer's disease: A $\beta$ , tau, and epigenetic alterations. *Mol. Neurobiol.* 55, 3021–3032. doi: 10.1007/s12035-017-0533-3
- Li, Q., Yang, H., Wang, P., Liu, X., Lv, K., and Ye, M. (2022). XGBoost-based and tumor-immune characterized gene signature for the prediction of metastatic status in breast cancer. *J. Transl. Med.* 20:177. doi: 10.1186/s12967-022-03369-9
- Mattsson-Carlgen, N., Andersson, E., Janelidze, S., Ossenkoppele, R., Insel, P., Strandberg, O., et al. (2020). A $\beta$  deposition is associated with increases in soluble and phosphorylated tau that precede a positive Tau PET in Alzheimer's disease. *Sci. Adv.* 6:eaz2387. doi: 10.1126/sciadv.aaz2387
- Pang, K., Jiang, R., Zhang, W., Yang, Z., Li, L. L., Shimozawa, M., et al. (2022). An App knock-in rat model for Alzheimer's disease exhibiting A $\beta$  and tau pathologies, neuronal death and cognitive impairments. *Cell Res.* 32, 157–175. doi: 10.1038/s41422-021-00582-x
- Parreño Torres, A., Roncero-Parra, C., Borja, A. L., and Mateo-Sotos, J. (2023). Inter-hospital advanced and mild Alzheimer's disease classification based on electroencephalogram measurements via classical machine learning algorithms. *J. Alzheimers Dis.* 95, 1667–1683. doi: 10.3233/JAD-230525
- Petersen, R. C., Aisen, P. S., Beckett, L. A., Donohue, M. C., Gamst, A. C., Harvey, D. J., et al. (2010). Alzheimer's Disease Neuroimaging Initiative (ADNI): clinical characterization. *Neurology* 74, 201–209. doi: 10.1212/WNL.0b013e3181cb3e25
- Pichet Binette, A., Theaud, G., Rheault, F., Roy, M., Collins, D. L., Levin, J., et al. (2021). Bundle-specific associations between white matter microstructure and A $\beta$  and tau pathology in preclinical Alzheimer's disease. *Elife* 10:e62929. doi: 10.7554/eLife.62929
- Sadleir, K. R., and Vassar, R. (2023). Connections between ApoE, sleep, and A $\beta$  and tau pathologies in Alzheimer's disease. *J. Clin. Invest.* 133:171838. doi: 10.1172/JCI171838
- Scheltens, P., De Strooper, B., Kivipelto, M., Holstege, H., Chételat, G., Teunissen, C. E., et al. (2021). Alzheimer's disease. *Lancet* 397, 1577–1590. doi: 10.1016/S0140-6736(20)32205-4
- Shah, J., Rahman Siddiquee, M. M., Krell-Roesch, J., Syrjanen, J. A., Kremers, W. K., Vassilaki, M., et al. (2023). Neuropsychiatric symptoms and commonly used biomarkers of Alzheimer's disease: a literature review from a machine learning perspective. *J. Alzheimers Dis.* 92, 1131–1146. doi: 10.3233/JAD-221261
- Shaw, L. M., Vanderstichele, H., Knapik-Czajka, M., Clark, C. M., Aisen, P. S., Petersen, R. C., et al. (2009). Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects. *Ann. Neurol.* 65, 403–413. doi: 10.1002/ana.21610
- Stern, A. M., Van Pelt, K. L., Liu, L., Anderson, A. K., Ostaszewski, B., Mapstone, M., et al. (2023). Plasma NT1-tau and A $\beta$ (42) correlate with age and cognitive function in two large Down syndrome cohorts. *Alzheimers Dement.* 19, 5755–5764. doi: 10.1002/alz.13382
- Wang, Y., Sun, Y., Wang, Y., Jia, S., Qiao, Y., Zhou, Z., et al. (2023a). Urine metabolomics phenotyping and urinary biomarker exploratory in mild cognitive impairment and Alzheimer's disease. *Front. Aging Neurosci.* 15:1273807. doi: 10.3389/fnagi.2023.1273807
- Wang, Y., Sun, Y., Wang, Y., Jia, S., Qiao, Y., Zhou, Z., et al. (2023b). Identification of novel diagnostic panel for mild cognitive impairment and Alzheimer's disease: findings based on urine proteomics and machine learning. *Alzheimers Res. Ther.* 15:191. doi: 10.1186/s13195-023-01324-4
- Wang, Z., Ye, J., Zhang, K., Ding, L., Granzier-Nakajima, T., Ranasinghe, J. C., et al. (2022). Rapid biomarker screening of Alzheimer's disease by interpretable machine learning and graphene-assisted raman spectroscopy. *ACS Nano* 16, 6426–6436. doi: 10.1021/acsnano.2c00538
- Winchester, L. M., Harshfield, E. L., Shi, L., Badhwar, A., Khleifat, A. A., Clarke, N., et al. (2023). Artificial intelligence for biomarker discovery in Alzheimer's disease and dementia. *Alzheimers Dement.* 19, 5860–5871. doi: 10.1002/alz.13390
- Yakoub, Y., Ashton, N. J., Strikwerda-Brown, C., Montoliu-Gaya, L., Karikari, T. K., Kac, P. R., et al. (2023). Longitudinal blood biomarker trajectories in preclinical Alzheimer's disease. *Alzheimers Dement.* 19, 5620–5631. doi: 10.1002/alz.13318
- Yang, J., Feng, X., Laine, A. F., and Angelini, E. D. (2020). Characterizing Alzheimer's disease with image and genetic biomarkers using supervised topic models. *IEEE J. Biomed. Health Inform.* 24, 1180–1187. doi: 10.1109/JBHI.2019.2928831
- Yi, F., Yang, H., Chen, D., Qin, Y., Han, H., Cui, J., et al. (2023). XGBoost-SHAP-based interpretable diagnostic framework for alzheimer's disease. *BMC Med. Inform. Decis. Mak.* 23, 137. doi: 10.1186/s12911-023-02238-9
- Yue, S., Li, S., Huang, X., Liu, J., Hou, X., Zhao, Y., et al. (2022). Machine learning for the prediction of acute kidney injury in patients with sepsis. *J. Transl. Med.* 20, 215. doi: 10.1186/s12967-022-03364-0
- Zhang, H., Wei, W., Zhao, M., Ma, L., Jiang, X., Pei, H., et al. (2021). Interaction between A $\beta$  and tau in the pathogenesis of Alzheimer's disease. *Int. J. Biol. Sci.* 17, 2181–2192. doi: 10.7150/ijbs.57078
- Zhang, Y., Miao, Y., Tan, J., Chen, F., Lei, P., and Zhang, Q. (2023). Identification of mitochondrial related signature associated with immune microenvironment in Alzheimer's disease. *J. Transl. Med.* 21, 458. doi: 10.1186/s12967-023-04254-9
- Zhang, Z., Li, X. G., Wang, Z. H., Song, M., Yu, S. P., Kang, S. S., et al. (2021).  $\delta$ -Secretase-cleaved tau stimulates A $\beta$  production via upregulating STAT1-BACE1 signaling in Alzheimer's disease. *Mol. Psychiatry* 26, 586–603. doi: 10.1038/s41380-018-0286-z