Received: 5 November 2024 | Revised: 1 February 2025 | Accepted: 3 February 2025

DOI: 10.1002/alz.70044

Alzheimer's & Dementia®
THE JOURNAL OF THE ALZHEIMER'S ASSOCIATION

RESEARCH ARTICLE

Circulating proteomic biomarkers for cerebral amyloid angiopathy screening and risk stratification

Jiajie Xu[1] | **Ya Su**[1] | **Yuhui Sha**[2] | **Jiayu Fu**[1] | **Tingmeng Yan**[1] | **Feifan Xu**[1] | **Han Tang**[1] | **Yunqing Ying**[1] | **Yiwei Xia**[1] | **Qiang Dong**[1,3] | **M. Edip Gurol**[4] | **Jun Ni**[2] | **Xin Cheng**[1]

[1]Department of Neurology, National Center for Neurological Disorders, National Clinical Research Centre for Aging and Medicine, Huashan Hospital, Fudan University, Shanghai, P. R. China

[2]Department of Neurology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, P. R. China

[3]State Key Laboratory of Medical Neurobiology, Fudan University, Shanghai, P. R. China

[4]J Philip Kistler Stroke Research Center, Department of Neurology, Massachusetts General Hospital, Boston, Massachusetts, USA

Correspondence
Xin Cheng, Department of Neurology, National Center for Neurological Disorders, National Clinical Research Centre for Aging and Medicine, Huashan Hospital Fudan University, No.12 Wulumuqi Zhong Road, Shanghai 200040, P. R. China.
Email: chengxin@fudan.edu.cn

Jun Ni, Department of Neurology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, No. 1 Shuaifuyuan, Dongdan, Dongcheng District, Beijing 100032, P. R. China.
Email: pumchnijun@163.com

Funding information
National Natural Science Foundation of China, Grant/Award Numbers: 81971123, 82201467; Noncommunicable Chronic Diseases-National Science and Technology Major Project, Grant/Award Number: 2023ZD0504903; Shanghai Municipal Health Commission, Grant/Award Number: 2022XD022; Sailing Program of Shanghai Science and Technology Committee, Grant/Award Number: 22YF1405500; Shanghai "Rising Stars of Medical Talents" Youth Development Program, Grant/Award Number: SHWSRS (2023)_062; CAMS Innovation Fund for Medical Sciences (CIFMS), Grant/Award Number: 2024-I2M-C&T-B-007; Brain Science and Brain Diseases Youth

Abstract

INTRODUCTION: We systematically characterized plasma protein profiles in cerebral amyloid angiopathy (CAA) using proteomics and identified a hub protein panel for disease diagnosis and risk stratification.

METHODS: A total of 146 patients with probable CAA and 128 community-dwelling controls were prospectively enrolled. Plasma samples underwent proteomic analysis, and the hub proteins were validated in two validation cohorts. Machine learning algorithms were applied to construct and validate the performance of the circulating panel.

RESULTS: We identified 166 differentially expressed proteins in patients with CAA. Six hub proteins were selected to form the circulating panel, demonstrating excellent performance to distinguish patients from controls in all cohorts. Additionally, the risk stratification system derived from the hub protein panel accurately identified patients at high risk for new-onset lobar intracerebral hemorrhage.

DISCUSSION: Our findings revealed distinctive circulating protein signatures in CAA and established a validated hub protein panel aiding in CAA screening and risk stratification.

KEYWORDS
cerebral amyloid angiopathy, disease screening, machine learning algorithm, plasma biomarker, proteomic analysis, risk stratification

Highlights

- We identified plasma protein signatures in CAA using proteomics.

Jiajie Xu, Ya Su, and Yuhui Sha contributed equally to this work.

Innovation Program of Shanghai Zhou Liangfu Medical Development Foundation

- A circulating hub protein panel was developed and validated, demonstrating high accuracy for disease screening.
- The hub protein panel effectively stratified CAA patients according to risk of future intracerebral hemorrhage.
- The circulating panel offers potential for CAA screening and risk stratification.

1 | BACKGROUND

Cerebral amyloid angiopathy (CAA) is a major cause of spontaneous lobar intracerebral hemorrhage (ICH) and cognitive decline in the elderly, characterized by the progressive deposition of amyloid beta (Aβ) in cerebral small vessels.[1] Notably, patients with CAA-related ICH (CAA-ICH) show higher recurrence rates compared to those with hypertension-related ICH, leading to a poorer prognosis.[2,3] Therefore, effective disease screening and risk stratification to identify high-risk individuals are of particular importance. The Boston criteria provide a diagnostic framework for CAA based on clinical manifestations and neuroimaging markers.[4–6] While the Boston criteria version 2.0 has achieved higher sensitivity without compromising specificity,[6] there is still room for improvement, especially in asymptomatic patients or those without prior ICH.[7]

Recent studies have increasingly focused on biomarkers for disease screening and prognostic assessment. For example, in a cohort of Dutch-type hereditary CAA, cerebrospinal fluid (CSF) levels of Aβ_{40} and Aβ_{42} were found to decrease early in the disease course.[8] A meta-analysis further revealed elevated CSF total tau protein levels in patients with CAA compared to healthy controls.[9] Other CSF biomarkers, such as neurofilament light chain (NfL) and glial fibrillary acidic protein (GFAP), have also been explored as potential biomarkers.[10] Given the invasiveness of lumbar puncture, there is a growing interest in peripheral blood biomarkers for CAA, as they are more accessible and less invasive. For instance, elevated serum NfL and GFAP levels have been observed in patients with hereditary CAA compared to controls.[10] Our previous work also showed that serum levels of NfL were associated with ICH recurrence in sporadic CAA-ICH.[11] In addition, we found that serum levels of matrix metalloproteinases (MMPs) and chitinase-3-like-1 protein (CHI3L1) might be helpful in CAA diagnosis and prognostic prediction.[12,13] These findings indicate that circulating biomarkers might reflect CAA pathology and could be valuable in future clinical practice.

Nonetheless, most explorations to date have focused on certain proteins and failed to capture the comprehensive spectrum of peripheral protein alterations in CAA.[11–14] With advancements in techniques, proteomic analysis provides an opportunity to improve our understanding of CAA and identify new biomarkers. Previous studies revealed distinct CSF protein expression patterns in patients with CAA and in rat models through proteomic analysis.[15] Additionally, proteomic analysis has the potential to uncover novel biomarkers for prognostic assessment.

Given the high risk of ICH occurrence and recurrence in patients with CAA, we aimed to establish a circulating protein panel for disease screening and risk stratification. First, we depicted the plasma protein profiles of patients with CAA and explored novel protein signatures. Then we employed machine learning algorithms to select and cross-validate a panel of six hub proteins for disease screening. Next, we used the same hub protein panel to construct a predictive model and evaluate its performance in predicting new-onset ICH events. Finally, we validated the diagnostic and predictive performance of the hub protein panel in both internal and external validation cohorts, hoping to provide an in-depth understanding of CAA pathogenesis and develop a clinically applicable strategy for prognostic assessment.

2 | METHODS

2.1 | Study participants

A summary of the participant enrollment and allocation scheme is presented in Figure S1. Patients with a diagnosis of probable CAA according to the Boston criteria version 1.5[5] were prospectively recruited from two centers, and those with available plasma samples were enrolled in this study. Age- and sex-matched community-dwelling controls without a history of stroke or dementia with available plasma samples were recruited from the Shanghai Aging Study.[16] Detailed inclusion and exclusion criteria were described previously.[17]

To establish and validate the hub protein panel, participants were divided into three independent cohorts: the derivation cohort, the internal validation cohort, and the independent external validation cohort (Figure S1). The derivation cohort and the internal validation cohort comprised patients with CAA prospectively recruited from Huashan Hospital, Fudan University (Shanghai, China), and controls from the Shanghai Aging Study between January 2015 and July 2020. These participants were randomly divided into the derivation cohort and the internal validation cohort at a ratio of 7:3. The external derivation cohort comprised patients with CAA recruited from Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, and Peking Union Medical College (Beijing, China), and controls from the Shanghai Aging Study during October 2017 and July 2023.

The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of Huashan Hospital, Fudan University (KY2015-263 and KY2021-810), and Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, and Peking Union Medical College (I-23PJ878). All

participants or their legal representatives provided written informed consents.

2.2 | Follow-up and the primary outcome

All patients were regularly followed up every 6 months by phone or clinic visits until the occurrence of primary outcome, death, or December 2023. The primary outcome of the study was the new-onset symptomatic lobar ICH, including recurrent CAA-ICH and first-ever CAA-ICH. Representative examples of the primary outcome are presented in Figure S2. Recurrent CAA-ICH was defined as a new non-traumatic lobar ICH lesion confirmed on computed tomography (CT) or susceptibility-weighted imaging (SWI) with corresponding symptoms in patients with CAA-ICH.[18] First-ever CAA-ICH was defined as a new symptomatic and non-traumatic lobar ICH in patients without a history of previous ICH during the follow-up period.[19]

2.3 | MRI acquisition and neuroimaging marker evaluation

In the derivation cohort and internal validation cohort, patients with CAA were scanned on a Siemens MAGNETOM Verio 3.0 T scanner, while in the external validation cohort, patients were scanned on a GE Discovery MR750 W 3.0 T scanner. All controls in the study were scanned on a GE Discovery MR750 3.0 T scanner. T1-weighted, T2-weighted, FLAIR, and SWI sequences were acquired, and detailed parameters are provided in Supplementary Materials.

Two experienced neurologists who were blinded to clinical information evaluated small vessel disease (SVD) markers according to the STandards for ReportIng Vascular changes on nEuroimaging (STRIVE) 2 criteria.[20] A senior neurologist was consulted in cases of disagreement. SVD markers measured in the study included perivascular spaces in the centrum semiovale (CSO-PVS), lobar lacunes, white matter hyperintensity (WMH) Fazekas scale, lobar cerebral microbleeds (CMBs), cortical superficial siderosis (cSS) scale, and total MRI burden of SVD for CAA; evaluation methods are described in Supplementary Materials.

2.4 | Global cognitive status assessment

The Chinese version of the Mini-Mental State Examination (MMSE) was utilized to assess the global cognitive function of participants in the study by an experienced neurologist. Seventeen patients in center one and six patients in center two had symptoms including a disorder of consciousness or slurred speech, in which condition the cognitive test could hardly be conducted. For the missing data, we carried out the random forest (RF) regression approach[21] to impute values.

2.5 | Plasma sample preparation for proteomics

Venous blood samples were collected within 24 h of enrollment on the same day as the MRI was performed. The blood samples were centrifuged at 1000 × g for 10 min twice at 4°C, and plasma aliquots were stored at −80°C. Samples were gradually brought to room temperature for measurement. High-abundance proteins in the plasma samples were removed using an Agilent Multiple Affinity Removal System Column according to the manufacturer's protocol. Protein concentrations in the supernatant were quantified with the BCA Protein Assay Kit (Bio-Rad, USA). Proteins were then digested with trypsin following reduction with dithiothreitol (DTT) and alkylation with iodoacetamide (IAM). After precipitation with chilled acetone, proteins were resuspended in ammonium bicarbonate and digested overnight at 37°C. Peptides were desalted using a C18 cartridge for further analysis.

For biomarker discovery, untargeted proteomic analysis was performed on plasma samples from participants in the derivation cohort and internal validation cohort, while targeted proteomic analysis was carried out on plasma samples from those in the external validation cohort to verify the biomarkers.

RESEARCH IN CONTEXT

1. **Systematic review**: The authors conducted a literature review using PubMed. Blood proteins are emerging as promising biomarkers in neurological disorders. Some studies have identified several peripheral proteins associated with the diagnosis of CAA, but few have systematically characterized the peripheral blood protein profiles of CAA.
2. **Interpretation**: This study performed a proteomic analysis and comprehensively profiled plasma proteins in CAA. A panel of six hub proteins, selected by machine learning, could distinguish patients from controls with high accuracy. In addition, the circulating hub protein panel successfully predicted patients at high risk for new-onset lobar intracerebral hemorrhage. These findings indicate the potential utility of this panel for CAA screening and risk stratification.
3. **Future directions**: Longitudinal studies are needed to validate the applicability of the hub protein panel in larger populations and to explore its potential in tracking disease progression. Future research should also investigate the underlying mechanisms linking these proteins to CAA pathology, enhancing our understanding of disease progression and discovering novel treatment targets.

2.6 | Untargeted proteomic analysis

The pooled samples were fractioned by an offline high-pH reverse-phase HPLC separation system (Agilent). For liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis, two data acquisition modes, data-dependent analysis (DDA) and data-independent analysis

(DIA), were utilized in the study. The detailed processes of untargeted proteomic analysis are detailed in Supplementary Materials. Expression levels of proteins were adjusted for age, sex, hypertension, and diabetes using a linear regression model. Proteins with false discovery rate (FDR)-adjusted *p* value (q value) < 0.05 and fold change (FC) > 1.2 or < 0.83 were declared differentially expressed.

2.7 | Differentially expressed protein visualization and enrichment analysis

Visualization of differentially expressed proteins was performed using volcano plots, principal component analysis (PCA) plots, and hierarchical clustering heatmaps, all generated in R (version 4.1.2). Gene Ontology (GO) enrichment analysis of differentially expressed proteins was conducted using the PANTHER database (http://www.pantherdb.org/). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis was performed using the KOBAS online tool (http://bioinfo.org/kobas).

2.8 | Hub protein selection

Correlations between differentially expressed proteins were determined using pairwise Pearson's correlation coefficients. An unsupervised hierarchical agglomerative clustering analysis was then conducted to identify clusters using the Ward.D2 method with the Euclidean distances in the correlation matrix.[22] The optimal number of clusters was determined using the gap statistic[23] (Figure S3). Additionally, the t-distributed stochastic neighbor embedding (t-SNE) dimensionality reduction analysis, which captured local similarities between data points and generated a low-dimensional mapping, was performed to further validate the clustering results.[24] Different colors were assigned to data points according to identified clusters, allowing for clear distinction between clusters. For each cluster, the candidate protein with the smallest q value was defined as the hub protein.

2.9 | Targeted proteomic analysis

To verify and quantify hub proteins, the 4D-parallel reaction monitoring (PRM) method was performed on plasma samples from external validation cohort. After enriching low abundance protein, samples were digested into peptide fragments as described earlier. The process of targeted proteomic analysis of hub proteins is described in detail in Supplementary Materials.

2.10 | Hub protein panel development for disease diagnosis

To develop a classifier for disease diagnosis, a set of machine learning algorithms, including the logistic regression,[25] elastic net regression,[26] support vector machine (SVM),[27] neural network,[28] recursive partitioning,[29] RF,[30] adaptive boosting (AdaBoost),[31] and eXtreme gradient boosting (XGBoost)[32] were applied. The input features for the classifier included age, sex, and selected hub proteins of both patients and controls. In the modeling process, participants in the derivation cohort were randomly split into a training set (70%) for model development and a testing set (30%) for model validation. During model development, hyperparameter tuning with 10-fold cross-validation was employed on the training set to verify the robustness of the model. The performance of each classifier was evaluated with area under the receiver operating characteristic (ROC) curve (AUC), accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) in both training and testing sets. The classifier that achieved the highest AUC in the testing set was selected as the final model for disease diagnosis. The final model was subsequently validated in the internal validation and external cohorts. The relative importance of hub proteins in the final classifier was estimated with SHapley Additive Explanations (SHAP) values.[33] AUCs of different classifiers were compared using the Delong test.[34]

2.11 | Hub protein panel exploration for risk stratification

To assess the risk stratification ability of the hub protein panel, a prognostic prediction model was developed using the same machine learning algorithm in the final diagnosis classifier. In this predictive model, the input features included selected hub protein panel, age, and sex of patients with CAA. Given the limited sample size of patients, hyperparameter tuning with cross-validation was also applied across all patients in the derivation cohort to optimize hyperparameters and develop the predictive model. The primary outcome predictive performance of the model was evaluated with AUC, accuracy, sensitivity, specificity, PPV, and NPV. Next, an individual CAA risk score was calculated for each patient based on outputs of the predictive model. Patients were then stratified according to their risk scores. Specifically, the Youden index was applied to determine the optimal cutoff value of risk scores.[35] Patients with risk scores below the cutoff value were stratified into the low-risk group; those with scores above the cutoff value were assigned to the high-risk group. Finally, to evaluate the performance of the risk stratification system, the primary outcome, global cognitive function, and neuroimaging markers were compared between patients in the high-risk and low-risk groups. Results were further validated in patients from both internal and external validation cohorts.

2.12 | Statistical analysis

All statistical analyses, machine learning strategies, and data visualizations were conducted using R software version 4.1.2. Continuous data were described as mean ± standard deviation (SD) or median and interquartile range (IQR). Categorical data were presented as counts

TABLE 1 Baseline characteristics of the study population.

	Derivation cohort			Internal validation cohort			External validation cohort		
	CAA ($n = 81$)	Control ($n = 68$)	p	CAA ($n = 35$)	Control ($n = 30$)	p	CAA ($n = 30$)	Control ($n = 30$)	p
Age (years)	69.69 ± 9.11	69.69 ± 6.65	1.000	71.03 ± 8.40	71.03 ± 6.68	0.998	69.87 ± 8.81	67.77 ± 6.44	0.296
Sex, male	47 (58.0)	41 (60.3)	0.910	24 (68.6)	19 (63.3)	0.856	18 (60.0)	15 (50.0)	0.604
Hypertension	61 (75.3)	19 (27.9)	<0.001	20 (57.1)	9 (30.0)	0.052	16 (53.3)	12 (40.0)	0.438
Diabetes mellitus	22 (27.2)	7 (10.3)	0.017	7 (20.0)	2 (6.7)	0.233	7 (23.3)	4 (13.3)	0.505
Antiplatelet drug use	25 (30.9)	12 (17.6)	0.095	8 (22.9)	5 (16.7)	0.756	3 (10.0)	1 (3.3)	0.605
Anticoagulant drug use	3 (3.7)	0 (0.0)	0.309	2 (5.7)	0 (0.0)	0.542	1 (3.3)	0 (0.0)	1.000
MMSE	20 [15, 25]	29 [27, 29]	<0.001	21 [20, 22.5]	29 [27, 30]	<0.001	23 [16.25, 26]	29 [27.25, 30]	<0.001
Total MRI burden of SVD	3 [3, 4]	0 [0, 1]	<0.001	3 [2, 4]	0 [0, 0]	<0.001	3.5 [3, 5]	0 [0, 1]	<0.001

Notes: Data are shown as number (%) for categorical variables, mean ± SD for normally distributed continuous variables, or median (interquartile range) for non-normally distributed continuous variables.
Abbreviations: CAA, cerebral amyloid angiopathy; IQR, interquartile range.; MMSE, Mini-Mental State Examination; SVD, small vessel disease.

and percentages (n [%]). Baseline differences between patients and controls in all cohorts were calculated with Student's t-test, Mann–Whitney U test, and χ^2 test. After normalization, the significance of proteomic data was evaluated using a linear regression model adjusted for age, sex, hypertension, and diabetes, with Benjamini–Hochberg FDR correction. Comparisons of hub proteins between patients and controls, as well as clinical variables between two risk stratification groups, were analyzed using a linear regression model, with age and sex as covariates. As to the follow-up data, Kaplan–Meier survival analysis with a log-rank test was applied to compare the occurrence of the primary outcome between high-risk and low-risk groups, reflecting the ability of the risk stratification system in prognostic prediction. The significance level was set at $\alpha = 0.05$.

Furthermore, a sensitivity analysis was conducted to explore the performance of the hub protein panel in prognostic assessment in addition to neuroimaging markers, and three models were established. The neuroimaging marker model included age and SVD markers that were significant in the univariate analysis. The protein panel model included age and hub proteins. The combined model included age, significant SVD markers, and hub proteins. The net reclassification index (NRI) and integrated discrimination improvement (IDI) index were used to evaluate the reclassification performance, with NRI or IDI > 0 indicating an improvement in the predictive ability of the model.[36]

3 | RESULTS

The study design is illustrated in Figure 1. A total of 146 patients with CAA and 128 community-dwelling controls were included in the final analysis. The baseline demographic characteristics of the derivation cohort, internal validation cohort, and external validation cohort are presented in Tables 1 and S1. Two groups were well matched for age and sex across all cohorts, while in the derivation cohort, patients with CAA had a higher prevalence of hypertension (75.3% vs 27.9%, $p < 0.001$; Table 1) and diabetes (27.2% vs 10.3%, $p = 0.017$; Table 1) compared to controls. Moreover, all patients exhibited a significantly heavier burden of SVD (all $p < 0.001$; Table 1).

3.1 | Overview of altered plasma protein expression profiles in CAA

In the derivation cohort, untargeted proteomic analysis identified 952 plasma proteins, with 166 proteins showing differential expression patterns after adjustment for age, sex, hypertension, and diabetes. Among them, 141 were upregulated and 25 were downregulated in patients with CAA compared to controls (Figure 2A). The PCA mapping demonstrated clear discrimination between patients and controls based on these differentially expressed proteins (Figure 2B), further corroborated by hierarchical clustering analysis (Figure 2C).

Next, GO-term enrichment analysis revealed that differentially expressed proteins were predominantly involved in immune response (FDR-adjusted $p = 4.31\times10^{-15}$), response to stimulus (FDR-adjusted $p = 8.49\times10^{-5}$), cell adhesion (FDR-adjusted $p = 2.89\times10^{-4}$), and protein-lipid complex remodeling processes (FDR-adjusted $p = 2.05\times10^{-3}$) (Figure 2D and Table S2). In addition, KEGG pathway analysis illustrated significant enrichment of pathways related to the lysosome (FDR-adjusted $p = 4.10\times10^{-2}$), platelet activation (FDR-adjusted $p = 4.10\times10^{-2}$), and cholesterol metabolism (FDR-adjusted $p = 4.10\times10^{-2}$) in CAA (Figure 2E and Table S3). These findings highlight the complex pathophysiology underlying CAA and offer a detailed view of the plasma proteomic landscape in these patients.

3.2 | Establishment of circulating hub protein panel for disease screening

The workflow of establishing and validating the hub protein panel for disease diagnosis was presented in Figure 3A. To explore the coregulation patterns among differentially expressed proteins, a

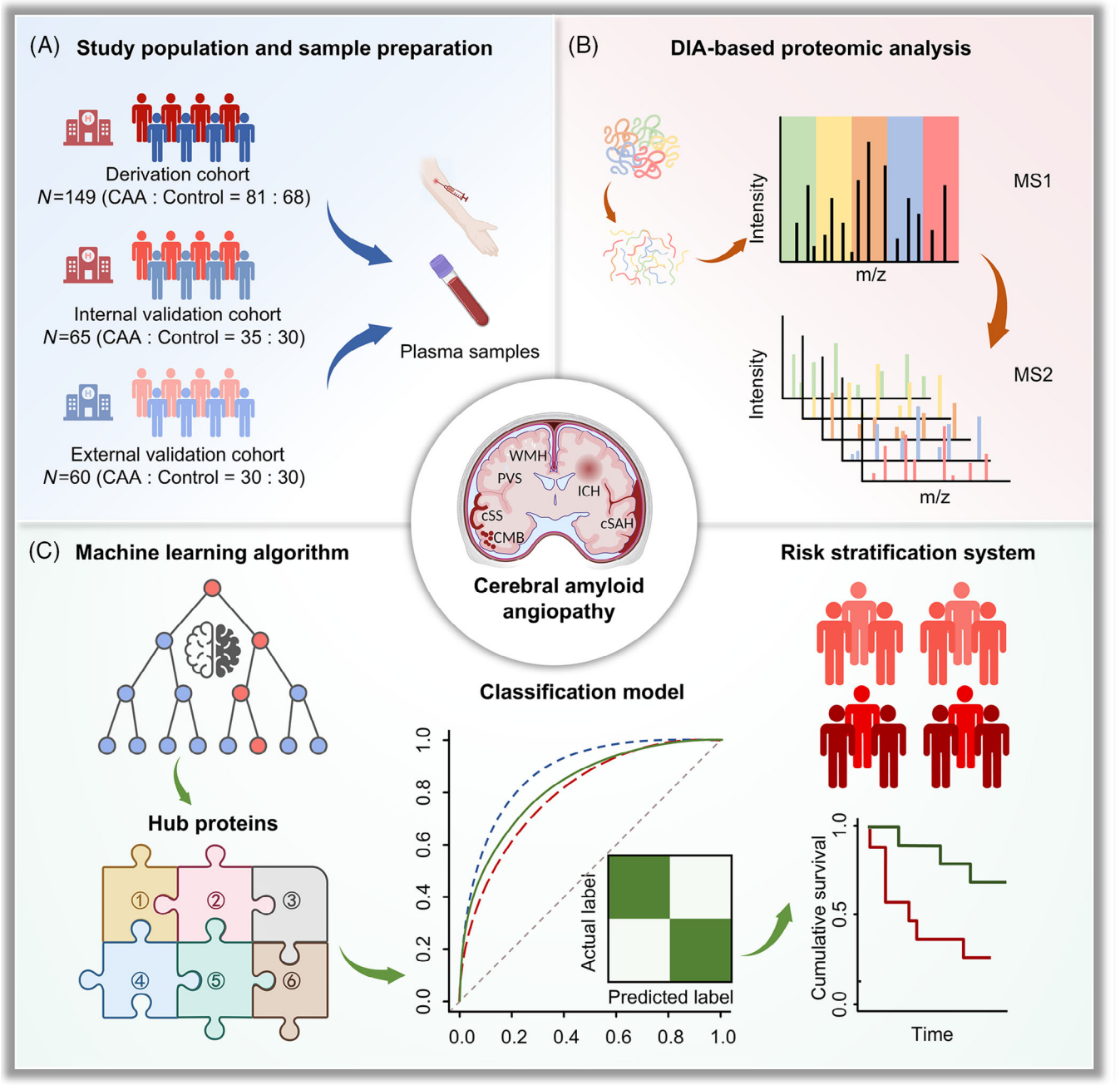

FIGURE 1 Schematic overview of study design. (A) Summary of study population and sample preparation. Patients with CAA and controls were prospectively enrolled and divided into derivation, internal validation, and external validation cohorts. Venous blood samples were collected from each participant within 24 h of enrollment, and plasma was prepared immediately. (B) Workflow of DIA-based proteomics analysis, including peptide preparation, peptide separation, and DIA LC-MS/MS data acquisition. (C) Workflow of bioinformatic analysis and model construction for disease diagnosis and risk stratification. Machine learning algorithms were used to select hub proteins and establish classifiers as well as the risk stratification system. CAA, cerebral amyloid angiopathy; DIA, data-independent acquisition; LC-MS/MS, liquid chromatography-tandem mass spectrometry.

correlation network analysis was performed, and hierarchical agglomerative clustering analysis identified six unique plasma protein clusters (Figure 3B, C). Notably, major biological processes of each cluster were similar to those of an entire protein set (Tables S2 and S4). To further confirm heterogeneity among clusters and homogeneity within each cluster, the t-SNE dimensionality reduction analysis demonstrated clear separations among clusters (Figure 3D). To be specific, proteins within the same cluster were adjacent in the low-dimensional space, while proteins in different clusters were far apart. This clustering pattern was consistent with the result of hierarchical agglomerative clustering, indicating the coregulation of proteins within each cluster. The hub protein in each cluster was defined as the one with the most

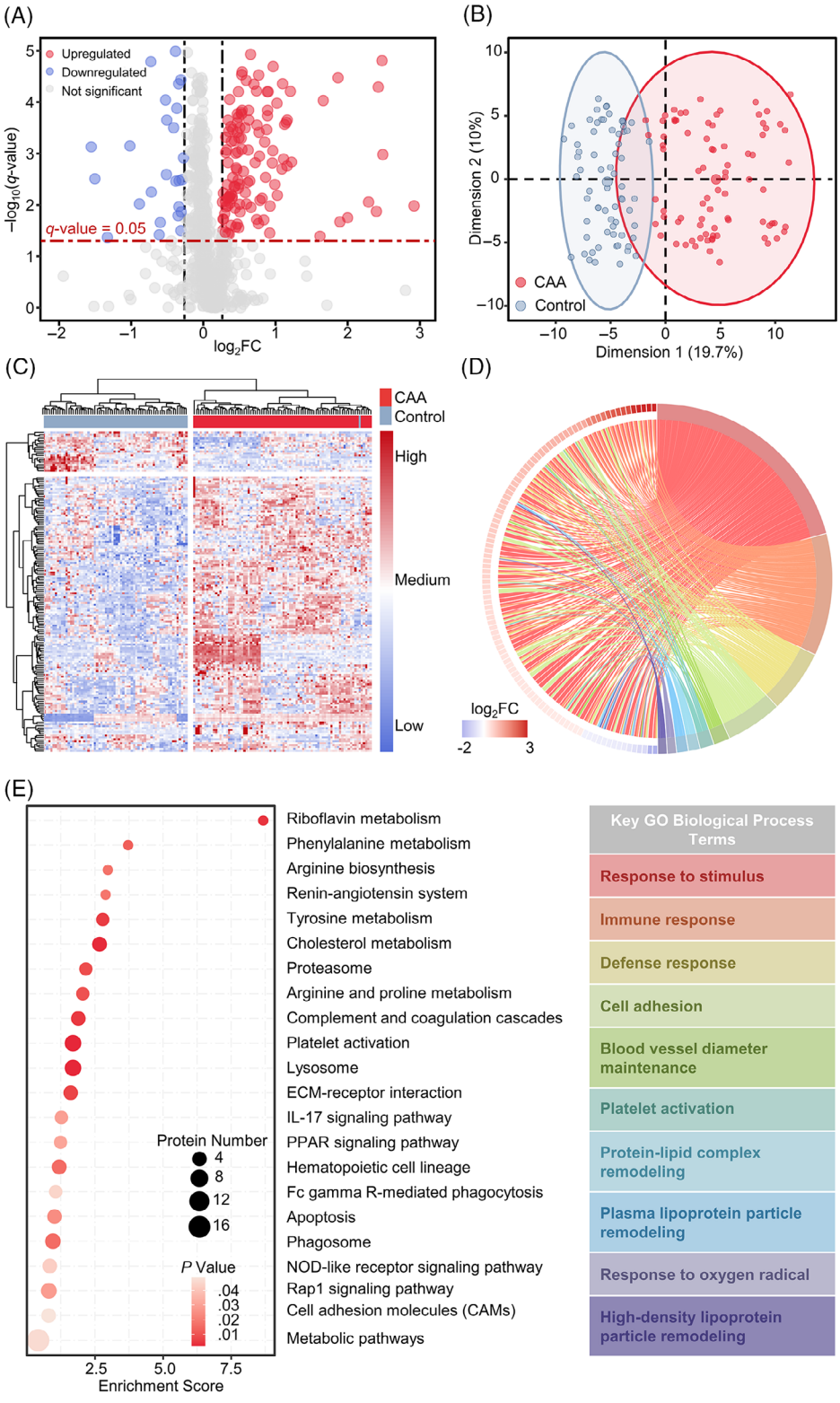

FIGURE 2 Circulating proteomic profiles of CAA. (A) Volcano plot displaying differentially expressed proteins in plasma between patients with CAA and controls. Each point represents a protein, with purple points representing downregulated proteins, red points representing upregulated proteins, and gray points representing proteins without significant difference. (B) PCA plot showing distribution of differentially expressed proteins. Each point represents a sample, with blue points for controls and red points for CAA. (C) Heatmap showing expression levels of differentially expressed proteins in all participants. Each column represents a sample, each row a protein. (D) Chord diagram displaying representative GO biological process in terms of differentially expressed proteins. Proteins with upregulated and downregulated expression levels indicated in red and purple, respectively. (E) Bubble plot displaying KEGG pathway analysis of differentially expressed proteins. CAA, cerebral amyloid angiopathy; FC, fold change; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PCA, principal component analysis.

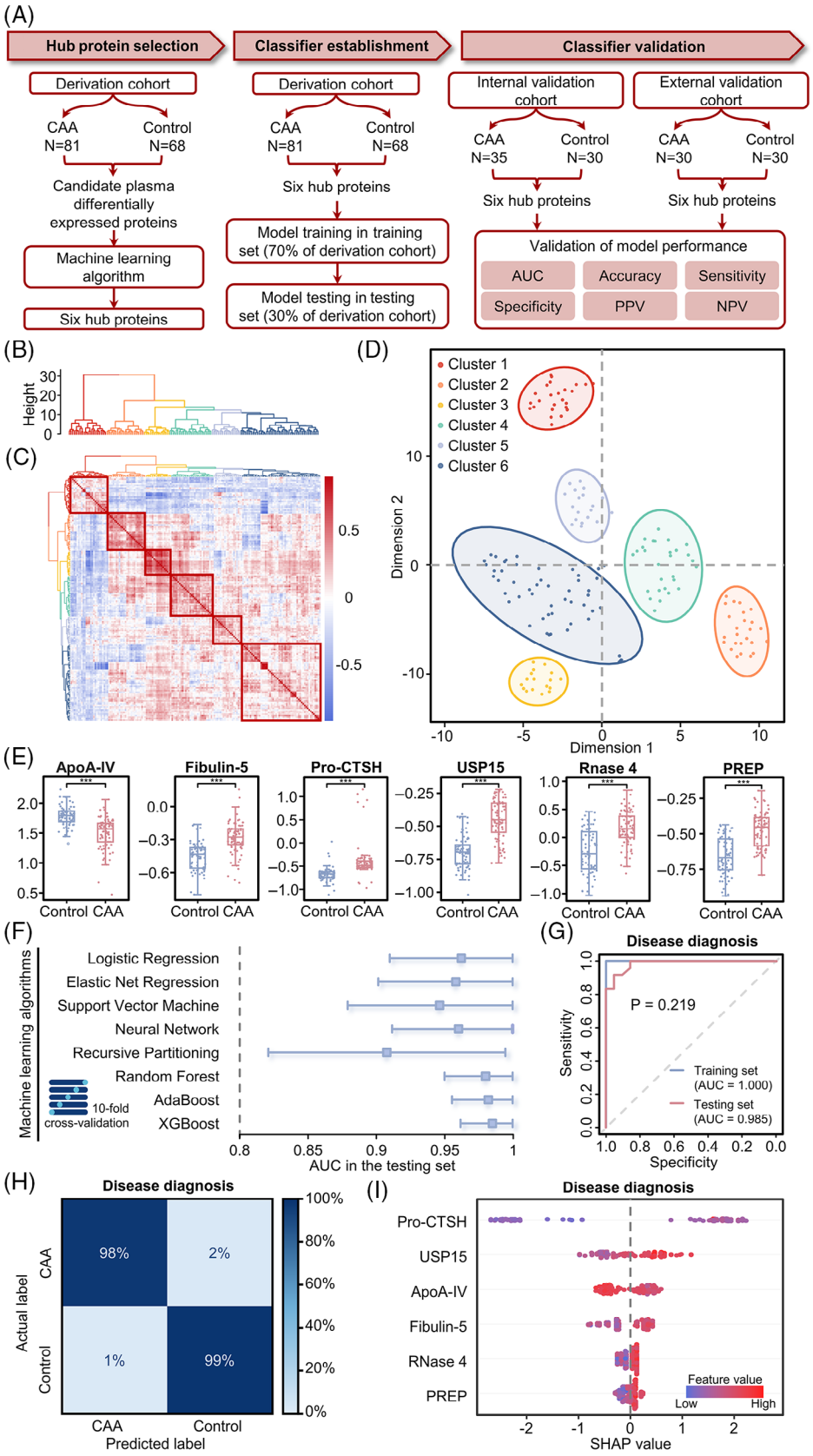

FIGURE 3 Discovery of circulating biomarkers and establishment of diagnosis model. (A) Workflow of hub protein selection and diagnosis model establishment and validation in the study. (B) Dendrogram showing results of agglomerative hierarchical clustering of proteins based on pairwise correlations. Each column denotes a protein and proteins in the same cluster are depicted in one color. (C) Heatmap showing pairwise correlations among differentially expressed proteins. Red squares represent clusters identified by agglomerative hierarchical clustering. Each row and column denote a protein; the order of proteins corresponds to the dendrogram in (B). (D) t-SNE plot displaying differentially expressed

significant difference between patients and controls (Table S4). Of six hub proteins, fibulin-5, pro-cathepsin H (Pro-CTSH), ubiquitin carboxyl-terminal hydrolase 15 (USP15), ribonuclease 4 (RNASE 4), and prolyl endopeptidase (PREP) were significantly upregulated in patients with CAA compared to controls, while apolipoprotein A-IV (ApoA-IV) was downregulated in CAA (Figure 3E and Table S4).

Using these six hub proteins, we developed several classifiers trained on 70% of the derivation cohort (the training set) and validated on the remaining 30% (the testing set) to distinguish patients from controls. As shown in Figure 3F, all classifiers demonstrated excellent diagnostic performance in the testing set, with AUCs exceeding 0.8 (Figure 3F and Table S5). Among these classifiers, the XGBoost model showed the best performance in both training and testing sets, with AUCs of 1.000 and 0.985, respectively (Figure 3G and Table S5). Other performance metrics are listed in Table S6. The XGBoost model also demonstrated a sensitivity of 98% and a specificity of 99% in the whole derivation cohort, as shown by the confusion matrix (Figure 3H). Furthermore, the SHAP value was applied to estimate the contribution of each hub protein to the diagnostic model. As displayed in Figure 3I, the relative importance of Pro-CTSH, USP15, ApoA-IV, Fibulin-5, RNASE 4, and PREP gradually decreased across the derivation cohort. Similar findings were observed in the training and testing sets, respectively (Figure S4). Taken together, these findings indicate that the panel of six hub proteins could be useful for differentiating CAA from controls, and these proteins might capture essential changes in disease. Ultimately, the XGBoost model was selected as the final classifier based on its comprehensive diagnostic performance.

3.3 | Development of risk stratification system with hub proteins

Next, whether these circulating hub protein panels could be used for risk stratification and prognostic prediction in CAA was further investigated. As shown in Figure 4A, 145 patients with CAA in the study were followed up for a median (IQR) time of 39 (21 to 59) months, with 29 (20%) patients developing the primary outcome (16 in the derivation cohort, nine in the internal validation cohort, and four in the external validation cohort), and one patient in the external validation was lost to follow-up. In the derivation cohort, patients with primary outcome were older (75.06 vs 68.37, $p = 0.008$) and had a more severe burden of SVD markers, including cSS scale (1 vs 0, $p = 0.009$), CSO-PVS (3 vs 2, $p = 0.003$), and total MRI burden of SVD (5 vs 3, $p = 0.002$; Table S7). In contrast, demographic characteristics and neuroimaging markers were comparable between two validation cohorts, probably due to their small sample sizes (Table S8).

In terms of prognostic prediction, the circulating hub protein panel, combined with age and sex, demonstrated excellent predictive ability in the derivation cohort using XGBoost algorithm, with an AUC of 0.977, a sensitivity of 95%, and a specificity of 100% (Figure 4B, C and Table S9). Then the individual risk score was calculated for each patient in the derivation cohort, and patients were stratified into a high-risk group and a low-risk group according to the cutoff value of risk scores (defined by the Youden index). It is noteworthy that patients in the high-risk group showed more severe total MRI burden of SVD (4 vs 3, $p = 0.009$) and CSO-PVS (3 vs 2, $p = 0.048$) than those in the low-risk group, after adjustment for age and sex (Figure 4D), further supporting the performance of the risk stratification system. With regard to prognostic prediction, patients in the high-risk group had an increased risk for new-onset CAA-ICH events (log-rank test, $p < 0.001$; Figure 4E).

Furthermore, a sensitivity analysis was conducted to explore the performance of the hub protein panel in prognostic assessment in addition to neuroimaging markers in the derivation cohort. The neuroimaging marker model included age, cSS scale, CSO-PVS, and total MRI burden of SVD. The hub protein panel model did not outperform the neuroimaging marker model alone (NRI = 0.202, $p = 0.191$; IDI = 0.031, $p = 0.687$; Table 2). However, the combined model integrating hub proteins with the neuroimaging markers significantly improved integrated discrimination abilities indicated by NRI and IDI (NRI = 0.370, $p = 0.007$; IDI = 0.175, $p = 0.003$), showing the potential prognostic value of the hub protein panel.

3.4 | Internal and external validation of circulating proteomic panel for disease screening and risk stratification

To further verify the diagnosis performance and risk stratification ability of the hub protein panel, the same classifier model and risk stratification system were applied to the internal and external validation cohorts. First, similar expression patterns of plasma hub proteins were observed in patients with CAA in the internal validation cohort (Figure 5A and Table S4). In the external validation cohort, targeted proteomic analysis showed consistent expression levels for four out of six hub proteins, whereas the remaining two proteins did not exhibit significant differences (Figure 5E and Table S4).

proteins clustered by identified clusters. Each point denotes a protein, clusters are indicated by colored shading. (E) Comparisons of expressions of six hub proteins in patients with CAA and controls, adjusted for age and sex. Each point represents a sample. (F) Forest plot showing AUCs using different machine learning algorithms combined with 10-fold cross-validation in testing set of derivation cohort. (G) ROC curves showing performance of hub protein panel with XGBoost algorithm in classifying CAA and controls in training and testing sets. (H) Confusion matrix illustrating performance of XGBoost classifier in classifying CAA and controls in whole derivation cohort. (I) SHAP value plot showing relative importance of six hub proteins in disease diagnosis in whole derivation cohort. Hub proteins are ordered by importance in classifying CAA and controls. Each point represents a sample. ApoA-IV, apolipoprotein A-IV; AUC, area under the curve; CAA, cerebral amyloid angiopathy; NPV, negative predictive value; PPV, positive predictive value; PREP, prolyl endopeptidase; Pro-CTSH, pro-cathepsin H; RNASE 4, ribonuclease 4; ROC, receiver operating characteristic; t-SNE, t-distributed stochastic neighbor embedding; USP15, ubiquitin carboxyl-terminal hydrolase 15; AdaBoost, Adaptive Boosting; XGBoost, eXtreme Gradient Boosting; SHAP, SHapley Additive exPlanations.

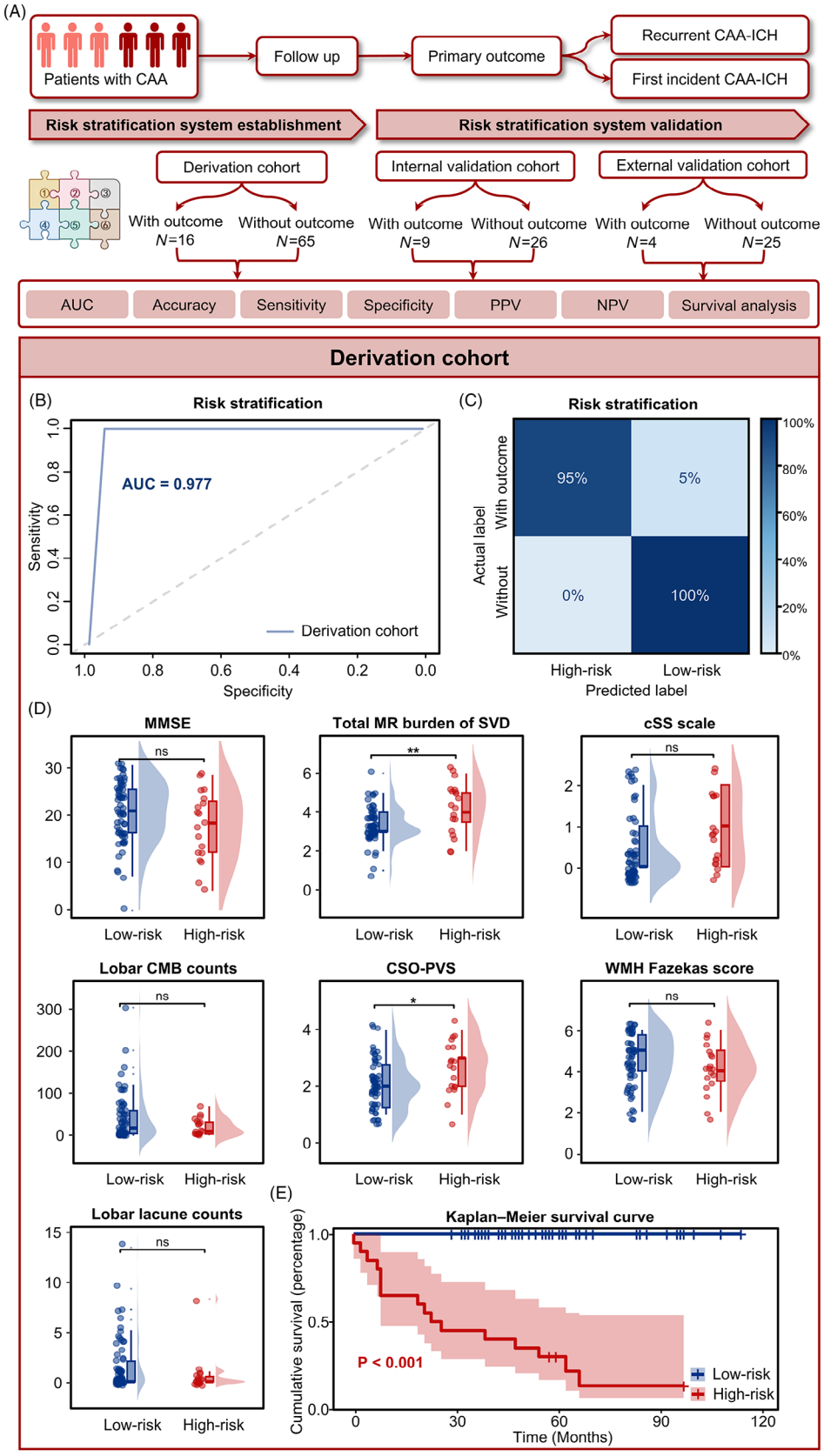

FIGURE 4 Follow-up strategy and construction of risk stratification system. (A) Workflow of follow-up strategy in patients with CAA in the study. (B) ROC curve showing performance of hub protein panel with XGBoost algorithm in predicting primary outcome in whole derivation cohort. (C) Confusion matrix illustrating predictive ability of hub protein panel with XGBoost algorithm for primary outcome in whole derivation cohort. (D) Comparisons of global cognitive function and neuroimaging markers in patients with CAA between low-risk and high-risk groups stratified by

TABLE 2 Performance metrics of models in predicting primary outcome.

	Neuroimaging marker model	Protein panel model	Combined model
NRI			
Derivation cohort	Reference	0.202 (−0.100 to 0.504)	0.370 (0.103 to 0.637)
Internal validation cohort	Reference	0.549 (0.287 to 0.811)	0.549 (0.287 to 0.811)
External validation cohort	Reference	0.350 (−0.129 to 0.829)	1.000 (1.000 to 1.00)
IDI			
Derivation cohort	Reference	0.031 (−0.118 to 0.180)	0.175 (0.060 to 0.290)
Internal validation cohort	Reference	0.615 (0.445 to 0.785)	0.615 (0.445 to 0.785)
External validation cohort	Reference	0.250 (−0.018 to 0.518)	0.991 (0.962 to 1.019)

Notes: AUC, NRI, and IDI were presented as value (95% CI). Imaging marker model was established based on age, cSS scale, CSO-PVS, and total MRI burden of SVD. Protein panel model was established based on age and six hub proteins. Combined model was established based on age, cSS scale, CSO-PVS, total MRI burden of SVD, and six hub proteins.

Abbreviations: AUC, area under curve; CSO-PVS, enlarged perivascular spaces in centrum semiovale; IDI, integrated discrimination improvement; NRI, net reclassification improvement; SVD, small vessel disease.

Next, the aforementioned XGBoost classifier was applied to the internal and external validation cohorts, yielding AUCs of 0.961 and 0.942 for disease diagnosis, respectively (Figure 5B, F). Other performance metrics, including accuracy, sensitivity, specificity, PPV, and NPV, also demonstrated the ability of the XGBoost model based on hub protein panel to distinguish between patients with CAA and controls (Figure 5C, G and Table S10). The SHAP value indicated the same relative importance of hub proteins in the classifier (Figure S5).

As to the validation of the risk stratification system, the predictive model based on the hub protein panel, age, and sex was used to predict the primary outcome in both validation cohorts, achieving AUCs of 0.808 and 0.740, respectively (Figure 5B, F and Table S9). The prognostic predictive performance of each hub protein indicated no significant difference in either validation cohort (Figure S6). Similar risk scores were obtained and stratified patients into the high-risk and low-risk groups. In validation cohorts, compared to patients in the low-risk group, those in the high-risk group were also associated with a higher risk of suffering from new-onset CAA-ICH (log-rank test, internal validation cohort: $p = 0.009$, external validation cohort: $p < 0.001$; Figure 5D, H), consistent with observations in the derivation cohort. However, the global cognitive function and neuroimaging markers were not significantly different between the two groups, except for a less severe degree of CSO-PVS in the high-risk group in the internal validation cohort and a higher count of lobar lacunes in those in the external validation cohort (Figure S7).

In the sensitivity analysis, the addition of six hub proteins to neuroimaging markers also improved prognosis reclassification capabilities in the internal validation cohort (NRI = 0.549, $p < 0.001$; IDI = 0.615, $p < 0.001$; Table 2) and the external validation cohort (NRI = 1.000, $p < 0.001$; IDI = 0.991, $p < 0.001$; Table 2). However, given the relatively small sample size of patients in the external validation cohort, the results should be interpreted with caution. Taken together, findings in two validation cohorts further underscore the potential efficacy of the circulating panel for helping CAA screening and risk stratification.

4 | DISCUSSION

Identifying novel biomarkers is crucial not only for understanding CAA pathogenesis but also for improving prognostic prediction. In this study, we comprehensively explored the plasma protein alterations in patients with CAA through proteomics. Notably, we developed and validated a circulating hub protein panel with machine learning algorithms for disease screening and risk stratification. The hub protein panel effectively distinguished CAA from controls and stratified patients based on their potential risk of new-onset CAA-ICH events. The consistency of these findings across derivation and validation cohorts reinforced the robustness and reliability of our results.

To date, few studies have systematically investigated blood-based biomarkers in CAA. Given their non-invasiveness and accessibility, blood-based biomarkers could serve as promising adjuncts to facilitate disease diagnosis, rather than a replacement for the Boston criteria.[37] Several studies have reported reduced plasma $A\beta_{40}$ and $A\beta_{42}$ during the presymptomatic stage in hereditary D-CAA.[14,38] However, studies with sporadic CAA have yielded inconsistent results,[39–42] likely due to differences in detection methods, disease stages, and sample sizes. In contrast, our study addressed these limitations by utilizing advanced proteomic techniques, enrolling patients with different disease severities, and validating findings across two independent cohorts.

risk scores in derivation cohort, adjusted for age and sex. Each point represents a sample. (E) Kaplan–Meier survival curve displaying risk of primary outcome between low-risk and high-risk groups stratified by risk scores in 81 CAA patients with follow-up data in derivation cohort. AUC, area under the curve; CAA, cerebral amyloid angiopathy; CMB, cerebral microbleed; CSO-PVS, centrum semiovale enlarged perivascular space; cSS, cortical superficial siderosis; ICH, intracerebral hemorrhage; MMSE, Mini-Mental State Examination; NPV, negative predictive value; PPV, positive predictive value; ROC, receiver operating characteristic; SVD, small vessel diseases; WMH, white matter hyperintensity.

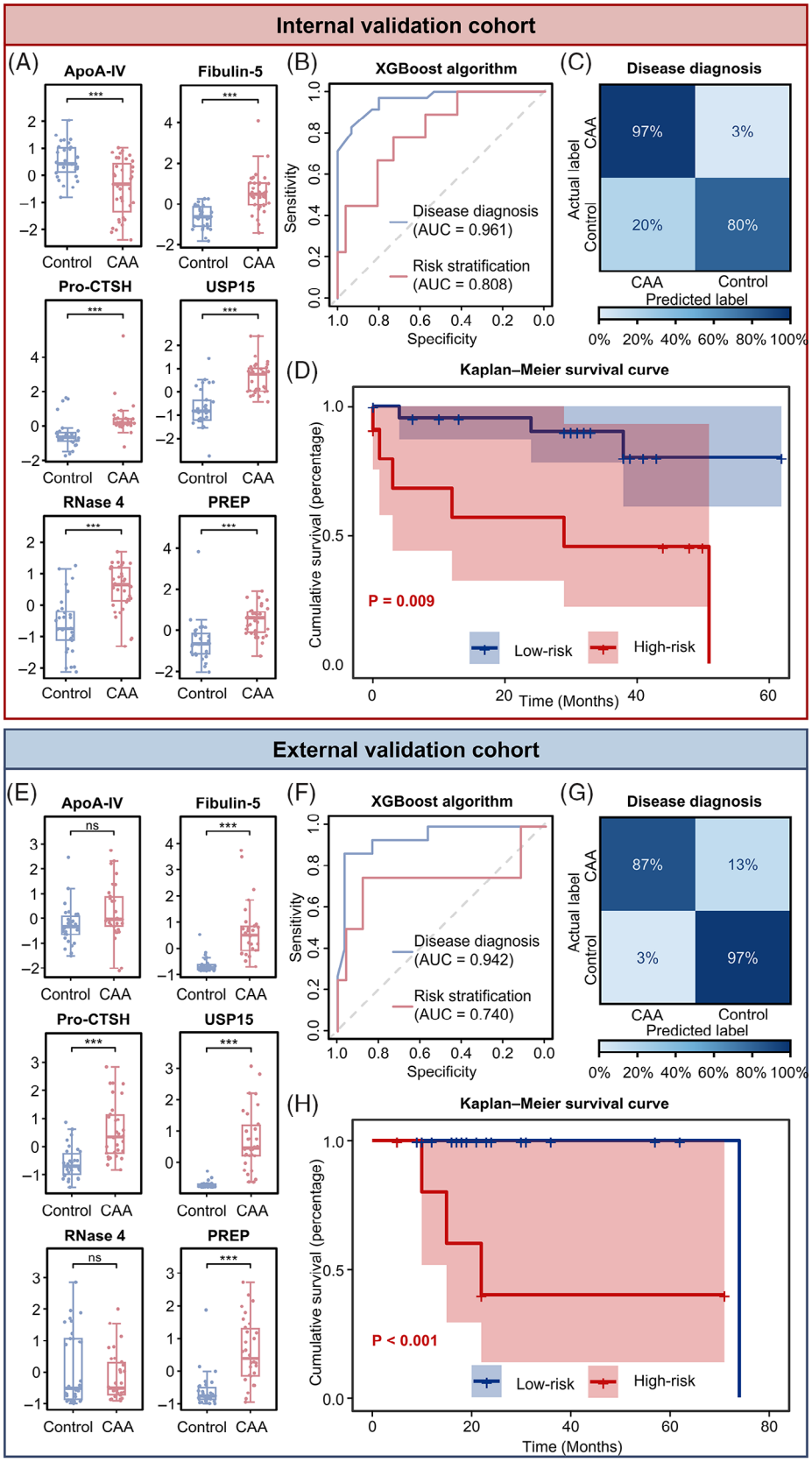

FIGURE 5 Internal and external validations of hub protein panel in disease diagnosis and risk stratification. (A) Comparisons of expressions of six hub proteins in patients with CAA and controls in internal validation cohort, adjusted for age and sex. Each point represents a sample. (B) ROC curves showing performance of hub protein panel with XGBoost algorithm in classifying CAA and controls, as well as predicting primary outcome in CAA in internal validation cohort. The blue curve denotes the AUC for disease diagnosis, and the pink curve denotes the AUC for risk

These approaches enhance our understanding of CAA pathogenesis and identify potential biomarkers that might complement existing screening methods.

Our DIA-based proteomic analysis identified 166 differentially expressed plasma proteins in patients with CAA. The functional enrichment analysis linked these proteins to biological processes, including immune response, response to stimuli, cell adhesion, and protein–lipid complex remodeling. A recent review highlighted the role of the immune system in CAA pathogenesis,[43] which was particularly evident in patients with CAA-related inflammation.[44] In addition, our previous studies demonstrated altered expression of MMPs and CHI3L1 in patients with CAA-ICH, which were associated with recurrence of lobar ICH,[12,13] supporting the role of neuroinflammation in disease progression. Cell adhesion is vital for maintaining the structural integrity and function of blood-brain barrier (BBB).[45] For instance, tight junction proteins are vital components of endothelial cell adhesion, ensuring the selective permeability and protective function of the BBB.[46] However, complement activation and subsequent altered expression of tight junction proteins in CAA could lead to BBB dysfunction.[47] Notably, alterations in lipids and lipoproteins are gradually recognized in CAA,[48] with previous studies finding alterations in apolipoproteins and cholesterol levels in patients with CAA-ICH.[48] Since apolipoprotein E (ApoE) is closely related to Aβ metabolism, its alteration may impact Aβ aggregation and deposition through lipid remodeling. Therefore, our proteomic findings reveal the complex and multifaceted potential pathophysiological processes of CAA.

In this study, hierarchical agglomerative clustering identified proteins with distinct expression patterns and ultimately selected six hub proteins. The classifier established using these proteins demonstrated its utility in disease screening, with SHAP values indicating their relative importance. Most of these proteins have been reported to be associated with Aβ pathology or neuroinflammation. Among them, Pro-CTSH, the inactive precursor form of cathepsin H (CTSH), ranked highest in relative importance across three cohorts. Pro-CTSH is a lysosomal cysteine protease, which is important for protein degradation and control of cellular energy metabolism.[49] Cathepsins have been implicated in the degradation of Aβ and neuroinflammation.[50,51] Recent multiomic analysis revealed that elevated expression of CTSH in the brain was associated with an increased risk of Alzheimer's disease.[52] Although the link between Pro-CTSH and CAA has not been reported previously, the upregulation of Pro-CTSH observed in our study might suggest enhanced lysosomal activity in response to Aβ accumulation and later neuroinflammation. USP15, a deubiquitinating enzyme, is proposed to regulate type I IFN response and neuroinflammation.[53] Upregulation of USP15 could activate the NF-κB signaling pathway, which might be involved in CAA pathology.[54,55] ApoA-IV is important in lipoprotein metabolism and regulation of plasma lipid transport. In humans, genes encoding ApoA-I, ApoC-III, and ApoA-IV on chromosome 11 tend to be co-regulated.[56] Redistribution of ApoC-III in plasma lipoproteins has been revealed in patients with CAA-ICH.[48] In addition, decreased expression levels of ApoA-IV might impact astrocyte-mediated Aβ clearance, leading to increased Aβ deposition and impaired cognitive function.[57] PREP is a serine peptidase that cleaves peptide bonds at proline residues.[58] According to previous studies, PREP might be involved in the cleavage of amyloid precursor protein (APP) and the production of Aβ.[59] The two remaining proteins, to our knowledge, have not been previously reported in CAA or Aβ pathology. They might contribute through mechanisms related to vascular integrity and immune response. Fibulin-5 is an extracellular matrix protein essential for cell adhesion and elastic fiber assembly.[60,61] Fibulin-5 was upregulated in vascular injury and cerebral ischemia-reperfusion injury, potentially playing a role in combating oxidative stress and improving BBB permeability.[62–64] Therefore, its elevation in CAA could reflect a compensatory response to oxidative stress, rather than a driving factor. RNase 4 belongs to the RNase A superfamily, yet its biological function remains unclear. This protein might be related to host defense and adaptive immune responses,[65,66] indicating an association with neuroinflammation.

Although no direct interactions between these hub proteins have been identified, they seem to interact within overlapping pathways, including the production and clearance of Aβ, vascular remodeling, neuroinflammation, and oxidative stress involved in CAA pathology. For instance, PREP, Pro-CTSH, and ApoA-IV might impact Aβ production and clearance, while Fibulin-5, USP15, and RNase 4 could play roles in regulating vascular integrity and neuroinflammatory processes. Interactions between neuroinflammation and impaired Aβ clearance could further amplify vascular injury and Aβ accumulation.[67] Future studies should aim to uncover the potential mechanisms of these proteins using rodent models, especially their roles in Aβ metabolism, vascular remodeling, and neuroinflammation. In addition, integrating proteomics with other -omics technologies, such as metabolomics and lipidomics, could provide further insights into the

stratification. (C) Confusion matrix illustrating predictive ability of hub protein panel with XGBoost algorithm for primary outcome in internal validation cohort. (D) Kaplan–Meier survival curve displaying risk of primary outcome between low-risk and high-risk groups stratified by risk scores in 35 CAA patients with follow-up data in internal validation cohort. (E) Comparisons of expressions of six hub proteins in patients with CAA and controls in external validation cohort, adjusted for age and sex. Each point represents a sample. (F) ROC curves showing performance of hub protein panel with XGBoost algorithm in classifying CAA and controls, as well as predicting primary outcome in CAA in external validation cohort. The blue curve denotes the AUC for disease diagnosis, and the pink curve denotes the AUC for risk stratification. (G) Confusion matrix illustrating predictive ability of hub protein panel with XGBoost algorithm for primary outcome in external validation cohort. (H) Kaplan–Meier survival curve displaying risk of primary outcome between low-risk and high-risk groups stratified by risk scores in 29 CAA patients with follow-up data in external validation cohort. ApoA-IV, apolipoprotein A-IV; AUC, area under the curve; CAA, cerebral amyloid angiopathy; PREP, prolyl endopeptidase; Pro-CTSH, pro-cathepsin H; RNASE 4, ribonuclease 4; ROC, receiver operating characteristic; USP15, ubiquitin carboxyl-terminal hydrolase 15; XGBoost, eXtreme Gradient Boosting.

regulatory networks and interactions among these proteins. Moreover, longitudinal tracking of these proteins could reveal their temporal dynamics during disease progression and support their potential application as biomarkers. These efforts could deepen our understanding of the molecular mechanisms underlying CAA and facilitate the development of novel therapeutic strategies.

Despite tight control of risk factors, such as blood pressure, patients with CAA still face a certain risk of ICH. Therefore, timely and accurate identification of high-risk patients is crucial for clinical management. Using longitudinal data, our study established a risk stratification system based on the hub protein panel and effectively stratified patients into high-risk and low-risk groups. High-risk patients exhibited a significantly increased risk of new-onset CAA-ICH, revealing the potential of this circulating hub protein panel as an ancillary tool for CAA screening and risk stratification. Across three cohorts, MMSE scores were comparable between the high-risk and low-risk groups, probably due to the weak association between global cognitive function and the risk of ICH occurrence. Moreover, while differences in certain neuroimaging markers between the high-risk and low-risk groups were observed, the small sample sizes of the two validation cohorts limited the robustness of these findings, warranting further validation in larger cohorts. In addition, the sensitivity analysis revealed that the addition of hub proteins improved the predictive efficacy of neuroimaging biomarkers. These findings suggest that hub proteins might provide additional information beyond neuroimaging markers in prognostic prediction. The risk stratification system based on the hub protein panel may have implications for clinical practice, for instance, guiding clinical management and patient selection for future trials.

Although our results provide novel insights into biomarker identification and mechanism exploration in CAA, the study has several limitations. First, the sample sizes of the two validation cohorts were small and only Chinese populations were included. Despite this, the consistent differential expression of the hub proteins and their performance in validation cohorts strongly suggest their relevance in disease pathology and highlight their potential as biomarkers. Data from multiple centers and larger sample sizes is warranted. Second, this study only enrolled community-dwelling controls and patients with CAA, without including other disease controls, such as arteriolosclerosis. The utility of the hub protein panel for the differential diagnosis of CAA will be investigated in future studies. In addition, several patients had missing cognitive data due to symptoms such as disorders of consciousness. To address this issue, the widely used RF method was applied to impute the missing values. Moreover, given the accessibility and generalizability of biomarkers, this study focused on plasma samples rather than CSF samples or brain tissue. Notably, the hub protein panel maintained its robust performance across all three cohorts, demonstrating its clinical feasibility. Finally, in the external validation cohort, targeted proteomics was performed to detect hub plasma proteins. Since omic technologies are not yet widely implemented in clinical practice, future efforts should explore more accessible detection methods to enhance clinical applicability.

In conclusion, this study identified significant plasma protein alterations in patients with CAA using proteomic analysis. We developed and validated a circulating hub protein panel with machine learning algorithms for CAA screening and risk stratification. Considering the importance of assessing future ICH risk in patients with CAA, these hub proteins not only offer insights into disease pathogenesis but also represent a promising tool for clinical evaluation. Our findings highlight the potential of plasma biomarkers to enhance our understanding of CAA and improve disease risk assessment.

ACKNOWLEDGMENTS

The authors thank all the participants who took part in our cohorts.

This study was funded by the National Natural Science Foundation of China (81971123, 82201467), Noncommunicable Chronic Diseases-National Science and Technology Major Project (2023ZD0504903), Shanghai Municipal Health Commission (2022XD022), Sailing Program of Shanghai Science and Technology Committee (22YF1405500), Shanghai "Rising Stars of Medical Talents" Youth Development Program (SHWSRS (2023)_062), CAMS Innovation Fund for Medical Sciences (CIFMS)(2024-I2M-C&T-B-007), and Brain Science and Brain Diseases Youth Innovation Program of Shanghai Zhou Liangfu Medical Development Foundation.

All participants or their legal representatives provided written informed consent.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest. Author disclosures are available in the supporting information.

ORCID

Jiajie Xu https://orcid.org/0000-0001-6819-4487
Xin Cheng https://orcid.org/0000-0001-7816-0547

REFERENCES

1. Qureshi AI, Tuhrim S, Broderick JP, Batjer HH, Hondo H, Hanley DF. Spontaneous intracerebral hemorrhage. *N Engl J Med*. 2001;344:1450-1460.
2. Fandler-Hofler S, Obergottsberger L, Ambler G, et al. Association of the presence and pattern of MRI markers of cerebral small vessel disease with recurrent intracerebral hemorrhage. *Neurology*. 2023;101:e794-e804.
3. Charidimou A, Imaizumi T, Moulin S, et al. Brain hemorrhage recurrence, small vessel disease type, and cerebral microbleeds: a meta-analysis. *Neurology*. 2017;89:820-829.
4. Knudsen KA, Rosand J, Karluk D, Greenberg SM. Clinical diagnosis of cerebral amyloid angiopathy: validation of the Boston criteria. *Neurology*. 2001;56:537-539.
5. Linn J, Halpin A, Demaerel P, et al. Prevalence of superficial siderosis in patients with cerebral amyloid angiopathy. *Neurology*. 2010;74:1346-1350.
6. Charidimou A, Boulouis G, Frosch MP, et al. The Boston criteria version 2.0 for cerebral amyloid angiopathy: a multicentre, retrospective, MRI-neuropathology diagnostic accuracy study. *Lancet Neurol*. 2022;21:714-725.
7. Switzer AR, Charidimou A, McCarter S, et al. Boston criteria v2.0 for cerebral amyloid angiopathy without hemorrhage: an MRI-neuropathologic validation study. *Neurology*. 2024;102:e209386.
8. Koemans EA, Rasing I, Voigt S, et al. Temporal ordering of biomarkers in Dutch-type hereditary cerebral amyloid angiopathy. *Stroke*. 2024;55:954-962.

9. Charidimou A, Friedrich JO, Greenberg SM, Viswanathan A. Core cerebrospinal fluid biomarker profile in cerebral amyloid angiopathy: a meta-analysis. *Neurology*. 2018;90:e754-e762.
10. Rasing I, Voigt S, Koemans EA, et al. Serum and cerebrospinal fluid neurofilament light chain and glial fibrillary acid protein levels in early and advanced stages of cerebral amyloid angiopathy. *Alzheimers Res Ther*. 2024;16:86.
11. Cheng X, Su Y, Wang Q, et al. Neurofilament light chain predicts risk of recurrence in cerebral amyloid angiopathy-related intracerebral hemorrhage. *Aging (Albany NY)*. 2020;12:23727-23738.
12. Xia M, Su Y, Fu J, et al. The use of serum matrix metalloproteinases in cerebral amyloid angiopathy-related intracerebral hemorrhage and cognitive impairment. *J Alzheimers Dis*. 2021;82:1159-1170.
13. Xu F, Xu J, Wang Q, et al. Serum YKL-40 as a predictive biomarker of cerebral amyloid angiopathy-related intracerebral hemorrhage recurrence. *J Alzheimers Dis*. 2024;99:503-511.
14. de Kort AM, Kuiperij HB, Jakel L, et al. Plasma amyloid beta 42 is a biomarker for patients with hereditary, but not sporadic, cerebral amyloid angiopathy. *Alzheimers Res Ther*. 2023;15:102.
15. Vervuurt M, Schrader JM, de Kort AM, et al. Cerebrospinal fluid shotgun proteomics identifies distinct proteomic patterns in cerebral amyloid angiopathy rodent models and human patients. *Acta Neuropathol Commun*. 2024;12:6.
16. Xia Y, Shen Y, Wang Y, et al. White matter hyperintensities associated with progression of cerebral small vessel disease: a 7-year Chinese urban community study. *Aging (Albany NY)*. 2020;12:8506-8522.
17. Xu J, Su Y, Fu J, et al. Glymphatic dysfunction correlates with severity of small vessel disease and cognitive impairment in cerebral amyloid angiopathy. *Eur J Neurol*. 2022;29:2895-2904.
18. Charidimou A, Boulouis G, Roongpiboonsopit D, et al. Cortical superficial siderosis and recurrent intracerebral hemorrhage risk in cerebral amyloid angiopathy: large prospective cohort and preliminary meta-analysis. *Int J Stroke*. 2019;14:723-733.
19. Charidimou A, Boulouis G, Xiong L, et al. Cortical superficial siderosis and first-ever cerebral hemorrhage in cerebral amyloid angiopathy. *Neurology*. 2017;88:1607-1614.
20. Duering M, Biessels GJ, Brodtmann A, et al. Neuroimaging standards for research into small vessel disease-advances since 2013. *Lancet Neurol*. 2023;22:602-618.
21. Stekhoven DJ, Buhlmann P. MissForest–non-parametric missing value imputation for mixed-type data. *Bioinformatics*. 2012;28:112-118.
22. Jiang Y, Zhou X, Ip FC, et al. Large-scale plasma proteomic profiling identifies a high-performance biomarker panel for Alzheimer's disease screening and staging. *Alzheimers Dement*. 2022;18:88-102.
23. Tibshirani R, Walther G, Hastie T. Estimating the number of clusters in a data set via the gap statistic. *J Royal Stat Soc B*. 2001;63:411-423.
24. Cieslak MC, Castelfranco AM, Roncalli V, Lenz PH, Hartline DK. t-Distributed Stochastic Neighbor Embedding (t-SNE): a tool for eco-physiological transcriptomic analysis. *Mar Genomics*. 2020;51:100723.
25. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. Springer; 2009.
26. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J Royal Stat Soc B*. 2005;67:301-320.
27. Vapnik V. The support vector method of function estimation. In: Suykens JAK, Vandewalle J, eds. *Nonlinear Modeling: Advanced Black-Box Techniques*. Springer US; 1998:55-85.
28. Gurney K. *An Introduction to Neural Networks*. CRC Press; 1997.
29. Therneau T, Atkinson E. An introduction to recursive partitioning using the RPART routines. *Mayo Clin*. 1997;61.
30. Breiman L. Random forests. *Machine Learn*. 2001;45:5-32.
31. Freund Y, Schapire RE. A Short Introduction to Boosting. *Journal of Japanese Society of Artificial Intelligence*. 1999;14:771-780.
32. Chen T, Guestrin C. XGBoost: a Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, California, USA: Association for Computing Machinery, 2016: 785-794.
33. Lundberg SM, Nair B, Vavilala MS, et al. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nat Biomed Eng*. 2018;2:749-760.
34. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44:837-845.
35. Youden WJ. Index for rating diagnostic tests. *Cancer*. 1950;3:32-35.
36. Alba AC, Agoritsas T, Walsh M, et al. Discrimination and calibration of clinical prediction models: users' guides to the medical literature. *JAMA*. 2017;318:1377-1384.
37. Muir RT, Ismail Z, Black SE, Smith EE. Comparative methods for quantifying plasma biomarkers in Alzheimer's disease: implications for the next frontier in cerebral amyloid angiopathy diagnostics. *Alzheimers Dement*. 2024;20:1436-1458.
38. Chatterjee P, Tegg M, Pedrini S, et al. Plasma amyloid-beta levels in a pre-symptomatic Dutch-type hereditary cerebral amyloid angiopathy pedigree: a cross-sectional and longitudinal investigation. *Int J Mol Sci*. 2021;22:2931.
39. Piccarducci R, Caselli MC, Zappelli E, et al. The role of amyloid-beta, tau, and alpha-Synuclein proteins as putative blood biomarkers in patients with cerebral amyloid angiopathy. *J Alzheimers Dis*. 2022;89:1039-1049.
40. Greenberg SM, Cho HS, O'Donnell HC, et al. Plasma beta-amyloid peptide, transforming growth factor-beta 1, and risk for cerebral amyloid angiopathy. *Ann N Y Acad Sci*. 2000;903:144-149.
41. Hernandez-Guillamon M, Delgado P, Penalba A, et al. Plasma beta-amyloid levels in cerebral amyloid angiopathy-associated hemorrhagic stroke. *Neurodegener Dis*. 2012;10:320-323.
42. Kim HJ, Park D, Yun G, et al. Screening for cerebral amyloid angiopathy based on serological biomarkers analysis using a dielectrophoretic force-driven biosensor platform. *Lab Chip*. 2021;21:4557-4565.
43. Munsterman D, Falcione S, Long R, et al. Cerebral amyloid angiopathy and the immune system. *Alzheimers Dement*. 2024;20(7):4999-5008.
44. Eng JA, Frosch MP, Choi K, Rebeck GW, Greenberg SM. Clinical manifestations of cerebral amyloid angiopathy-related inflammation. *Ann Neurol*. 2004;55:250-256.
45. Persidsky Y, Ramirez SH, Haorah J, Kanmogne GD. Blood-brain barrier: structural components and function under physiologic and pathologic conditions. *J Neuroimmune Pharmacol*. 2006;1:223-236.
46. Liu WY, Wang ZB, Zhang LC, Wei X, Li L. Tight junction in blood-brain barrier: an overview of structure, regulation, and regulator substances. *CNS Neurosci Ther*. 2012;18:609-615.
47. Hu M, Li T, Ma X, et al. Macrophage lineage cells-derived migrasomes activate complement-dependent blood-brain barrier damage in cerebral amyloid angiopathy mouse model. *Nat Commun*. 2023;14:3945.
48. Bonaterra-Pastra A, Fernandez-de-Retana S, Rivas-Urbina A, et al. Comparison of plasma lipoprotein composition and function in cerebral amyloid angiopathy and Alzheimer's disease. *Biomedicines*. 2021;9:7.
49. Reiser J, Adair B, Reinheckel T. Specialized roles for cysteine cathepsins in health and disease. *J Clin Invest*. 2010;120:3421-3431.
50. Mueller-Steiner S, Zhou Y, Arai H, et al. Antiamyloidogenic and neuroprotective functions of cathepsin B: implications for Alzheimer's disease. *Neuron*. 2006;51:703-714.
51. Cataldo AM, Nixon RA. Enzymatically active lysosomal proteases are associated with amyloid deposits in Alzheimer brain. *Proc Natl Acad Sci U S A*. 1990;87:3861-3865.
52. Ou YN, Yang YX, Deng YT, et al. Identification of novel drug targets for Alzheimer's disease by integrating genetics and proteomes from brain and blood. *Mol Psychiatry*. 2021;26:6065-6073.
53. Torre S, Polyak MJ, Langlais D, et al. USP15 regulates type I interferon response and is required for pathogenesis of neuroinflammation. *Nat Immunol*. 2017;18:54-63.

54. Zhou Q, Cheng C, Wei Y, et al. USP15 potentiates NF-kappaB activation by differentially stabilizing TAB2 and TAB3. *FEBS J*. 2020;287:3165-3183.
55. Vromman A, Trabelsi N, Rouxel C, Bereziat G, Limon I, Blaise R. beta-Amyloid context intensifies vascular smooth muscle cells induced inflammatory response and de-differentiation. *Aging Cell*. 2013;12:358-369.
56. Karathanasis SK. Apolipoprotein multigene family: tandem organization of human apolipoprotein AI, CIII, and AIV genes. *Proc Natl Acad Sci U S A*. 1985;82:6374-6378.
57. Cui Y, Huang M, He Y, Zhang S, Luo Y. Genetic ablation of apolipoprotein A-IV accelerates Alzheimer's disease pathogenesis in a mouse model. *Am J Pathol*. 2011;178:1298-1308.
58. Kalwant S, Porter AG. Purification and characterization of human brain prolyl endopeptidase. *Biochem J*. 1991;276(Pt 1):237-244.
59. Shinoda M, Toide K, Ohsawa I, Kohsaka S. Specific inhibitor for prolyl endopeptidase suppresses the generation of amyloid beta protein in NG108-15 cells. *Biochem Biophys Res Commun*. 1997;235:641-645.
60. Preis M, Cohen T, Sarnatzki Y, et al. Effects of fibulin-5 on attachment, adhesion, and proliferation of primary human endothelial cells. *Biochem Biophys Res Commun*. 2006;348:1024-1033.
61. Papke CL, Yanagisawa H. Fibulin-4 and fibulin-5 in elastogenesis and beyond: insights from mouse and human studies. *Matrix Biol*. 2014;37:142-149.
62. Guadall A, Orriols M, Rodriguez-Calvo R, et al. Fibulin-5 is up-regulated by hypoxia in endothelial cells through a hypoxia-inducible factor-1 (HIF-1alpha)-dependent mechanism. *J Biol Chem*. 2011;286:7093-7103.
63. Spencer JA, Hacker SL, Davis EC, et al. Altered vascular remodeling in fibulin-5-deficient mice reveals a role of fibulin-5 in smooth muscle cell proliferation and migration. *Proc Natl Acad Sci USA*. 2005;102:2946-2951.
64. Guo J, Cheng C, Chen CS, et al. Overexpression of Fibulin-5 attenuates ischemia/reperfusion injury after middle cerebral artery occlusion in rats. *Mol Neurobiol*. 2016;53:3154-3167.
65. Lu L, Li J, Moussaoui M, Boix E. Immune modulation by human secreted RNases at the extracellular space. *Front Immunol*. 2018;9:1012.
66. Bender K, Schwartz LL, Cohen A, et al. Expression and function of human ribonuclease 4 in the kidney and urinary tract. *Am J Physiol Renal Physiol*. 2021;320:F972-F983.
67. van den Brink H, Voigt S, Kozberg M, van Etten ES. The role of neuroinflammation in cerebral amyloid angiopathy. *EBioMedicine*. 2024;110:105466.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Xu J, Su Y, Sha Y, et al. Circulating proteomic biomarkers for cerebral amyloid angiopathy screening and risk stratification. *Alzheimer's Dement*. 2025;21:e70044. https://doi.org/10.1002/alz.70044