Journal of the American Heart Association

ORIGINAL RESEARCH

Cross-Platform Proteomics and Machine Learning Algorithms Nominate Plasma Biomarkers of Stroke Diagnosis

Shubham Misra, PhD; Wooyoung Eric Jang, PhD; Sebastian Sanchez, MD; Aditya Natu, MS; Prateek Kumar, PhD; Michael Liu, MD; Anmol Kaur, MBBS; Victor Torres Lopez, MS; Pinar Caglayan, BS; Rolando Garcia-Milian, MLS; Caroline M. Watson, PhD; Michael R. Frankel, MD; Guido J. Falcone, MD, ScD; Lauren H. Sansing, MD, MS; Srikant Rangaraju, MD, MS

BACKGROUND: Blood-based biomarkers for stroke subtyping could improve triage in emergency settings. We used cross-platform proteomics to identify plasma biomarkers differentiating major stroke diagnostic groups.

METHODS: We conducted a case–control study using 2 biorepositories. Plasma was collected in the emergency department from adults with suspected stroke before therapeutic intervention. Differentially enriched proteins were identified across acute ischemic stroke, intracerebral hemorrhage, transient ischemic attack, and stroke mimics using SomaScan discovery proteomics (Grady). Differentially enriched proteins were nominated using pairwise and multigroup comparisons and adjusted for clinical covariates. Protein panels were created using least absolute shrinkage and selection operator logistic regression. Internal validation used repeated nested cross-validation (rCV) and targeted mass spectrometry (MS), while external validation used data-independent acquisition mass spectrometry in an independent cohort (Yale).

RESULTS: We included 100 subjects (40 with acute ischemic stroke, 20 with intracerebral hemorrhage, 20 with transient ischemic attack, 20 with stroke mimics) in discovery and 80 subjects (20 per group) in external validation cohorts. SomaScan quantified 7307 proteins, of which 61 differentiated stroke subtypes. We identified 7 protein classifiers for acute ischemic stroke (rCV-area under the curve, 0.82 [95% CI, 0.78–0.86]), 6 for intracerebral hemorrhage (rCV-area under the curve, 0.70 [95% CI, 0.64–0.76]), 8 for transient ischemic attack (rCV-area under the curve, 0.78 [95% CI, 0.73–0.84]), and 7 for stroke mimics (rCV-area under the curve, 0.81 [95% CI, 0.77–0.86]). Targeted proteomics internally validated 11 proteins, and data-independent acquisition-mass spectrometry externally validated 32 proteins, including VTN (vitronectin), PLG (plasminogen), and S100A9 as top stroke mimics, transient ischemic attack, and intracerebral hemorrhage classifiers.

CONCLUSIONS: This study highlights plasma proteomics as a valuable tool for discovering protein biomarkers of stroke diagnosis. These findings support further validation in larger, multicenter cohorts to facilitate biomarker-guided stroke diagnosis in acute care.

Key Words: biomarkers ■ blood ■ diagnosis ■ differentiation ■ proteins ■ proteomics ■ stroke

Stroke is a leading cause of overall disease burden worldwide, accounting for high disability-adjusted life years.[1] Timely diagnosis of stroke and its subtypes is critical to determine appropriate treatment approaches, to improve functional outcomes, and to reduce mortality. The majority of patients presenting with acute neurological symptoms to emergency department settings can be broadly classified into 4

Correspondence to: Srikant Rangaraju, MD, MS, Department of Neurology, Yale University School of Medicine, 8300B, 300 George Street, New Haven CT 06511. Email: srikant.rangaraju@yale.edu

This article was sent to Shaan Khurshid, MD, MPH, Associate Editor, for review by expert referees, editorial decision, and final disposition.

Supplemental Material is available at https://www.ahajournals.org/doi/suppl/10.1161/JAHA.125.048249

For Disclosures and Sources of Funding: see page 17.

JAHA is available at: www.ahajournals.org/journal/jaha

CLINICAL PERSPECTIVE

What Is New?

- This study analyzed 7307 plasma protein targets using an aptamer-based platform from 100 individuals with suspected stroke diagnoses in the emergency setting and nominated 61 proteins that differentiated groups with acute ischemic stroke, intracerebral hemorrhage, transient ischemic attack, and stroke mimics, 11 of which were internally validated using targeted proteomics, and 32 were externally validated in an independent cohort of 80 patients with stroke using untargeted proteomics by mass spectrometry.

What Are the Clinical Implications?

- The findings highlight the potential of plasma proteomics as a valuable tool for discovering protein biomarkers of stroke diagnostic groups.
- These proteomic signatures, validated through cross-platform methods, have the potential to improve stroke diagnosis, guide early intervention strategies, and reduce diagnostic uncertainty in acute settings.

Nonstandard Abbreviations and Acronyms

AIS	acute ischemic stroke
BAK-270	biomarker assessment kit-270
DEP	differentially enriched proteins
DIA-MS	data-independent acquisition-mass spectrometry
GCS	Glasgow Coma Scale
ICH	intracerebral hemorrhage
LASSO	least absolute shrinkage and selection operator
MIM	stroke mimics
MS	mass spectrometry
NIHSS	National Institutes of Health Stroke Scale
rCV	repeated cross-validation

categories, namely acute ischemic stroke (AIS), intracerebral hemorrhage (ICH), transient ischemic attack (TIA), and stroke mimics (MIM).

A confident diagnosis needs a combination of clinical assessment and brain imaging using computed tomography (CT) or magnetic resonance imaging, often requiring time and significant health care use.[2] Time-sensitive treatment decisions are often made within the first few hours of presentation to the emergency department (ED). These include acute thrombolysis[3] and endovascular reperfusion treatments for AIS[4,5] and blood pressure reduction, anticoagulant reversal, and surgical evaluation for ICH.[6] Risk stratification of TIA can also guide urgent evaluation in the ED versus outpatient management. Lastly, a diagnosis of MIM, accounting for nearly 20% to 50% of acute neurological presentations,[7] is often one of exclusion and includes peripheral vertigo, migraine, seizures, and functional disorders.[8] As a result, many patients with acute neurological presentations undergo expedited and extensive testing before a diagnosis of MIM can be made, leading to significant health care use. Panels of blood-based biomarkers that differentiate these stroke groups may enhance clinical decision-making, complement CT imaging, and expedite stroke management, especially in resource-limited and time-constrained scenarios.

Several candidate biomarker studies using low-plex assays have identified potential blood biomarkers for differential stroke diagnosis[9–12]; however, none have been validated in clinical settings. There is a critical need to use high-throughput proteomics approaches that can simultaneously quantify thousands of proteins to provide an opportunity for unbiased exploratory biomarker nomination, ultimately leading to narrower panels of proteins that can be accurately and reliably quantified for clinical applications.[13,14] We undertook this cross-platform proteomics study to derive a set of plasma protein biomarkers associated with these distinct stroke diagnostic groups, focusing on plasma samples collected in the acute settings.

METHODS

Data Availability

Deidentified raw proteomic data with clinical covariates will be shared with investigators upon request. Lists of measured proteins, average abundances across the study, statistical parameters, and effect sizes are all provided in the supplemental data with this article. The targeted mass spectrometry (MS) data have been deposited to PeptideAtlas with the data set identifier: PASS05889. R codes used for analyses are available on the GitHub repository: https://github.com/Shubhammisra30/Plasma-Proteomics-Stroke.

Study Population

Discovery Cohort

We used an archived plasma repository that was collected at Grady Memorial Hospital, Atlanta, GA, USA. These samples had been prospectively collected from patients at the time of arrival as part of standard-of-care and used for a quality improvement project. In

2015, approval was obtained from the Institutional Review Board of Emory University (IRB00088621) with a waiver of consent to allow a retrospective study of this biorepository of 356 cases with banked plasma samples (Figure 1). We included patients aged ≥18 years who were evaluated in the ED from 2010 to 2014 with a suspected stroke diagnosis due to acute neurological symptoms and had at least 1 nonhemolyzed plasma sample frozen at −80 °C. Due to budgetary restrictions, we conducted convenience sampling and included 100 samples in the discovery cohort with a minimum sample size of 20 per group. The final diagnosis (AIS, ICH, TIA, or MIM) was adjudicated by a stroke neurologist based on medical chart review.

External Validation Cohort

We analyzed clinical and proteomics data from subjects with stroke aged ≥18 years, using blood samples from a prospective repository of acute brain injury (2014–2023) at Yale University (IRB1506016023). The biorepository included plasma samples from 234 subjects within 24 hours of stroke onset, comprising 133 with AIS, 51 with ICH, 27 with TIA, and 23 with MIM (Figure 1). From these, we selected 20 subjects per group, matched by age and sex. Upper limits of sample sizes for derivation and validation studies were limited by budgetary feasibility.

Data Collection

We collected baseline data on patient demographics (age, sex), vascular risk factors (hypertension, diabetes, coronary artery disease, atrial fibrillation, smoking), stroke mechanism, prior stroke, and recorded patient outcomes using Glasgow Coma Scale (GCS), modified Rankin Scale, and National Institutes of Health Stroke Scale (NIHSS) scores.

Objectives

Our goal was to identify differentially enriched proteins (DEPs) in patients with AIS, ICH, TIA, and MIM.

Sample Collection and Proteomics Procedures

Blood samples were drawn into EDTA-coated tubes at the time of ED presentation (for the discovery cohort) and within 24 hours (for the external validation cohort), before any treatments, centrifuged at 1500 *g* for 15 minutes at 4 °C, and plasma was stored as a single 2 to 4 mL aliquot at −80 °C until further analysis. In the discovery cohort, we performed aptamer-based proteomics using the plasma 7K SomaScan assay, as per the manufacturer's protocols, using technical duplicates.[15] We also conducted internal validation on the same samples using multiple reaction monitoring-based targeted proteomics.[16] We used the PeptiQuant Plus biomarker assessment kits (BAK-270) for targeted protein quantitation in undepleted and unenriched human plasma samples.[17,18] In the external validation cohort, we performed data-independent acquisition MS (DIA-MS) proteomics (liquid chromatography-MS/MS, Orbitrap Exploris 480). For DIA-MS, plasma was depleted using High-Select Top 14 Protein Depletion Resin as per the manufacturer's protocol. Full experimental details for MS are provided in the Supplemental Methods.[19–21]

Statistical Analysis

We assessed the normality of the data using the Shapiro–Wilk test. If normally distributed, we summarize continuous variables as mean±SD, otherwise, as median and interquartile range. We summarized categorical variables as number (percentage). We compared the baseline characteristics between the 4 groups using 1-way ANOVA or the Kruskal–Wallis test for continuous variables and the chi-square test for categorical variables. For pairwise comparisons, we used the *t* test or the Mann–Whitney *U* test to compare the difference in means or medians, respectively, between the 2 groups. The protein relative abundance values were $\log_2$-transformed.

Pairwise Comparisons

We performed pairwise comparisons and identified the DEPs using 3 protein selection criteria: (1) ±1.5-fold change and unadjusted *P* value <0.05 cutoffs (Welch's *t* test), (2) confirmed selection in Boruta random forest-based feature selection (Boruta package in R version 4.3.2),[22] and (3) ≥20% variation explained by stroke subtypes in the variance partitioning analysis (variancePartition R package). For all pairwise comparisons, we selected the top proteins that were commonly identified using 2 of the 3 protein selection approaches. We then conducted multivariable regularized logistic least absolute shrinkage and selection operator (LASSO) regression analysis, adjusting the top proteins for age, sex, NIHSS score, and GCS score in the univariable pairwise analysis (rlassologit function in the hdm R package).[23] We employed a regularized logistic regression method based on the double selection LASSO framework to minimize model overfitting and adjust for relevant covariates by shrinking less informative predictors to 0.[24] Before model fitting, collinearity among these variables was assessed using variance inflation factors, and no concerning collinearity was observed. Adjusted *P* values for the LASSO-selected predictors were then obtained using standard multivariable logistic regression. Additionally, we conducted the functional enrichment analysis using the Ingenuity Pathway Analysis software (QIAGEN Digital Insights, 2024) for all pairwise comparisons in the discovery cohort.

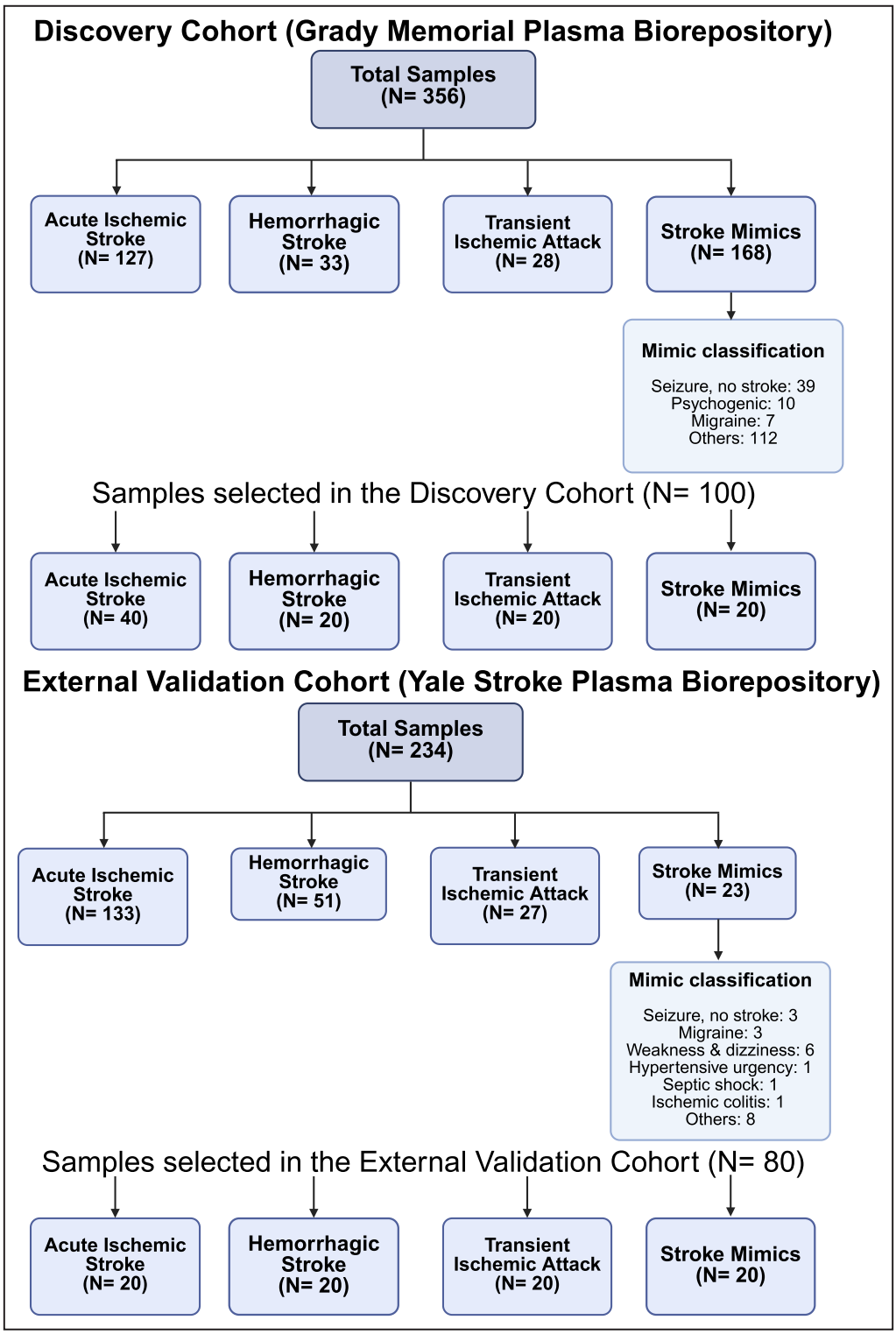

Figure 1. **Flow chart of patients included in the plasma repositories.**

We selected proteins with fold change ±1.2 and unadjusted *P* value <0.05 and compared the enriched pathways between all pairwise comparisons using the *Z*-score metric. The enriched pathways were identified using the following parameters: false discovery rate-corrected *P* value <0.1 and *Z*-score > 0.2.

Multigroup Comparisons

We used a supervised machine learning algorithm and performed feature selection by sparse partial least squares discriminant analysis between the 4 stroke subtypes (mixOmics R package version 4.3.2).[25] We performed 10-fold cross-validation repeated 50 times to tune the sparse partial lease squares discriminant analysis model and select the number of components and proteins for classifying stroke subtypes. The top 10 proteins from each component were selected as top proteins from the multigroup analysis.

Protein panels of stroke subtypes: The top-nominated proteins from the pairwise and multigroup comparisons were visualized using principal component analysis (factoextra R package) and heatmap (heatmap.2() function in gplots R package) for classifying the 4 stroke subtypes. From the top nominated hits, we performed regularized LASSO logistic regression to select top classifiers and develop protein panels for classifying AIS, ICH, TIA, and MIM (rlassologit function in the hdm R package).

Internal Validation

Internal validation was conducted using 2 approaches. (1) To evaluate the performance and generalizability of protein panels, we applied repeated nested 10-fold cross-validation (rCV) using LASSO logistic regression. Across 5 repetitions, models were trained on training folds and evaluated on test folds (total 50 outer folds). Predicted probabilities were pooled to compute the area under the receiver operating characteristic curves (AUC) and 95% CI using DeLong's method to quantify discrimination among stroke subtypes. Sensitivity and specificity were calculated at a 0.50 predicted probability, and positive predictive value and negative predictive value (NPVs) were calculated at different pretest probability cut-offs for classifying each stroke subtype. Performance metrics were validated via 2000 bootstrap resamples. A final model was fit on the full data set, and top protein classifiers were summarized by frequency across folds. (2) Additionally, we validated protein expression patterns using the BAK-270 targeted proteomics platform on the same samples, applying the same methodology of pairwise and multigroup comparisons. We determined the correlation of overlapping proteins between the SomaScan and the BAK-270 proteomics platform using Spearman's rank correlation.

External Validation

We externally validated protein expression patterns of the SomaScan hits using DIA-MS proteomics in an independent cohort. DIA-MS data were quantified using DIA-neural network,[26] $\log_2$ transformed, batch corrected using the Combat function (sva package in R), imputed for ≤30% missing values from random distribution using Perseus, and normalized by total area sum. The validated hits in pairwise and multigroup comparisons were identified using a P value <0.05, visualized using a principal component analysis plot and heatmap, and the discrimination ability was assessed using receiver operating characteristic curves.

RESULTS

Discovery Cohort Demographics

The plasma repository at Grady included plasma samples from 356 subjects with a suspected stroke diagnosis. Based on eligibility criteria, we selected 100 age- and sex-matched patients for our discovery cohort, divided into 4 subtypes: 40 with AIS, 20 with ICH, 20 with TIA, and 20 with MIM. AIS included 14 (35%) cardioembolic, 19 (47.5%) symptomatic large vessel, 5 (12.5%) small vessel disease, and 2 (5%) other subtypes. MIM included seizures (N=5), psychogenic/psychosis (N=5), migraine (N=5), hypertensive emergency (N=1), radiculopathy (N=1), weakness and dizziness (N=1), moyamoya disease without acute stroke (N=1), and others (N=1). The mean age was 58.5 (15.5) years, and 44 were male. We observed significant differences in mean age (younger in MIM), sex (more women in MIM), median NIHSS, modified Rankin Scale, and GCS scores across the subtypes (P<0.05), summarized in Table 1. The study flow diagram is shown in Figure 2.

Plasma Aptamer-Based Proteomics

We identified 7307 somamer targets, which included 6373 unique proteins with no missing values using the plasma 7K SomaScan assay, after removing control/noise somamers. Samples were processed in 2 batches (N=40 in Batch 1, N=60 in Batch 2) with equal representation of all 4 groups across both batches. Batch effects were removed using linear mixed models, and batch effects before and after correction were verified using variance partitioning analysis, ensuring that effects of other variables were not affected (Figure S1).

Pairwise Comparisons Between Stroke Subtypes

We conducted 6 pairwise comparisons: AIS versus ICH, AIS versus TIA, AIS versus MIM, ICH versus TIA, ICH versus MIM, and TIA versus MIM (Table 1). Using the 3 protein selection criteria, we identified 31 DEPs for AIS versus ICH, 50 for AIS versus TIA, 80 for AIS versus MIM, 62 for ICH versus TIA, 105 for ICH versus MIM, and 194 for TIA versus MIM (Tables S1 through S6). Next, we selected the top proteins identified by at least 2 of the 3 protein selection approaches and

Table 1. Baseline Characteristics of Patients Included in the Discovery Cohort of the Study

Variable	Total (N=100)	AIS (N=40)	ICH (N=20)	TIA (N=20)	MIM (N=20)	*P* value (all groups)	*P* value (AIS vs ICH)	*P* value (AIS vs TIA)	*P* value (AIS vs MIM)	*P* value (TIA vs MIM)	*P* value (ICH vs TIA)	*P* value (ICH vs MIM)
Age, y, mean±SD	58.49 ±15.5	62.95 ±16.09	56.75 ±11.04	62.25 ±15.51	47.55 ±12.98	0.001*	0.13	0.87	0.001*	0.002*	0.20	0.02*
Male sex, n (%)	44 (44)	19 (47.5)	10 (50)	12 (60)	3 (15)	0.02*	0.86	0.36	0.01*	0.003*	0.53	0.02*
Diabetes, n (%)	40 (40)	16 (40)	6 (30)	10 (50)	8 (40)	0.64	0.45	0.46	1.00	0.53	0.20	0.51
Hypertension, n (%)	81 (81%)	34 (85)	17 (85)	16 (80)	14 (70)	0.53	1.00	0.62	0.17	0.47	0.68	0.26
Coronary artery disease, n (%)	12 (12)	7 (17.5)	1 (5)	3 (15)	1 (5)	0.37	0.18	0.81	0.18	0.29	0.29	1.00
Atrial fibrillation, n (%)	12 (12)	7 (17.5)	1 (5)	4 (20)	0	0.11	0.18	0.81	N.E.	N.E.	0.15	N.E.
Smoking, n (%)	21 (21)	6 (15)	6 (30)	4 (20)	5 (25)	0.56	0.17	0.62	0.35	0.71	0.47	0.72
Prior stroke, n (%)	37 (37)	17 (42.5)	4 (20)	11 (55)	5 (25)	0.07	0.09	0.36	0.19	0.05	0.02*	0.71
Baseline National Institutes of Health Stroke Scale score, median (IQR)	7 (0–15.5)	11.5 (6–20.5)	8 (0–18.5)	1 (0–4)	1 (0–7.5)	<0.001*	0.21	<0.001*	<0.001*	0.71	0.02*	0.04*
Glasgow Coma Scale score, median (IQR)	14 (10.5–15)	13 (10–15)	13 (8–15)	15 (15–15)	14.5 (13–15)	0.002*	0.42	0.001*	0.20	0.05	0.001*	0.09
Discharge modified Rankin Scale score, median (IQR)	2 (1–4)	3 (2–4)	4 (2–5.5)	1.5 (0–2)	0 (0–2)	<0.001*	0.1	<0.001*	<0.001*	0.13	<0.001*	<0.001*

AIS indicates acute ischemic stroke; ICH, intracerebral hemorrhage; IQR, interquartile range; MIM, stroke mimics; N.E., not estimable; and TIA, transient ischemic attack.

*$P<0.05$.

adjusted them for age, sex, NIHSS score, and GCS score using regularized LASSO logistic regression (variance inflation factor <2 for collinearity). This resulted in 5 proteins for AIS versus ICH, 4 proteins for AIS versus TIA, 7 proteins for AIS versus MIM, 4 proteins for ICH versus TIA, 11 proteins for ICH versus MIM, and 8 proteins for TIA versus MIM (adjusted $P<0.05$) comparisons. We provide the adjusted P values for these top proteins in Table S7. See Figures S2 through S7 for DEPs identified in each pairwise comparison.

Functional Enrichment Analysis of DEPs With ±1.2-Fold Change and $P<0.05$

Pathways including the formation of fibrin clot, coagulation system, and integrin signaling were increased in stroke (AIS and ICH) compared with TIA. Extracellular matrix degradation, integrin cell surface interactions, posttranslational protein phosphorylation, and regulation of insulin-like growth factor were increased in stroke (AIS and ICH) compared with TIA and MIM. Wound healing signaling was increased in AIS compared with ICH, TIA, and MIM. Extracellular matrix organization and atherosclerosis signaling were increased in AIS, ICH, and TIA compared with MIM. The heatmap of enriched pathways is shown in Figure 3.

Multigroup Comparisons Identified a 90-Protein Panel for Classifying Stroke Subtypes

Using the sparse partial lease squares discriminant analysis approach, we identified 3 components that classified MIM from TIA (Component 1, 30 proteins), AIS from ICH and MIM (Component 2, 30 proteins), and ICH from MIM and TIA (Component 3, 30 proteins) (Table S8 and Figure S8). Because head CT is readily available and represents the gold standard to diagnose or rule out ICH in acute stroke settings, we also performed a secondary analysis using sparse partial lease squares discriminant analysis after excluding the group with ICH from our analysis. We identified proteins differentiating between AIS, TIA, and MIM: Component 1 (15 proteins) for MIM from TIA and AIS, and Component 2 (20 proteins) for AIS from TIA and MIM (Figure S9).

Nominated Proteins Highlight Top Classifiers and Protein Panels for Classifying Stroke Subtypes

We identified 39 top proteins from pairwise comparisons after adjusting for clinical covariates. After accounting for 3 duplicated proteins across comparisons, 36 unique

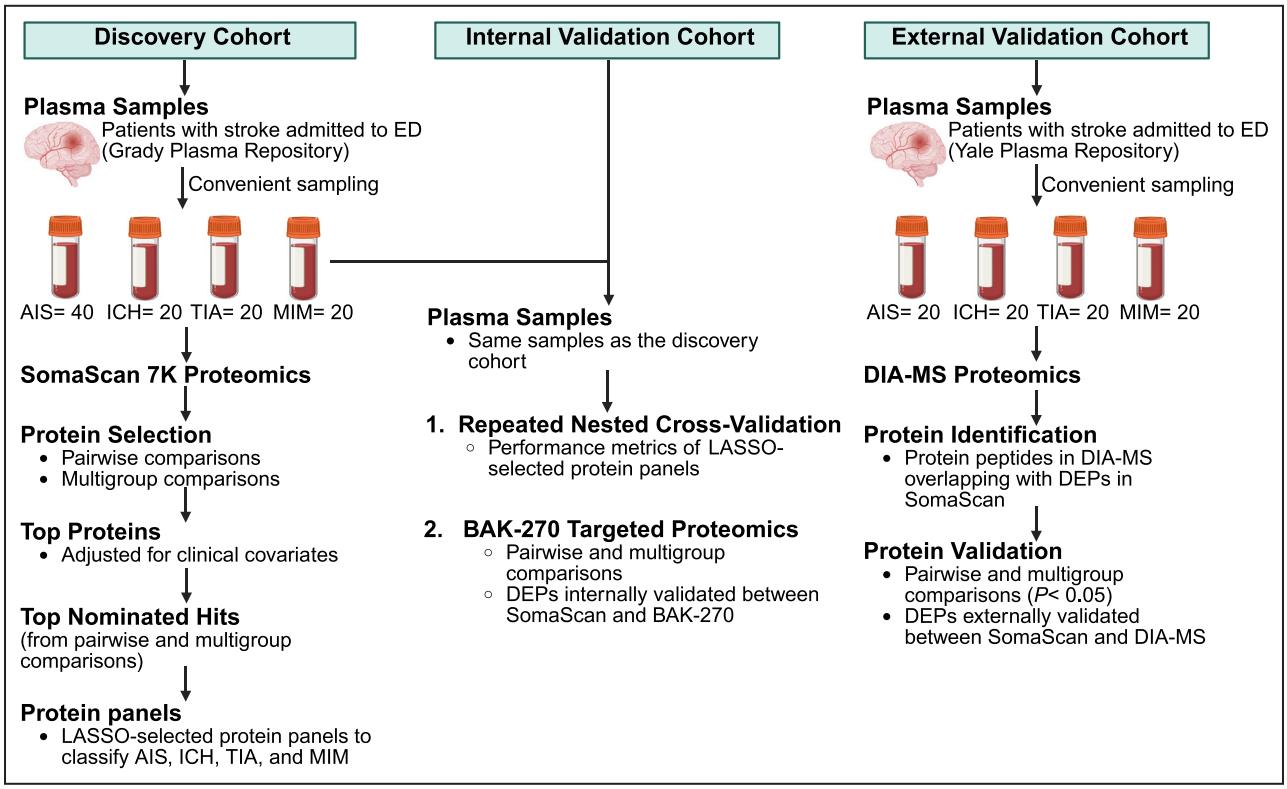

Figure 2. Plasma proteomics study flow diagram.
AIS indicates acute ischemic stroke; BAK-270, biomarker assessment kit-270; DEP, differentially enriched protein; DIA-MS, data independent acquisition-based mass spectrometry; ED, emergency department; ICH, intracerebral hemorrhage; LASSO, least absolute shrinkage and selection operator; MIM, stroke mimics; and TIA, transient ischemic attack.

proteins remained. Additionally, we selected 30 top proteins from multigroup comparisons (top 10 proteins from each component). Among the 66 top proteins identified across both approaches, 5 (EGLN3 [Egl-9 family hypoxia-inducible factor 3], UBE2B [ubiquitin conjugating enzyme E2B], PSG3 [pregnancy-specific beta-1-glycoprotein 3], PLG [plasminogen], and CDKN1B [cyclin-dependent kinase inhibitor 1B]) overlapped. After consolidating results and removing duplicates, we nominated 61 unique protein targets that accurately classified the stroke subtypes (Figure 4). The discriminative performance (AUCs) of these proteins is summarized in Table S9. Notably, the majority of proteins (27/61) distinguish MIM from AIS, ICH, and TIA (AUCs ≥0.70).

After using LASSO logistic regression for feature selection on these 61 nominated hits, we identified 7 protein classifiers for AIS (TIMP3 [tissue inhibitor of metalloproteinases-3], CR2 [complement receptor 2], APLP1 [amyloid precursor like protein 1], AXIN2 [axis inhibition protein 2], IL1RAP [interleukin-1 receptor accessory protein], FGFBP3 [fibroblast growth factor binding protein 3], LILRA6 [leukocyte immunoglobulin-like receptor A6]), 6 for ICH (PSG3, IZUMO4 [izumo sperm-egg fusion protein 4], UBE2B, S100A9, CDKN1B, SQSTM1 [sequestosome-1]), 8 for TIA (HSPA9 [heat shock protein family A member 9], CTRB2 [chymotrypsinogen B2], ADGRB2 [adhesion G protein-coupled receptor B2], UNC5C [netrin receptor UNC5C], NDUFV2 [NADH dehydrogenase [ubiquinone] flavoprotein 2, mitochondrial], TENM3 [teneurin transmembrane protein 3], PLG, CRIP1 [cysteine-rich intestinal protein 1]), and 7 for MIM (CXCL12 [stromal cell-derived factor 1], VTN [vitronectin], INIP [INTS3- and NABP-interacting protein], FAM20C [family with sequence similarity 20, member C], EGLN3, CTSB [cathepsin B], LILRA6).

Repeated Nested Cross-Validation and Bootstrap Resampling of Protein Panels

We performed 10-fold nested cross-validation repeated 5 times (ie, 50 sampling reiterations) to evaluate the performance metrics and generalizability of the protein panels identified using LASSO logistic regression. In each iteration, we trained on 90% of data and tested on the remaining 10%, with each subject serving as a validation case exactly once per repeat. After repeated nested cross-validation across 50 sampling reiterations, the 4 protein panels distinguished AIS from ICH/TIA/MIM with an average AUC of 0.82 (95% CI, 0.78–0.86) and NPV of 75.2% (95% CI, 72%–78.5%) (Figure 5A), ICH from AIS/TIA/MIM with an average AUC of 0.70 (95% CI, 0.64–0.76) and NPV of 84.1% (95% CI, 82.3%–86%)

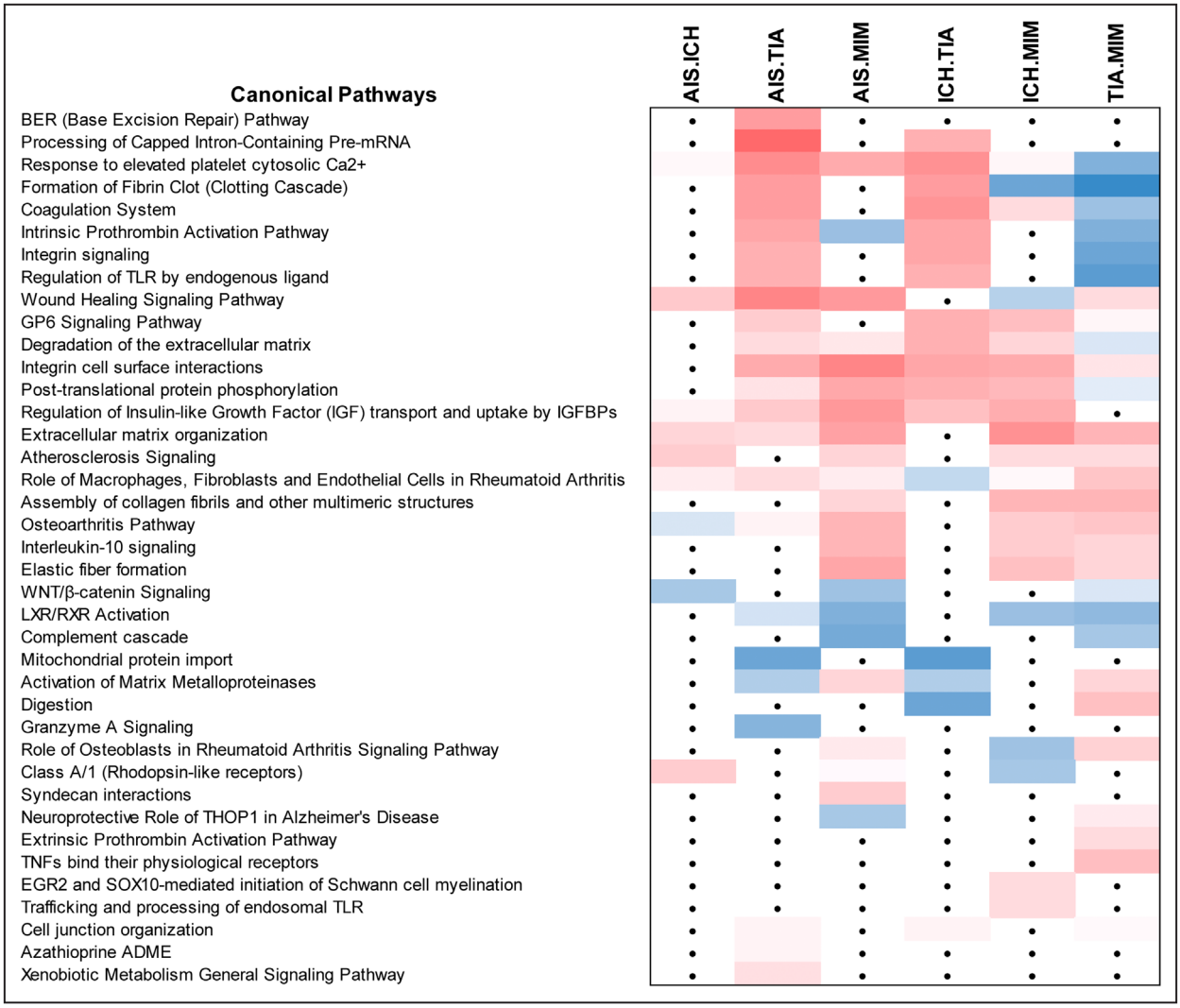

Figure 3. Functional enrichment analysis of canonical pathways differentially enriched in each pairwise comparison. The heatmap represents upregulated pathways in red and downregulated pathways in blue. • indicates insignificance threshold (false discovery rate adjusted $P>0.1$). ADME indicates absorption, distribution, metabolism, and excretion; AIS, acute ischemic stroke; EGR2, early growth response protein 2; GP6, glycoprotein VI; ICH, intracerebral hemorrhage; IGFBP, insulin-like growth factor-binding protein 2; LXR, liver X receptor; MIM, stroke mimics; RXR, retinoid X receptor; THOP1, thimet oligopeptidase; TIA, transient ischemic attack; TLR, toll-like receptor; and TNF, tumor necrosis factor.

(Figure 5B), TIA from AIS/ICH/MIM with an average AUC of 0.78 (95% CI, 0.73–0.84) and NPV of 87.9% (95% CI, 85.8%–90.1%) (Figure 5C), and MIM from AIS/ICH/TIA with an average AUC of 0.81 (95% CI, 0.77–0.86) and NPV of 87.6% (95% CI, 85.4%–89.6%) (Figure 5D). The performance metrics are given in Table 2, and the selection frequency of top protein classifiers across 50 sampling reiterations is given in Figure S10.

Internal Validation Using Targeted Proteomics (BAK-270) Validated 11 Proteins

We next used the same 99 of 100 samples for targeted proteomics using the BAK-270 platform for cross-platform validation. One TIA sample was missing and thus excluded from the validation analysis (N=99). We measured 249 proteins across the 4 stroke subtypes, performed random forest-based imputation for proteins with ≤30% missing values,[27] and removed 33 proteins with >30% missing values, resulting in 216 proteins with a complete data set. Of these, 185 proteins overlapped between the SomaScan and BAK-270 platforms, with 75 showing moderate to high correlation ($r_s \geq 0.5$) with the SomaScan data set (Table S10).

We validated 11 proteins using targeted MS, with VTN, IGFBP2 (insulin-like growth factor-binding protein 2), CNDP1 (carnosine dipeptidase 1), B2M (beta-2-microglobulin), SERPIND1 (serpin family D member 1), PLG, S100A9, and F2 in pairwise comparisons

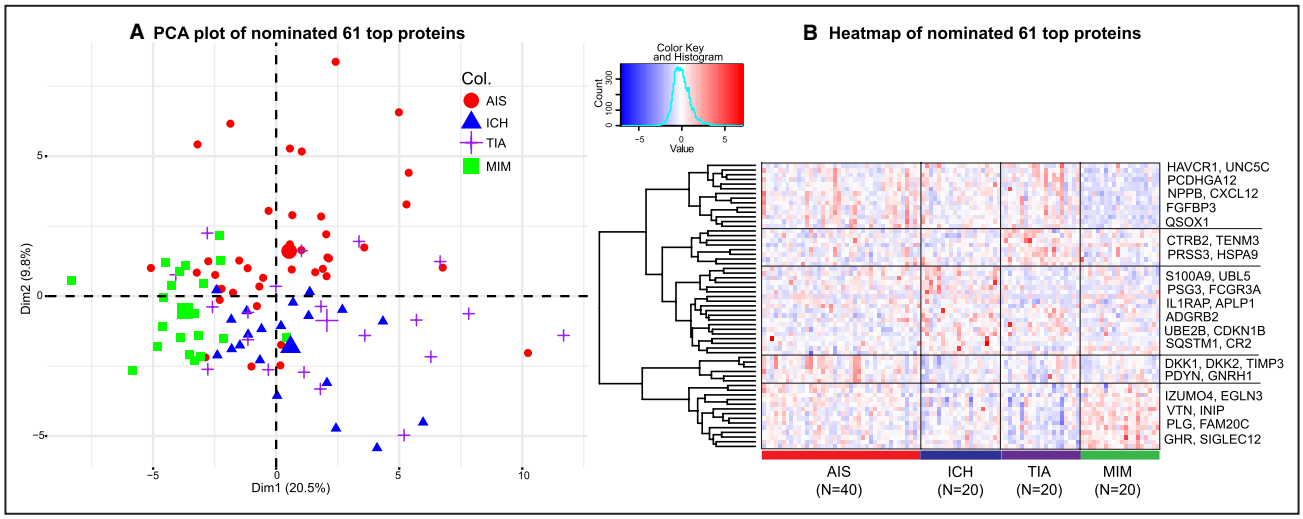

Figure 4. Biomarker panel of nominated 61 top proteins to classify stroke subtypes.
A, PCA plot and (**B**) Heatmap of the nominated 61 top proteins to differentiate the stroke subtypes. The heatmap depicts the names of 35 of the 61 nominating hits for visualization. ADGRB2 indicates adhesion G protein-coupled receptor B2; AIS, acute ischemic stroke; APLP1, amyloid precursor like protein 1; CDKN1B, cyclin-dependent kinase inhibitor 1B; CR2, complement receptor 2; CTRB2, chymotrypsinogen B2; CXCL12, stromal cell-derived factor 1; DKK1,2, Dickkopf 1,2; EGLN3, Egl-9 family hypoxia-inducible factor 3; FAM20C, family with sequence similarity 20, member C; FCGR3A, Fc gamma receptor IIIa; FGFBP3, fibroblast growth factor binding protein 3; GHR, growth hormone receptor; GNRH1, gonadotropin-releasing hormone 1; HAVCR1, hepatitis A virus cellular receptor 1; HSPA9, heat shock protein family A member 9; ICH, intracerebral hemorrhage; IL1RAP, interleukin-1 receptor accessory protein; INIP, INTS3- and NABP-interacting protein; IZUMO4, izumo sperm-egg fusion protein 4; MIM, stroke mimics; NPPB, natriuretic peptide B; PCA, principal component analysis; PCDHGA12, protocadherin gamma-A12; PDYN, prodynorphin; PLG, plasminogen; PRSS3, protease serine 3; PSG3, pregnancy-specific beta-1-glycoprotein 3; QSOX1, quiescin sulfhydryl oxidase 1; SIGLEC12, sialic acid-binding Ig-like lectin 12; SQSTM1, sequestosome-1; TENM3, teneurin transmembrane protein 3; TIA, transient ischemic attack; TIMP3, tissue inhibitor of metalloproteinases-3; UBE2B, ubiquitin conjugating enzyme E2B; UBL5, ubiquitin-like protein 5; and VTN, vitronectin.

(Table S11 and Figure S11) and VTN, PLG, apo A4 (apolipoprotein A4), CPN2 (carboxypeptidase N subunit 2), PGLYRP2 (peptidoglycan recognition protein 2), and S100A9 in multigroup comparisons (Figure S12). Of the 11 validated hits, 4 were identified as top nominating hits in SomaScan (VTN, PLG, SERPIND1, and S100A9). Higher VTN and PLG levels differentiated MIM from AIS, ICH, and TIA with an AUC >0.70 in both discovery and internal validation (Figure S13).

External Validation Using DIA-MS Validated 32 Proteins That Distinguish Stroke Diagnostic Groups

The Yale stroke plasma biorepository included samples from 234 patients with a suspected stroke diagnosis. Based on eligibility criteria, we selected 80 age- and sex-matched patients for our external validation cohort, classified into the following groups: 20 with AIS, 20 with ICH, 20 with TIA, and 20 with MIM. AIS included 12 (60%) cardioembolic, 4 (20%) symptomatic large vessel, 1 (5%) small vessel disease, and 3 (15%) other subtypes. MIM included weakness and dizziness (N=6), migraine (N=3), seizure (N=3), hypertensive urgency (N=1), septic shock (N=1), ischemic colitis (N=1), encephalopathy (N=1), and others (N=4). The mean age was 67.9 (17.4) years, and 45 were male. We observed significant differences in median NIHSS, modified Rankin Scale, and GCS scores across the subtypes (P<0.05) summarized in Table 3. Compared with the discovery cohort, the external validation cohort showed no significant differences in age or sex but included subjects with milder AIS and a higher proportion of severe ICH.

For the external validation cohort, we analyzed peptide-level data to ensure consistency between the epitope-based SomaScan measurements in the discovery cohort and the peptide-based targeted proteomics measurements in the internal validation cohort. After removing peptides with >30% missing values, we included 7153 peptide measurements from 638 proteins in the final analysis. The batching was done at the level of depletion using a standardized kit and the same volume of plasma, although all 80-sample processing after depletion, including digestion and DIA-MS, was performed together. Samples were processed in 3 batches (N=33 in Batch 1, N=26 in Batch 2, and N=21 in Batch 3), and batch effects were corrected using the *Combat* function in R (Figure S14). Across SomaScan and DIA-MS platforms, 502 proteins were commonly quantified. Among these, 68 proteins encompassing 1343 peptides in DIA-MS were DEPs in SomaScan (discovery cohort) based on any pairwise or multigroup comparison and were selected for external validation.

In the independent cohort, we validated 45 differentiating protein peptides across pairwise and multigroup comparisons, including 5 duplicates, resulting in 40 unique peptides (Table 4 and Table S12). These 40 peptides corresponded to 32 proteins that were DEPs in SomaScan and successfully validated by DIA-MS in an independent cohort, including top nominating hits such as VTN, PLG, LCP1 (lymphocyte cytosolic

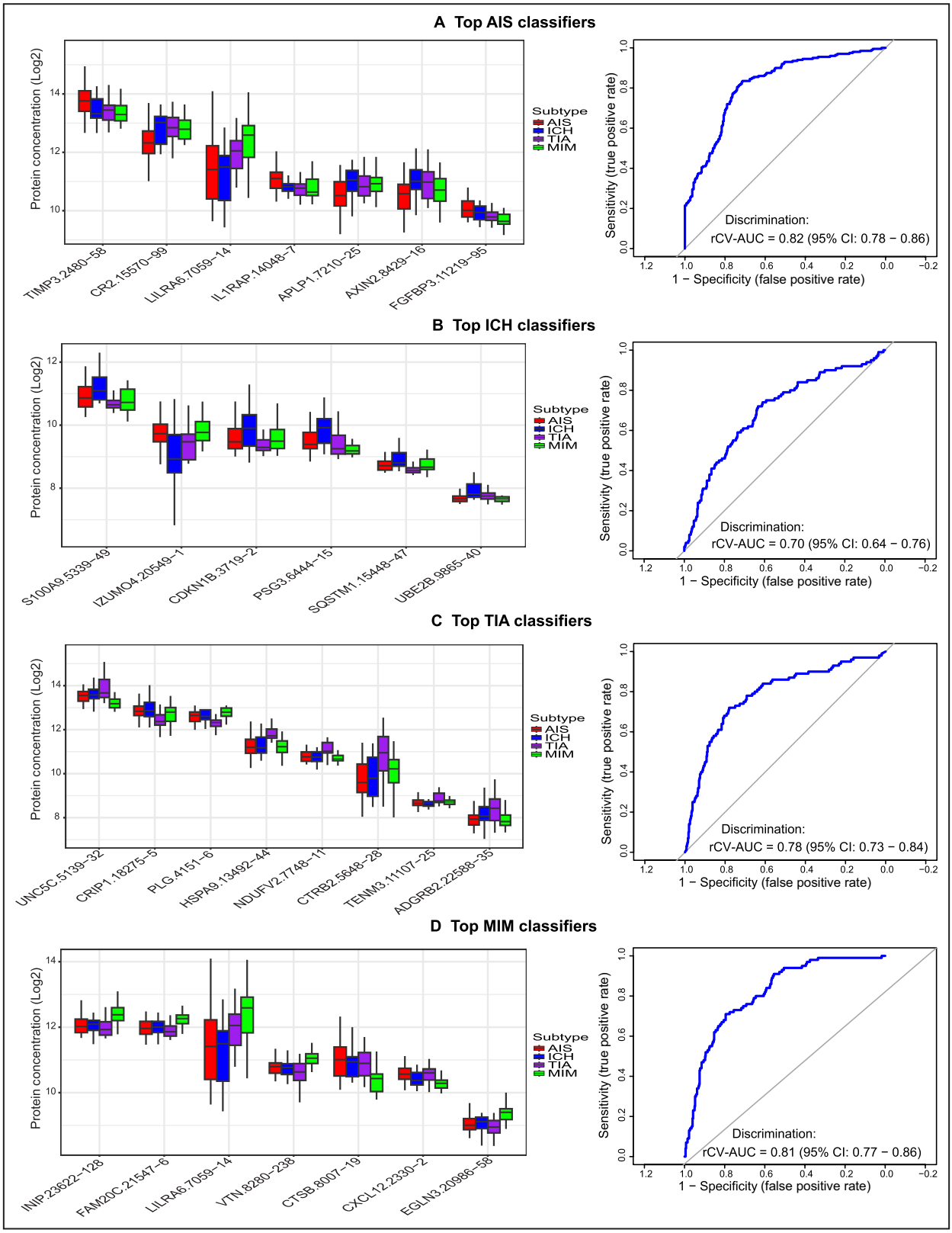

Figure 5. Top classifiers of stroke subtypes.
Box plots (left) and receiver operating characteristic curves after repeated nested 10-fold cross-validation (right) of (**A**) top AIS classifiers, (**B**) top ICH classifiers, (**C**) top TIA classifiers, and (**D**) top MIM classifiers. ADGRB2 indicates adhesion G protein-coupled receptor B2; AIS, acute ischemic stroke; APLP1, amyloid precursor like protein 1; AUC, area under the curve; AXIN2, axis inhibition protein 2; CDKN1B, cyclin-dependent kinase inhibitor 1B; CR2, complement receptor 2; CRIP1, cysteine-rich intestinal protein 1; CTRB2, chymotrypsinogen B2; FGFBP3, fibroblast growth factor binding protein 3; HSPA9, heat shock protein family A member 9; ICH, intracerebral hemorrhage; IL1RAP, interleukin-1 receptor accessory protein; IZUMO4, izumo sperm-egg fusion protein 4; LILRA6, leukocyte immunoglobulin-like receptor A6; MIM, stroke mimics; PLG, plasminogen; PSG3 pregnancy-specific beta-1-glycoprotein 3; rCV, repeated cross-validation; SQSTM1, sequestosome-1; TENM3, teneurin transmembrane protein 3; TIA, transient ischemic attack; TIMP3, tissue inhibitor of metalloproteinases-3; and UBE2B, ubiquitin conjugating enzyme E2B.

protein 1), S100A9, QSOX1 (quiescin sulfhydryl oxidase 1), SERPIND1, and FAM20C (Figures S15 through S19). Figure 6 shows the principal component analysis plot and heatmap of these 40 protein peptides that classify the 4 stroke subtypes after adjusting for NIHSS score and atrial fibrillation.

DISCUSSION

We employed cross-platform aptamer-based proteomics followed by targeted and untargeted proteomics and supervised machine learning algorithms and nominated 61 plasma proteins that were differentially

Table 2. Performance Evaluation of Protein Panels Using Repeated Nested Cross-Validation (Outer Loop) and Bootstrapping

Top classifiers	Protein panels	AUC (95% CI)	rCV-AUC (95% CI)	Pretest probability	rCV-sensitivity (95% CI)	rCV-specificity (95% CI)	rCV-PPV (95% CI)	rCV-NPV (95% CI)
AIS vs ICH+TIA+MIM	Tissue inhibitor of metalloproteinases-3, complement receptor 2, amyloid precursor like protein 1, axis inhibition protein 2, interleukin-1 receptor accessory protein, fibroblast growth factor binding protein 3, leukocyte immunoglobulin-like receptor A6	0.93 (0.87–0.98)	0.82 (0.78–0.86)	40%	59.5% (52.8–66.2%)	81.7% (77.2–86%)	68.4% (63–74.2%)	75.2% (72–78.5%)
ICH vs AIS+TIA+MIM	Pregnancy-specific beta-1-glycoprotein 3, izumo sperm-egg fusion protein 4, ubiquitin conjugating enzyme E2B, S100A9, cyclin-dependent kinase inhibitor 1B, sequestosome-1	0.91 (0.82–0.99)	0.70 (0.64–0.76)	20%	32% (23.1–41.6%)	90.3% (87.3–93.1%)	45.1% (34.8–55.5%)	84.1% (82.3–86%)
TIA vs ICH+TIA+MIM	Heat shock protein family A member 9, chymotrypsinogen B2, adhesion G protein-coupled receptor B2, netrin receptor UNC5C, NADH dehydrogenase [ubiquinone] flavoprotein 2, mitochondrial, teneurin transmembrane protein 3, plasminogen, cysteine-rich intestinal protein 1	0.97 (0.94–1.00)	0.78 (0.73–0.84)	20%	51% (41.4–61.1%)	88.8% (85.6–91.7%)	53.1% (44.9–61.3%)	87.9% (85.8–90.1%)
MIM vs AIS+ICH+TIA	Stromal cell-derived factor 1, vitronectin, INTS3- and NABP-interacting protein, family with sequence similarity 20 member C, Egl-9 family hypoxia-inducible factor 3, cathepsin B, leukocyte immunoglobulin-like receptor A6	0.97 (0.94–1.00)	0.81 (0.77–0.86)	20%	49% (39.1–58.3%)	89.8% (86.5–92.7%)	54.4% (45.4–63.9%)	87.6% (85.4–89.6%)

Ten-fold nested cross-validation (outer loop) repeated 5 times and 2000 bootstrapping CIs. AIS indicates acute ischemic stroke; AUC, area under the curve; ICH, intracerebral hemorrhage; MIM, stroke mimics; NPV, negative predictive value; PPV, positive predictive value; rCV, repeated cross-validation; and TIA, transient ischemic attack.

Table 3. Baseline Characteristics of Patients Included in the External Validation Cohort of the Study

Variable	Total (N=80)	AIS (N=20)	ICH (N=20)	TIA (N=20)	MIM (N=20)	*P* value (all groups)	*P* value (AIS vs ICH)	*P* value (AIS vs TIA)	*P* value (AIS vs MIM)	*P* value (TIA vs MIM)	*P* value (ICH vs TIA)	*P* value (ICH vs MIM)
Age, y, mean±SD	67.85±17.41	68.3±19.27	69.65±15.47	69.05±14.79	64.4±20.32	0.89	0.91	0.89	0.55	0.41	0.90	0.46
Male sex, n (%)	45 (56.25)	10 (50)	11 (55)	15 (75)	9 (45)	0.24	0.75	0.10	0.75	0.05	0.19	0.53
Sample collection time from stroke onset in h, median (IQR)	10.29 (3.07–18.15)	12.75 (2.21–18.5)	7.13 (4.50–12.34)	17.94 (6.75–18.94)	3.07 (1.23–18.69)	0.11	0.29	0.17	0.58	0.16	0.006*	0.52
Diabetes, n (%)	19 (23.75)	6 (30)	3 (15)	5 (25)	5 (25)	0.73	0.26	0.72	0.72	1.00	0.43	0.43
Hypertension, n (%)	61 (76.25)	17 (85)	15 (75)	15 (75)	14 (70)	0.73	0.43	0.43	0.26	0.72	1.00	0.72
Coronary artery disease, n (%)	16 (20)	7 (35)	2 (10)	3 (15)	4 (20)	0.22	0.06	0.14	0.29	0.68	0.63	0.38
Atrial fibrillation, n (%)	20 (25)	9 (45)	2 (10)	5 (25)	4 (20)	0.07	0.01*	0.19	0.09	0.71	0.21	0.38
Smoking, n (%)	31 (38.75)	7 (35)	7 (35)	8 (40)	9 (45)	0.90	1.00	0.74	0.52	0.75	0.74	0.52
Prior stroke, n (%)	17 (21.25)	4 (20)	4 (20)	3 (15)	6 (30)	0.70	1.00	0.68	0.47	0.26	0.68	0.47
Baseline National Institutes of Health Stroke Scale score, median (IQR)	2 (0–5)	3 (1–5)	13 (3–19.5)	0 (0–1)	1 (0–3.5)	<0.001*	0.01*	<0.001*	0.07	0.04*	<0.001*	<0.001*
Glasgow Coma Scale score, median (IQR)	15 (15–15)	15 (15–15)	15 (12–15)	15 (15–15)	15 (15–15)	0.001*	0.02*	0.15	0.64	0.08	0.001*	0.03*
Discharge modified Rankin Scale score, median (IQR)	1 (0–4)	1.5 (0–4)	4 (2.5–6)	0 (0–0.5)	1 (0.5–1.5)	<0.001*	0.003*	0.005*	0.51	0.004*	<0.001*	<0.001*

AIS indicates acute ischemic stroke; ICH, intracerebral hemorrhage; IQR, interquartile range; MIM, stroke mimics; and TIA, transient ischemic attack.
*$P<0.05$.

abundant across 4 major stroke categories (AIS, ICH, TIA, and MIM). Notably, we developed a panel of 7 proteins that serve as top classifiers for AIS, 6 for ICH, 8 for TIA, and 7 for MIM. Repeated nested cross-validation of these protein panels across 50 sampling reiterations demonstrated good discrimination ability (AUCs ≥0.70). To validate hits emerging from aptamer-based proteomics, we performed internal validation using targeted MS on the same cohort and confirmed the utility of 11 key proteins, including VTN and PLG, in differentiating MIM from AIS, ICH, and TIA with an AUC >0.70 (Figure S13). Additionally, external validation using DIA-MS on an independent cohort of 80 samples (20/group) validated 32 proteins, including top-nominating hits such as VTN, PLG, S100A9, FAM20C, LCP1, QSOX1, and SERPIND1. These validated proteins represent a robust set of candidate biomarkers for stroke diagnostic subtyping, particularly because they emerge despite inherent differences in proteomics platforms, differences across study sites, and clinical characteristics in derivation and external validation cohorts. The proteins are intended to serve as candidate biomarkers for future studies in larger cohorts, as well as using targeted and lower-plex assay platforms, which may be more readily applicable for use in clinical scenarios. These protein panels have important implications for improving clinical decision-making and patient outcomes in emergency settings. Beyond diagnostic classifications, our results using expanded lists of classifiers also provide biological insights into how molecular mechanisms of ischemic and hemorrhagic stroke in the brain are reflected in the blood via peripheral coagulation and inflammatory cascades. Our motivation to analyze the data using prespecified pairwise and multigroup approaches is to provide biomarker candidates applicable to different clinical contexts. For example, we envision that some plasma biomarkers may be used to distinguish TIA from MIM after ICH has been excluded by CT imaging and after AIS has been excluded by brain magnetic resonance imaging. In such cases, a biomarker of MIM or TIA can be very useful in triaging patients who need expedited TIA evaluation and follow-up versus those with MIM diagnoses that need nonstroke-related management.

Our study nominated 4 distinct protein panels for stroke subtype classification. A panel of 7 proteins (TIMP3, CR2, APLP1, AXIN2, IL1RAP, FGFBP3, and LILRA6) differentiated AIS from ICH, TIA, and MIM with rCV-AUC of 0.82 (95% CI, 0.78–0.86). Our panel was relatively effective in identifying patients that do not have AIS (high specificity 82% [95% CI, 77%–86%] and NPV 75% [95% CI, 72%–79%]), making it reliable for confirming when a patient does not have AIS and is less prone to misclassifying non-AIS as AIS, which could potentially lead to harmful treatment (eg, tissue plasminogen activator administered to a patient with ICH).

Table 4. Differentially Enriched Proteins Externally Validated Between SomaScan and DIA-MS in Pairwise Comparisons

S. No	Protein	SomaScan			DIA-MS			
		FC	*P* value	AUC (95% CI)	FC	*P* value	AUC (95% CI)	No. of validated peptides per protein*
AIS vs ICH								
1	ITIH4	1.16	<0.001	0.77 (0.63–0.89)	1.43	0.001	0.79 (0.64–0.94)	16
2	CKM	0.63	0.049	0.64 (0.49–0.79)	0.55	0.026	0.71 (0.54–0.87)	1
AIS vs TIA								
3	C3	0.53	0.016	0.70 (0.56–0.84)	0.18	<0.001	0.87 (0.75–0.99)	32
4	FN1.3435–53	1.60	0.001	0.74 (0.60–0.87)	3.07	<0.001	0.82 (0.69–0.95)	12
	FN1.3434–34	1.67	0.001	0.73 (0.59–0.86)				
	FN1.20064–24	1.66	0.002	0.73 (0.60–0.86)				
	FN1.4131–72	1.81	0.006	0.70 (0.57–0.84)				
5	PLG	1.26	<0.001	0.79 (0.68–0.91)	2.89	0.001	0.81 (0.67–0.95)	13
6	Vitronectin	1.55	0.017	0.71 (0.56–0.86)	1.56	0.005	0.74 (0.56–0.91)	5
AIS vs MIM								
7	C3	0.57	0.039	0.67 (0.53–0.81)	0.20	<0.001	0.93 (0.84–1.00)	34
8	Lymphocyte cytosolic protein 1	0.84	0.001	0.76 (0.61–0.90)	0.75	0.003	0.77 (0.62–0.92)	8
9	Protein tyrosine phosphatase receptor type J	0.80	0.001	0.82 (0.69–0.94)	0.75	0.047	0.68 (0.51–0.85)	1
ICH vs TIA								
10	Amylase alpha 2A	0.71	0.002	0.80 (0.65–0.94)	0.57	0.002	0.78 (0.63–0.93)	1
11	CKM	1.73	0.042	0.67 (0.50–0.84)	2.72	<0.001	0.83 (0.69–0.96)	1
12	C-type lectin domain family 3 member B	1.24	0.005	0.79 (0.64–0.93)	3.21	<0.001	0.81 (0.65–0.96)	4
13	F13 B chain	1.66	0.011	0.82 (0.68–0.96)	4.32	<0.001	0.78 (0.60–0.95)	3
14	F13 A1 chain	1.66	0.011	0.82 (0.68–0.96)	1.43	0.005	0.77 (0.62–0.93)	3
15	Histidine-rich glycoprotein	1.46	0.001	0.81 (0.67–0.94)	2.58	<0.001	0.88 (0.78–0.99)	7
16	FN1.3435–53	1.94	0.002	0.82 (0.68–0.95)	2.86	0.046	0.69 (0.52–0.85)	1
	FN1.3434–34	2.00	0.002	0.81 (0.67–0.95)				
	FN1.20064–24	2.04	0.001	0.80 (0.66–0.94)				
	FN1.4131–72	2.17	0.003	0.77 (0.62–0.91)				
17	PLG	1.35	0.001	0.84 (0.71–0.97)	6.44	<0.001	0.91 (0.81–1.00)	20
18	Regenerating islet-derived 1 alpha	0.64	0.035	0.65 (0.48–0.82)	0.64	0.037	0.69 (0.52–0.85)	1
19	S100A9	1.41	<0.001	0.86 (0.74–0.97)	3.29	0.001	0.78 (0.63–0.93)	3
20	Transgelin-2	0.86	0.054	0.72 (0.55–0.89)	0.40	0.023	0.69 (0.52–0.85)	1
ICH vs MIM								
21	CD93	1.32	0.001	0.81 (0.67–0.95)	2.78	0.008	0.78 (0.63–0.93)	1
22	Collagen alpha-1(VI) chain	1.77	0.025	0.67 (0.50–0.85)	1.90	0.005	0.79 (0.64–0.94)	1
23	Carboxypeptidase N subunit 2	0.84	0.005	0.75 (0.59–0.90)	0.67	<0.001	0.70 (0.53–0.86)	6
24	Coagulation factor XI	0.84	0.001	0.80 (0.65–0.94)	0.75	0.012	0.75 (0.59–0.91)	4
25	Hyaluronic acid binding protein 2	0.76	0.001	0.79 (0.65–0.94)	0.76	<0.001	0.80 (0.65–0.94)	5
26	Interleukin-6 signal transducer	1.14	0.011	0.78 (0.62–0.93)	2.20	<0.001	0.82 (0.67–0.97)	1
27	ITIH4	0.88	0.003	0.77 (0.62–0.92)	0.37	<0.001	0.92 (0.83–1.00)	18
28	Laminin subunit alpha-2	1.24	0.011	0.73 (0.57–0.89)	1.49	0.004	0.77 (0.61–0.92)	1
29	Melanoma cell adhesion molecule	1.31	0.001	0.82 (0.68–0.96)	2.80	<0.001	0.90 (0.80–0.99)	4
30	PLG	0.85	0.004	0.76 (0.60–0.91)	0.66	<0.001	0.87 (0.74–0.99)	9
31	Quiescin sulfhydryl oxidase 1	1.21	0.001	0.82 (0.67–0.97)	2.64	<0.001	0.80 (0.65–0.95)	3
32	Serpin family D member 1	0.75	0.001	0.77 (0.62–0.92)	0.76	0.006	0.74 (0.58–0.89)	6

(*Continued*)

Table 4. Continued

S. No	Protein	SomaScan			DIA-MS			
		FC	*P* value	AUC (95% CI)	FC	*P* value	AUC (95% CI)	No. of validated peptides per protein*
TIA vs MIM								
33	F13B	0.50	0.005	0.78 (0.62–0.95)	0.44	0.010	0.72 (0.56–0.89)	1
34	Fibrinogen alpha chain	0.64	0.013	0.74 (0.57–0.90)	0.22	0.001	0.76 (0.60–0.91)	1
35	PLG	0.61	0.001	0.92 (0.83–1.00)	0.78	0.014	0.71 (0.54–0.87)	1
36	Thrombospondin-1	0.67	0.039	0.66 (0.48–0.83)	0.28	0.001	0.79 (0.65–0.93)	20

AIS indicates acute ischemic stroke; AUC, area under the curve; FC, fold change; C3, complement component 3; CKM, creatine kinase; DIA-MS, data independent acquisition-based mass spectrometry; F13, coagulation factor XIII; FN1, fibronectin 1; ICH, intracerebral hemorrhage; ITIH4, inter-alpha-trypsin inhibitor heavy chain H4; MIM, stroke mimics; PLG, plasminogen; and TIA, transient ischemic attack.

*Indicates the total number of significant peptides validated for a given protein in the same direction. The peptide with most significant *P* value was selected as the representative for the protein.

These proteins are involved in biological processes such as extracellular matrix remodeling, inflammation, and Wnt signaling, all of which are involved in ischemic brain injury.[28] For example, TIMP3's involvement in vascular inflammation and blood–brain barrier disruption supports its role as a marker for vascular injury in AIS, corroborating previous studies linking elevated TIMP3 levels in response to AIS.[29] A recent proteomics study identified C3 (complement component 3), ICAM-2 (intercellular adhesion molecule 2), PLGLA (plasminogen-like protein A), STXBP5 (syntaxin-binding protein 5), and IGHV3-6 (immunoglobulin heavy variable 3-6) as top classifiers in differentiating AIS from ICH and MIM with 75% sensitivity and 95% NPV; however, they did not include TIA as a diagnostic group.[30] Of these, only C3 and ICAM-2 were quantified in our study using SomaScan, with lower C3 levels in AIS differentiating it from TIA and MIM (Tables S2 and S3). Another recent proteomics study reported GFAP (glial fibrillary acidic protein), BCAN (brevican), SNAP25 (synaptosomal associated protein 25), and SPOCK1 (testican-1) as the most promising biomarkers for differentiating AIS from ICH with AUCs >0.75.[31] The study also found lower levels of UBE2B and APLP1 in AIS compared with

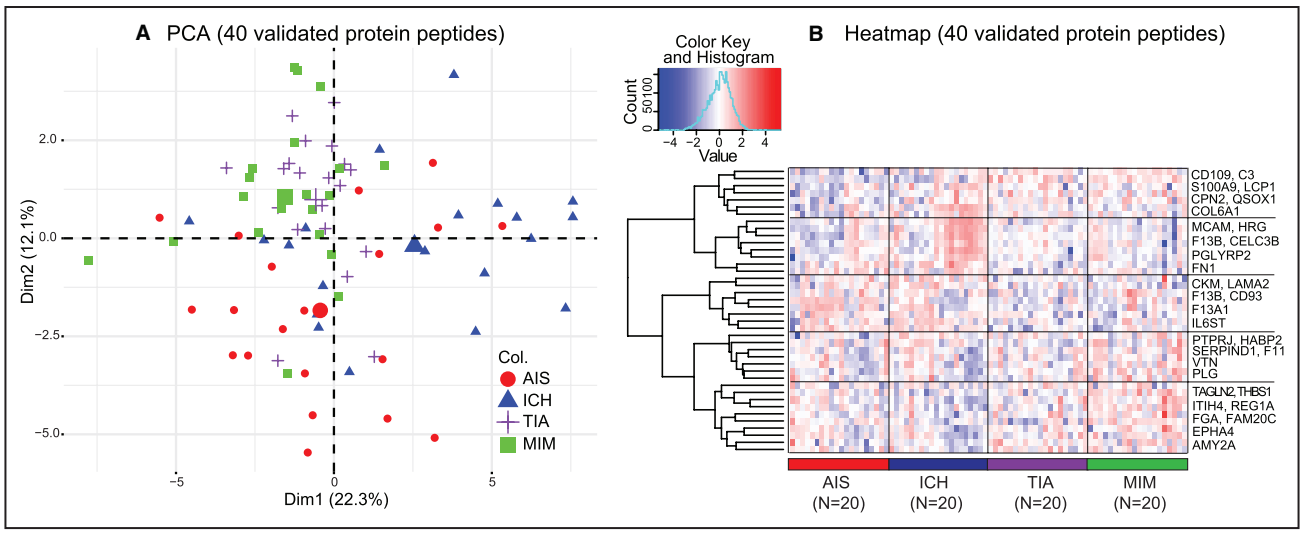

Figure 6. Biomarker panel of 40 externally validated protein peptides to classify stroke subtypes.
A, PCA plot and (**B**) Heatmap of the 40 externally validated protein peptides after adjusting for National Institutes of Health Stroke Scale score and atrial fibrillation to differentiate stroke subtypes. AIS indicates acute ischemic stroke; AMY2A, amylase alpha 2A; C3, complement component 3; CELC3B, tetranectin; CKM, creatine kinase; COL6A1, collagen alpha-1(VI) chain; CPN2, carboxypeptidase N subunit 2; EPHA4, ephrin type-A receptor 4; F11, coagulation factor XI; F13A1, coagulation factor XIII A1 chain; F13B, coagulation factor XIII B chain; FAM20C, family with sequence similarity 20, member C; FGA, fibrinogen alpha chain; FN1, fibronectin 1; HABP2, hyaluronic acid binding protein 2; HRG, hormone growth receptor; ICH, intracerebral hemorrhage; IL6ST, interleukin-6 signal transducer; ITIH4, inter-alpha-trypsin inhibitor heavy chain H4; LAMA2, laminin subunit alpha-2; LCP1, lymphocyte cytosolic protein 1; MCAM, melanoma cell adhesion molecule; MIM, stroke mimics; PCA, principal component analysis; PGLYRP2, peptidoglycan recognition protein 2; PLG, plasminogen; PTPRJ, protein tyrosine phosphatase receptor type J; QSOX1, quiescin sulfhydryl oxidase 1; REG1A, regenerating islet-derived 1 alpha; SERPIND1, serpin family D member 1; TAGLN2, transgelin-2; THBS1, thrombospondin-1; TIA, transient ischemic attack; and VTN, vitronectin.

ICH, consistent with our findings (Table S1). Previous studies using low-plex assays have identified protein biomarker panels with promising results for AIS diagnosis.[10,32–34] Notably, a biomarker panel including GFAP, RBP4 (retinol binding protein 4), and NT-proBNP (NT-terminal pro-B-type natriuretic peptide) had 100% specificity to distinguish AIS from ICH.[10] However, the diagnostic utility of these biomarkers was not tested against other clinically relevant comparators, including TIA and MIM.

For ICH, we identified a panel of 6 proteins (PSG3, IZUMO4, UBE2B, S100A9, SQSTM1, and CDKN1B) as top classifiers with rCV-AUC of 0.70 (95% CI, 0.64–0.76). S100A9 was also internally validated using targeted proteomics and externally validated using DIA-MS. This panel correctly identified non-ICH cases with high specificity (90%; 95% CI, 87%–93%) and NPV (84%; 95% CI, 82%–86%). S100A9, a known marker of inflammation, showed increased levels consistent with previous studies.[35] Although CDKN1B's role in ICH is unclear, it is involved in cell cycle regulation and nerve cell survival in subarachnoid hemorrhage.[36] Previous studies have shown that increased GFAP levels are promising for ICH diagnosis,[9,37–39] with a pooled specificity of 95% for differentiating ICH from AIS and MIM.[40] In our study, GFAP levels were higher in ICH compared with TIA and MIM (Tables S4 and S5) but not AIS. Given that a CT scan has nearly 100% specificity to rule out ICH but modest sensitivity (~30%) to diagnose AIS in the acute phase,[41] we conducted a secondary analysis excluding ICH cases. This analysis identified a panel of 35 proteins that differentiated AIS from MIM and TIA (Figure S9). This protein panel may be used to supplement neuroimaging in AIS diagnosis, where an ICH has already been excluded by CT. It is also important to note that despite high heterogeneity within the ICH group, nominated ICH biomarkers in our study were independent of baseline NIHSS score, suggesting that these biomarkers are less likely to be simply surrogates of the degree of neurological injury due to hemorrhage but rather biomarkers of unique biological aspects of ICH pathogenesis.

For TIA, early diagnosis is critical, as timely diagnostic evaluation and appropriate interventions can prevent up to 80% of stroke recurrences within the first 48 hours of the initial event.[42] We identified a panel of 8 proteins (TENM3, HSPA9, CTRB2, ADGRB2, UNC5C, NDUFV2, CRIP1, and PLG) as top TIA classifiers with rCV-AUC of 0.78 (95% CI, 0.73–0.84). This panel demonstrated high specificity (89% [95% CI, 86%–92%]) and NPV (88% [95% CI, 86%–90%]) for identifying non-TIA cases, indicating that it reliably recognizes patients who do not have TIA. PLG, involved in fibrinolysis and platelet activation, has been linked to transient ischemic events, consistent with our findings[43] and was also externally validated in the independent cohort as a TIA classifier (Table 4). CRIP1 is linked to hypertension pathophysiology and increased stroke risk,[44] which aligns with our findings showing decreased CRIP1 levels in TIA compared with AIS and ICH. Previous studies have identified proteins, including FABP (fatty-acid-binding protein), NPPB (natriuretic peptide B), NR2 antibodies, PARK7 (Parkinson disease protein 7), PON3 (paraoxonase 3), IGFBP3, apo CIII, IL-6 (interleukin-6), and vWF (von Willebrand factor) as promising biomarker candidates for TIA diagnosis.[12,45–47] Our study confirmed higher levels of NPPB (or BNP) and FABP1 in patients with TIA, whereas we identified IGFBP2 instead of IGFBP3 as a promising biomarker candidate to distinguish TIA from MIM. Although not part of our study, future work may further identify biomarkers that distinguish high-risk from low-risk patients with TIA to facilitate inpatient versus outpatient management decisions.

MIM account for a large proportion of acute neurological emergencies, which are often evaluated as "stroke codes" and require the recruitment of several health care resources emergently. Across centers in the United States, MIM accounts for a large fraction (>40%–50%) of stroke codes in the emergency setting. Therefore, validated biomarkers of MIM can be very useful in guiding acute triage decisions in the hyperacute setting as well as after initial evaluations and imaging studies are completed. Our study found that the majority of top-nominated proteins (27/61) in the discovery cohort differentiated MIM from other groups (AIS, ICH, TIA) (Table S9). Of these, our panel of 7 proteins (INIP, FAM20C, VTN, CXCL12, EGLN3, LILRA6, and CTSB) classified MIM from AIS, ICH, and TIA with rCV-AUC of 0.81 (95% CI, 0.77–0.86). This panel demonstrated high specificity (90% [95% CI, 87%–93%]) and high NPV (88% [95% CI, 85%–90%]) for identifying non-MIM cases and can be valuable in clinical settings to detect real strokes and rule out MIM. Importantly, higher levels of VTN in MIM compared with other stroke subtypes were internally and externally validated in an independent cohort. To our knowledge, this is the first study to highlight VTN's role as a strong MIM classifier. Limited evidence from a mouse stroke model indicates VTN may inhibit poststroke neurogenesis and aggravate neurological dysfunction.[48] PLG, a top classifier in the TIA panel, also showed increased levels in MIM in the discovery, internal, and external validation cohorts (Tables S11 and S12), consistent with previous findings.[43] Previous studies have identified biomarkers like NPPB, D-dimer, MMP-9 (matrix metalloproteinase-9), GFAP, S100B, apo CIII, and caspase-3 for differentiating MIM from real stroke, but

none achieved high classification accuracy.[9,32] Our study also identified lower NPPB (or BNP) levels in MIM compared with AIS and TIA (Tables S3 and S6).

Our study has several strengths. We used multiple statistical approaches, including hypothesis testing, variance partition, supervised machine learning, regularized LASSO logistic regression, internal validation using repeated nested cross-validation and targeted proteomics, and external validation using DIA-MS to identify DEPs across stroke subtypes. If externally validated in large-scale longitudinal studies, the protein panels identified in our study have the potential to reduce diagnostic uncertainty and improve the precision of treatments. For example, the TIA and MIM panels may aid in acute triage decision-making and optimizing resource allocation, particularly in the emergency department, where 20% to 50% of acute stroke emergencies are stroke mimics.[7] In the stroke setting, acute neuroimaging with head CT remains the gold standard to confirm or exclude ICH, and advanced CT-based imaging is essential for AIS characterization and treatment selection. Our findings are not intended to replace established neuroimaging workflows, guide endovascular therapy decisions, or alter standard clinical management pathways. Rather, our study provides proof-of-concept evidence that circulating plasma proteins capture biologically meaningful differences across major stroke diagnostic groups. A more plausible near-term application of these proteomic signatures is as a screening or biological stratification tool in research and early translational contexts, enabling molecular phenotyping, cohort enrichment, and hypothesis generation for future studies. Any clinical implementation would require extensive cross-platform external validation and the development of targeted biomarker panels compatible with rapid point-of-care multiplex assays using small plasma volumes.[39,49,50] Such platforms would allow reproducible, high-throughput, and cost-effective quantification of validated proteins directly within hospital laboratory workflows. We also acknowledge that our data do not evaluate feasibility or impact in prehospital settings, nor do they address prediction in healthy individuals. Dedicated prospective studies are warranted to assess operational utility and additional diagnostic value beyond standard clinical assessment and imaging, and potential roles in risk prediction.

Despite promising findings, our study has limitations. We conducted convenient sampling, and data on the exact time from symptom onset to blood sample collection were unavailable for the discovery cohort. However, we collected samples upon emergency department presentation before administering any therapeutic intervention. Despite relatively small sample sizes from each center, rigorous reproducibility supports the generalizability of our findings. Our derivation and validation cohorts were separated by study time frames that may have affected the biomarker replication rates across the 2 cohorts due to potential changes in treatment paradigms. Many peptides identified by SomaScan were not confirmed by BAK-270 and DIA-MS, as our targeted and untargeted MS validations were not prespecified. Targeted assays to detect these validated proteins in larger cohorts are needed. Given the asymmetry between the small sample size and the measurement of >7000 protein targets, the possibility of type 1 error cannot be ignored. However, we employed multiple statistical approaches to identify the most robust DEPs for stroke classification. There are also several inherent technical biases of aptamer-based and targeted and untargeted MS approaches, which may have affected the ability to cross-validate hits and facilitate the development of personalized diagnostic and therapeutic strategies. Additionally, the case–control design precludes causal inference, and longitudinal studies are needed to validate the prognostic utility of the identified biomarkers. We acknowledge that in our study, the groups with MIM and ICH are likely to be more heterogeneous than the groups with TIA and AIS, encompassing seizures, functional disorders, and migraines, and future studies are needed to determine whether proteins can further differentiate within subgroups with MIM. This is also likely the explanation for why TIA and MIM were better differentiated in our derivation cohort as compared with the validation cohort. Additionally, AIS and ICH are heterogeneous in cause, and future studies should explore cause-specific protein profiles in larger cohorts based on TOAST (Trial of ORG 10172 in Acute Stroke Treatment) classification and ICH subtypes (eg, lobar versus parenchymal, presence of intraventricular hemorrhage). Because we excluded traumatic brain injury and other forms of brain hemorrhage (subarachnoid, epidural, subdural), it is unclear if plasma biomarkers can be used for these conditions, although it is expected that brain CT imaging will most definitely be the gold standard in these settings. Although our study focused on plasma biomarkers, future studies incorporating other biofluids and multiomics approaches may provide a more comprehensive understanding of stroke pathophysiology.

CONCLUSIONS

Our study identified distinct plasma protein panels for classifying stroke subtypes (AIS, ICH, TIA, and MIM) with high accuracy. These proteomic signatures, validated through cross-platform methods, have the potential to improve stroke diagnosis, guide early intervention strategies, and reduce diagnostic uncertainty

in emergency settings. Further larger multicentric studies with targeted assays using these candidate biomarkers are needed to elucidate their clinical utility in guiding therapeutic decision-making and improving patient outcomes.

ARTICLE INFORMATION

Received November 26, 2025; accepted January 30, 2026.

Affiliations

Department of Neurology, Yale University School of Medicine, New Haven, CT (S.M., W.E.J., S.S., P.K., A.K., V.T.L., P.C., G.J.F., L.H.S., S.R.); Department of Human Genetics (A.N.) and Department of Neurology (P.K., M.L., C.M.W., M.R.F., S.R.), Emory University School of Medicine, Atlanta, GA; Lady Hardinge Medical College and affiliated hospitals, New Delhi, India (A.K.); and Bioinformatics Support Hub, Cushing/Whitney Medical Library, Yale University School of Medicine, New Haven, CT (R.G.).

Sources of Funding

None.

Disclosures

The authors declare no competing interests. The subject matter of this manuscript has been the basis for a pending provisional patent application filing.

Supplemental Material

Data S1. Supplemental Methods
Tables S1–S12
Figures S1–S19
Data Sets

REFERENCES

1. Collaborators GBDNSD. Global, regional, and national burden of disorders affecting the nervous system, 1990-2021: a systematic analysis for the Global Burden of Disease study 2021. *Lancet Neurol*. 2024;23:344–381. doi: 10.1016/S1474-4422(24)00038-3
2. Nowinski WL, Gupta V, Qian G, He J, Poh LE, Ambrosius W, Chrzan RM, Polonara G, Mazzoni C, Mol M, et al. Automatic detection, localization, and volume estimation of ischemic infarcts in noncontrast computed tomographic scans: method and preliminary results. *Investig Radiol*. 2013;48:661–670. doi: 10.1097/RLI.0b013e31828d8403
3. Emberson J, Lees KR, Lyden P, Blackwell L, Albers G, Bluhmki E, Brott T, Cohen G, Davis S, Donnan G, et al. Effect of treatment delay, age, and stroke severity on the effects of intravenous thrombolysis with alteplase for acute ischaemic stroke: a meta-analysis of individual patient data from randomised trials. *Lancet*. 2014;384:1929–1935. doi: 10.1016/S0140-6736(14)60584-5
4. Nogueira RG, Jadhav AP, Haussen DC, Bonafe A, Budzik RF, Bhuva P, Yavagal DR, Ribo M, Cognard C, Hanel RA, et al. Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct. *N Engl J Med*. 2018;378:11–21. doi: 10.1056/NEJMoa1706442
5. Albers GW, Marks MP, Kemp S, Christensen S, Tsai JP, Ortega-Gutierrez S, McTaggart RA, Torbey MT, Kim-Tenser M, Leslie-Mazwi T, et al. Thrombectomy for stroke at 6 to 16 hours with selection by perfusion imaging. *N Engl J Med*. 2018;378:708–718. doi: 10.1056/NEJMoa1713973
6. Ma L, Hu X, Song L, Chen X, Ouyang M, Billot L, Li Q, Malavera A, Li X, Munoz-Venturelli P, et al. The third intensive care bundle with blood pressure reduction in acute cerebral haemorrhage trial (INTERACT3): an international, stepped wedge cluster randomised controlled trial. *Lancet*. 2023;402:27–40. doi: 10.1016/S0140-6736(23)00806-1
7. Buck H, Akhtar N, Alrohimi A, Khan K, Shuaib A. Stroke mimics: incidence, aetiology, clinical features and treatment. *Ann Med*. 2021;53:420–436. doi: 10.1080/07853890.2021.1890205
8. Pohl M, Hesszenberger D, Kapus K, Meszaros J, Feher A, Varadi I, Pusch G, Fejes E, Tibold A, Feher G. Ischemic stroke mimics: a comprehensive review. *J Clin Neurosci*. 2021;93:174–182. doi: 10.1016/j.jocn.2021.09.025
9. Misra S, Montaner J, Ramiro L, Arora R, Talwar P, Nath M, Kumar A, Kumar P, Pandit AK, Mohania D, et al. Blood biomarkers for the diagnosis and differentiation of stroke: a systematic review and meta-analysis. *Int J Stroke*. 2020;15:704–721. doi: 10.1177/1747493020946157
10. Bustamante A, Penalba A, Orset C, Azurmendi L, Llombart V, Simats A, Pecharroman E, Ventura O, Ribo M, Vivien D, et al. Blood biomarkers to differentiate ischemic and hemorrhagic strokes. *Neurology*. 2021;96:e1928–e1939. doi: 10.1212/WNL.0000000000011742
11. Kalra LP, Khatter H, Ramanathan S, Sapehia S, Devi K, Kaliyaperumal A, Bal D, Sebastian I, Kakarla R, Singhania A, et al. Serum GFAP for stroke diagnosis in regions with limited access to brain imaging (BE FAST India). *Eur Stroke J*. 2021;6:176–184. doi: 10.1177/23969873211010069
12. Guaman-Pilco D, Chocano E, Pala E, Lamana-Vallverdu M, Penalba A, Garcia-Rodriguez P, Rubiera M, Bustamante A, Rovira A, Perez-Sanchez S, et al. H-FABP as a biomarker in transient ischemic attack. *J Cardiovasc Transl Res*. 2024;18:40–47. doi: 10.1007/s12265-024-10552-4
13. Hochrainer K, Yang W. Stroke proteomics: from discovery to diagnostic and therapeutic applications. *Circ Res*. 2022;130:1145–1166. doi: 10.1161/CIRCRESAHA.122.320110
14. Misra S, Singh P, Nath M, Bhalla D, Sengupta S, Kumar A, Pandit AK, Aggarwal P, Srivastava AK, Mohania D, et al. Blood-based protein biomarkers for the diagnosis of acute stroke: a discovery-based SWATH-MS proteomic approach. *Front Neurol*. 2022;13:989856. doi: 10.3389/fneur.2022.989856
15. Candia J, Daya GN, Tanaka T, Ferrucci L, Walker KA. Assessment of variability in the plasma 7k SomaScan proteomics assay. *Sci Rep*. 2022;12:17147. doi: 10.1038/s41598-022-22116-0
16. Zhang H, Liu Q, Zimmerman LJ, Ham AJ, Slebos RJ, Rahman J, Kikuchi T, Massion PP, Carbone DP, Billheimer D, et al. Methods for peptide and protein quantitation by liquid chromatography-multiple reaction monitoring mass spectrometry. *Mol Cell Proteomics*. 2011;10:M110 006593. doi: 10.1074/mcp.M110.006593
17. Gaither C, Popp R, Mohammed Y, Borchers CH. Determination of the concentration range for 267 proteins from 21 lots of commercial human plasma using highly multiplexed multiple reaction monitoring mass spectrometry. *Analyst*. 2020;145:3634–3644. doi: 10.1039/c9an01893j
18. Percy AJ, Borchers CH. Detailed method for performing the ExSTA approach in quantitative bottom-up plasma proteomics. *Methods Mol Biol*. 2021;2228:353–384. doi: 10.1007/978-1-0716-1024-4_25
19. Gold L, Ayers D, Bertino J, Bock C, Bock A, Brody EN, Carter J, Dalby AB, Eaton BE, Fitzwater T, et al. Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS One*. 2010;5:e15004. doi: 10.1371/journal.pone.0015004
20. Tin A, Yu B, Ma J, Masushita K, Daya N, Hoogeveen RC, Ballantyne CM, Couper D, Rebholz CM, Grams ME, et al. Reproducibility and variability of protein analytes measured using a multiplexed modified aptamer assay. *J Appl Lab Med*. 2019;4:30–39. doi: 10.1373/jalm.2018.027086
21. Rohloff JC, Gelinas AD, Jarvis TC, Ochsner UA, Schneider DJ, Gold L, Janjic N. Nucleic acid ligands with protein-like side chains: modified aptamers and their use as diagnostic and therapeutic agents. *Mol Ther Nucleic Acids*. 2014;3:e201. doi: 10.1038/mtna.2014.49
22. Kursa MB, Rudnicki WR. Feature selection with the Boruta package. *J Stat Softw*. 2010;36:1–13. doi: 10.18637/jss.v036.i11
23. Tibshirani R. Regression shrinkage and selection via the Lasso. *J Roy Stat Soc: Series B (Methodological)*. 2018;58:267–288. doi: 10.1111/j.2517-6161.1996.tb02080.x
24. Belloni A, Chernozhukov V, Hansen C. Inference on treatment effects after selection among high-dimensional controls†. *Rev Econ Stud*. 2013;81:608–650. doi: 10.1093/restud/rdt044
25. Le Cao KA, Boitard S, Besse P. Sparse PLS discriminant analysis: biologically relevant feature selection and graphical displays for multiclass problems. *BMC Bioinformatics*. 2011;12:253. doi: 10.1186/1471-2105-12-253
26. Demichev V, Messner CB, Vernardis SI, Lilley KS, Ralser M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods*. 2020;17:41–44. doi: 10.1038/s41592-019-0638-x
27. Stekhoven DJ, Buhlmann P. MissForest--non-parametric missing value imputation for mixed-type data. *Bioinformatics*. 2012;28:112–118. doi: 10.1093/bioinformatics/btr597
28. Wallace JA, Alexander S, Estrada EY, Hines C, Cunningham LA, Rosenberg GA. Tissue inhibitor of metalloproteinase-3 is associated

with neuronal death in reperfusion injury. *J Cereb Blood Flow Metab*. 2002;22:1303–1310. doi: 10.1097/01.WCB.0000040943.89393.c1
29. Xiao L, Zou X, Liang Y, Wang Y, Zeng L, Wu J. Evaluating the causal effects of TIMP-3 on Ischaemic stroke and intracerebral haemorrhage: a Mendelian randomization study. *Front Genet*. 2022;13:838809. doi: 10.3389/fgene.2022.838809
30. Marto JP, Carvalho AS, G. Mollet I, Mendonca M, Salavisa M, Meira B, Fernandes M, Serrazina F, Cabral G, Ventura R, et al. Proteomics to identify new blood biomarkers for diagnosing patients with acute stroke. *J Am Heart Assoc*. 2023;12:e030021. doi: 10.1161/JAHA.123.030021
31. Nunez-Jurado D, Fernandez-Vega A, Del Rio C, Penalba A, Llucia-Carol L, Muino-Acuna E, Ezcurra-Diaz G, Guasch-Jimenez M, Cullell N, Serrano-Heras G, et al. Plasma proteomics uncovers divergent molecular signatures in ischemic stroke and intracerebral hemorrhage. *Biomark Res*. 2025;13:136. doi: 10.1186/s40364-025-00848-1
32. Bustamante A, Lopez-Cancio E, Pich S, Penalba A, Giralt D, Garcia-Berrocoso T, Ferrer-Costa C, Gasull T, Hernandez-Perez M, Millan M, et al. Blood biomarkers for the early diagnosis of stroke: the stroke-Chip study. *Stroke*. 2017;48:2419–2425. doi: 10.1161/STROKEAHA.117.017076
33. Llombart V, Garcia-Berrocoso T, Bustamante A, Giralt D, Rodriguez-Luna D, Muchada M, Penalba A, Boada C, Hernandez-Guillamon M, Montaner J. Plasmatic retinol-binding protein 4 and glial fibrillary acidic protein as biomarkers to differentiate ischemic stroke and intracerebral hemorrhage. *J Neurochem*. 2016;136:416–424. doi: 10.1111/jnc.13419
34. Montaner J, Mendioroz M, Delgado P, Garcia-Berrocoso T, Giralt D, Merino C, Ribo M, Rosell A, Penalba A, Fernandez-Cadenas I, et al. Differentiating ischemic from hemorrhagic stroke using plasma biomarkers: the S100B/RAGE pathway. *J Proteome*. 2012;75:4758–4765. doi: 10.1016/j.jprot.2012.01.033
35. Feng L, Li M, Ren J, Li Y, Wang Q, Zhang P, Zhang X, Wang T, Li Y. Proteomic analysis reveals that Di dang decoction protects against acute intracerebral hemorrhage stroke in rats by regulating S100a8, S100a9 Col1a1, and Col1a2. *Neuropsychiatr Dis Treat*. 2021;17:3301–3314. doi: 10.2147/NDT.S331688
36. Chen D, Wang X, Huang J, Cui S, Zhang L. CDKN1B mediates apoptosis of neuronal cells and inflammation induced by oxyhemoglobin via miR-502-5p after subarachnoid hemorrhage. *J Mol Neurosci*. 2020;70:1073–1080. doi: 10.1007/s12031-020-01512-z
37. Misra S, Kumar A, Kumar P, Yadav AK, Mohania D, Pandit AK, Prasad K, Vibha D. Blood-based protein biomarkers for stroke differentiation: a systematic review. *Proteomics Clin Appl*. 2017;11:11. doi: 10.1002/prca.201700007
38. Paul JF, Ducroux C, Correia P, Daigneault A, Larochelle C, Stapf C, Gioia LC. Serum glial fibrillary acidic protein in acute stroke: feasibility to determine stroke-type, timeline and tissue-impact. *Front Neurol*. 2024;15:1470718. doi: 10.3389/fneur.2024.1470718
39. Kalra LP, Zylyftari S, Blums K, Barthelmes S, Baum H, Meckel S, Heilgeist A, Luger S, Foerch C. Rapid diagnosis of intracerebral hemorrhage in patients with acute stroke by measuring prehospital GFAP levels on a point-of-care device (DETECT). *Neurology*. 2025;105:e213823. doi: 10.1212/WNL.0000000000213823
40. Kumar A, Misra S, Yadav AK, Sagar R, Verma B, Grover A, Prasad K. Role of glial fibrillary acidic protein as a biomarker in differentiating intracerebral haemorrhage from ischaemic stroke and stroke mimics: a meta-analysis. *Biomarkers*. 2020;25:1–8. doi: 10.1080/1354750X.2019.1691657
41. Patel SC, Levine SR, Tilley BC, Grotta JC, Lu M, Frankel M, Haley EC Jr, Brott TG, Broderick JP, Horowitz S, et al. Lack of clinical significance of early ischemic changes on computed tomography in acute stroke. *JAMA*. 2001;286:2830–2838. doi: 10.1001/jama.286.22.2830
42. Coutts SB. Diagnosis and management of transient ischemic attack. *Continuum (Minneap Minn)*. 2017;23:82–92. doi: 10.1212/CON.0000000000000424
43. Penn AM, Saly V, Trivedi A, Lesperance ML, Votova K, Jackson AM, Croteau NS, Balshaw RF, Bibok MB, Smith DS, et al. Differential proteomics for distinguishing ischemic stroke from controls: a pilot study of the SpecTRA project. *Transl Stroke Res*. 2018;9:590–599. doi: 10.1007/s12975-018-0609-z
44. Zeller T, Schurmann C, Schramm K, Muller C, Kwon S, Wild PS, Teumer A, Herrington D, Schillert A, Iacoviello L, et al. Transcriptome-wide analysis identifies novel associations with blood pressure. *Hypertension*. 2017;70:743–750. doi: 10.1161/HYPERTENSIONAHA.117.09458
45. Dolmans LS, Rutten FH, Koenen NCT, Bartelink MEL, Reitsma JB, Kappelle LJ, Hoes AW. Candidate biomarkers for the diagnosis of transient ischemic attack: a systematic review. *Cerebrovasc Dis*. 2019;47:207–216. doi: 10.1159/000502449
46. Penn AM, Bibok MB, Saly VK, Coutts SB, Lesperance ML, Balshaw RF, Votova K, Croteau NS, Trivedi A, Jackson AM, et al. Validation of a proteomic biomarker panel to diagnose minor-stroke and transient ischaemic attack: phase 2 of SpecTRA, a large scale translational study. *Biomarkers*. 2018;23:793–803. doi: 10.1080/1354750X.2018.1499130
47. Guamán-Pilco D, Palà E, Lamana-Vallverdú M, Penalba A, García-Rodríguez P, Rubiera M, Bustamante A, Pérez-Sánchez S, Montaner J. A panel of blood biomarkers for the diagnosis of transient ischemic attacks. *CVIA*. 2025;10. doi: 10.15212/cvia.2024.0061
48. Jia C, Lovins C, Malone HM, Keasey MP, Hagg T. Female-specific neuroprotection after ischemic stroke by vitronectin-focal adhesion kinase inhibition. *J Cereb Blood Flow Metab*. 2022;42:1961–1974. doi: 10.1177/0271678X221107871
49. Nadim G, Laursen CB, Pietersen PI, Wittrock D, Sorensen MK, Nielsen LB, Rasmussen CH, Christensen HM, Helmerik S, Jorgensen G, et al. Prehospital emergency medical technicians can perform ultrasonography and blood analysis in prehospital evaluation of patients with chronic obstructive pulmonary disease: a feasibility study. *BMC Health Serv Res*. 2021;21:290. doi: 10.1186/s12913-021-06305-7
50. Alsaikhan F. Stroke biomarkers: the promise of point-of-care testing. *Clin Chim Acta*. 2025;579:120654. doi: 10.1016/j.cca.2025.120654