Journal of the American Heart Association

ORIGINAL RESEARCH

Proteomics to Identify New Blood Biomarkers for Diagnosing Patients With Acute Stroke

João Pedro Marto, MD; Ana Sofia Carvalho, PhD; Inês G. Mollet, PhD; Marcelo Mendonça, MD, PhD; Manuel Salavisa, MD; Bruna Meira, MD; Marco Fernandes, MD; Filipa Serrazina, MD; Gonçalo Cabral, MD; Rita Ventura, MD; André Sobral-Pinho, MD; Hans C. Beck, PhD; Helena L. A. Vieira, PhD; Miguel Viana-Baptista, MD, PhD*; Rune Matthiesen, PhD*

BACKGROUND: Blood biomarkers are a potential tool for early stroke diagnosis. We aimed to perform a pilot and exploratory study on untargeted blood biomarkers in patients with suspected stroke by using mass spectrometry analysis.

METHODS AND RESULTS: This was a prospective observational study of consecutive patients with suspected stroke admitted within 6 hours of last being seen well. Blood samples were collected at admission. Patients were divided into 3 groups: ischemic stroke (IS), intracerebral hemorrhage (ICH), and stroke mimics. Quantitative analysis from mass spectrometry data was performed using a supervised approach. Biomarker-based prediction models were developed to differentiate IS from ICH and ICH+stroke mimics. Models were built aiming to minimize misidentification of patients with ICH as having IS. We included 90 patients, one-third within each subgroup. The median age was 71 years (interquartile range, 57–81 years), and 49 participants (54.4%) were women. In quantitative analysis, C3 (complement component 3), ICAM-2 (intercellular adhesion molecule 2), PLGLA (plasminogen like A), STXBP5 (syntaxin-binding protein 5), and IGHV3-64 (immunoglobulin heavy variable 3-64) were the 5 most significantly dysregulated proteins for both comparisons. Biomarker-based models showed 88% sensitivity and 89% negative predictive value for differentiating IS from ICH, and 75% sensitivity and 95% negative predictive value for differentiating IS from ICH+stroke mimics. ICAM-2, STXBP5, PLGLA, C3, and IGHV3-64 displayed the highest importance score in our models, being the most informative for identifying patients with stroke.

CONCLUSIONS: In this proof-of-concept and exploratory study, our biomarker-based prediction models, including ICAM-2, STXBP5, PLGLA, C3, and IGHV3-64, showed 75% to 88% sensitivity for identifying patients with IS, while aiming to minimize misclassification of ICH. Although our methodology provided an internal validation, these results still need validation in other cohorts and with different measurement techniques.

Key Words: biomarkers ■ intracerebral hemorrhage ■ ischemic stroke ■ proteomics

Stroke is a leading cause of disability and mortality worldwide.[1] For ischemic stroke (IS) treatment, 2 recanalization therapies, namely intravenous thrombolysis with recombinant tissue plasminogen activator (tPA) and mechanical thrombectomy, have been demonstrated to improve patients' outcomes.[2,3] Treatment effect is time-dependent, with a shorter time from stroke onset to recanalization associated with better outcomes.[2,3] In the case of intracerebral hemorrhage (ICH), early blood pressure lowering may reduce hematoma growth and improve clinical outcomes.[4]

Correspondence to: João Pedro Marto, MD, Department of Neurology, Hospital de Egas Moniz, Centro Hospitalar Lisboa Ocidental, Rua da Junqueira n° 126, 1349-019, Lisbon, Portugal. Email: joao.pedro.seabra.marto@gmail.com

*M. Viana-Baptista and R. Matthiesen contributed equally.

This manuscript was sent to Jose R. Romero, MD, 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.123.030021

For Sources of Funding and Disclosures, see page 12.

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

CLINICAL PERSPECTIVE

What Is New?

- In our pilot and exploratory study of untargeted blood biomarkers in patients with suspected stroke by using mass spectrometry analysis, our proteomic analysis and biomarker-based prediction models, including intercellular adhesion molecule 2, syntaxin-binding protein 5, plasminogen like A, complement component 3, and immunoglobulin heavy variable 3-64, showed 75% to 88% sensitivity for identifying patients with ischemic stroke, while aiming to minimize misclassification of patients with intracerebral hemorrhage as having ischemic stroke.

What Are the Clinical Implications?

- In a hospital setting, after imaging confirms the absence of an intracranial hemorrhage, our results could help in guiding the decision to administer intravenous thrombolysis by assisting in the differentiation of ischemic stroke from stroke mimics, which can be especially relevant in patients with risk factors for bleeding.
- These are preliminary results that need further validation in larger cohorts, in different clinical scenarios, and with other measurement techniques.

Nonstandard Abbreviations and Acronyms

C3	complement component 3
GFAP	glial fibrillary acid protein
ICAM-1	intercellular adhesion molecule 1
ICAM-2	intercellular adhesion molecule 2
ICH	intracerebral hemorrhage
IGHV3-64	immunoglobulin heavy variable 3-64
IS	ischemic stroke
mRS	modified Rankin Scale
MS	mass spectrometry
NIHSS	National Institutes of Health Stroke Scale
PLGLA	plasminogen like A
PLS	partial least squares
PSM	peptide spectrum match
RBP4	retinol-binding protein 4
SM	stroke mimics
STXBP5	syntaxin-binding protein 5

For early tPA treatment, mobile stroke units allow the administration of tPA at the prehospital level, reducing time to treatment and improving patients' outcomes.[5] However, the need for an ambulance equipped with a computed tomography scanner to accurately rule out ICH, together with a specialized health professional team, raises relevant financial and technical limitations that affect its widespread applicability.

Blood-based biomarkers represent a potential alternative for early and rapid stroke diagnosis. Accurate and timely differentiation of IS, ICH, and stroke mimics (SM) could improve pre-hospital pathways and referral systems and potentially allow early treatment administration.

Different biomarkers or biomarker panels have shown promising results. GFAP (glial fibrillary acid protein), RBP4 (retinol-binding protein 4), N-terminal pro-B-type natriuretic peptide (NT-proBNP), endostatin, caspase-3, and D-dimer have demonstrated some level of accuracy in the differential diagnosis of stroke in the acute setting.[6–10] These biomarkers were identified based on previous experimental and clinical data, but, to date, no biomarker or biomarker panel with a clinically significant predictive accuracy for IS or ICH diagnosis has been validated.

Disclosing new potential biomarkers by using different new research approaches, such as mass spectrometry (MS)–based proteomics, may improve the reliability of the molecular diagnosis of stroke, contributing to the applicability of blood-based biomarkers in stroke care.

In this sense, there is only 1 preliminary study aimed at differentiating ICH from IS based on label-free proteomics of plasma from 3 patients with IS and 4 patients with ICH,[11] and not in the acute setting.

Herein, we performed a pilot and exploratory study on untargeted blood biomarkers in patients with suspected stroke. Our aim was to find new potential biomarkers that can help improve the sensitivity and specificity of blood-based biomarkers in the diagnosis of patients with suspected acute stroke.

METHODS

Study Design, Patient Selection, and Study Variables

The data that support the findings of this study are available upon reasonable request to the corresponding author. The authors have full access to all of the data in the study and take responsibility for their integrity and the data analysis.

This was a prospective, observational study of consecutive patients with a suspected stroke admitted to the Centro Hospitalar Lisboa Ocidental Emergency Department as stroke code patients, from November 2018 to December 2019. Patients were screened if they met the following criteria: (1) age 18 years and older; (2) admitted within 6 hours of symptom onset; and (3) no

history of a cerebrovascular or major vascular event within the 3 months (myocardial infarction, systemic embolism, venous thrombosis). Patients without stroke code activation were not included, nor were patients transferred from other hospitals.

After admission, a definitive clinical and imaging-supported diagnosis was established, and patients were divided into 3 groups: (1) IS; (2) ICH; and (3) SM. IS diagnosis was established according to the World Health Organization definition,[12] together with neuroimaging confirmation of a new ischemic brain lesion in a clinically relevant area. ICH was defined as an acute neurological syndrome caused by a confirmed brain parenchymal hemorrhage. SM diagnosis was established by trained neurologists and supported by ancillary tests deemed necessary in each case. Patients without a definitive diagnosis at hospital discharge, without good quality samples for MS analysis, or without signed informed consent were further excluded.

For the purpose of our exploratory study, we established a target of enrolling 30 patients within each subgroup. Therefore, recruitment for each subgroup was stopped when 30 patients met the inclusion criteria, had established definitive diagnoses, and provided good quality blood samples. For patients' inclusion flow chart, see Figure S1.

The following variables were collected: age, sex, vascular risk factors, admission blood pressure, and stroke severity at admission (assessed by the National Institutes of Health Stroke Scale [NIHSS]). In addition, clinical variables at hospital admission, previously studied to differentiate IS from ICH (headache, vomiting, decreased level of consciousness, epileptic seizure)[13] and IS from SM (isolated sensory deficit, facial palsy, history of seizures)[14,15] were collected. For patients with IS, presence of intracranial vessel occlusion, recanalization treatment, and stroke cause were recorded. For patients with ICH, probable cause was assessed. For SM, the final diagnosis was defined. In patients with stroke, outcome was evaluated at 3 months by the modified Rankin Scale (mRS) and defined as favorable if the mRS score was ≤2. Finally, time from symptom onset to blood collection was registered for all patients.

Standard Protocol Approvals, Registrations, and Patient Consents

The study protocol was approved by the Centro Hospital Lisboa Ocidental (registry number: 20170700050) and the NOVA Medical School's ethics committee (registry number: 85/2021/CEFCM). All patients or relatives signed a written informed consent, before or after blood sample collection. As stated, if informed consent was not obtained, patients were excluded from the study. This study was conducted according to the principles of the Declaration of Helsinki.

Data Availability

Anonymized data not published within this article will be made available by request, to the corresponding author, from any qualified investigator.

Statistical Analysis

We summarized continuous variables as median values with interquartile range and categorical variables as absolute numbers and percentages. We compared baseline characteristics between the 3 groups using Pearson χ^2 test and Fisher exact test for categorical variables and Kruskall–Wallis test for continuous variables, as appropriate. The Shapiro–Wilk test was used to assess the normality of continuous variables. All tests were 2-sided, and P values <0.05 were considered significant. Bonferroni correction was used for adjusting all P values in multiple comparisons. We performed statistical analysis with R statistical software version 4.0.3 (R Project Statistical Computing).

Blood Collection and Processing

Blood samples were collected at hospital admission before any therapeutic intervention or brain imaging, into EDTA tubes. Blood samples were centrifuged at 1500*g* for 15 minutes at 4 °C, and plasma samples were aliquoted in the presence of protease inhibitors and stored at −80 °C until further processing. The 14 most abundant proteins in plasma were eliminated using a commercially available kit (Agilent Human 14 Multiple Affinity Removal System Spin Cartridges for the High-Abundant Proteins from Human Proteomic Samples) following the manufacturer's instructions. Plasma samples were further treated for protein digestion. Protein solutions containing sodium dodecyl sulfate and dithiothreitol were loaded onto filtering columns and washed exhaustively with 8M urea in HEPES buffer.[16] Proteins were reduced with dithiothreitol and alkylated with iodoacetamide. Protein digestion was performed by overnight digestion with trypsin sequencing grade (Promega).

MS Analysis

Each sample was analyzed in duplicate except for 1, which was analyzed in triplicate. Peptide samples were analyzed by nano-liquid chromatography–MS/MS (Dionex RSLCnano 3000) coupled with an Exploris 480 Orbitrap mass spectrometer (Thermo Scientific). In brief, the samples (5 μL) were loaded onto a custom-made fused capillary precolumn (2 cm length, 360 μm outer diameter, 75 μm ID) packed with ReproSil Pur C18 5.0 μm resin (Dr Maish) with a flow of 5 μL per minute for 6 minutes. Trapped peptides were separated on a custom-made fused capillary column (25 cm length, 360 μm outer diameter, 75 μm inner diameter) packed

with ReproSil Pur C18 1.9 μm resin (Dr Maish) with a flow of 250 nL per minute using a linear gradient from 89% A (0.1% formic acid) to 32% B (0.1% formic acid in 80% acetonitrile) over 30 minutes. Mass spectra were acquired in positive ion mode applying automatic data-dependent switch between one Orbitrap survey MS scan in the mass range of 350 to 1200 *m/z* followed by higher-energy collision dissociation fragmentation and Orbitrap detection of fragment ions with a cycle time of 2 seconds between each master scan. MS and MS/MS maximum injection time were set to "auto," and higher-energy collision dissociation fragmentation in the ion routing multipole was performed at a normalized collision energy of 30%, and ion selection threshold was set to 10 000 counts. For 30 seconds, selected sequenced ions were dynamically excluded.

Protein Identification

The obtained data from the 181 liquid chromatography–MS runs were searched using virtual expert mass spectrometrist.[17,18] A standard human proteome database from UniProt (3AUP000005640) including permutated protein sequences, where Arg and Lys were not permutated, was used as a sequence database. For trypsin cleavages, a maximum of 4 missed cleavages were permitted. Carbamidomethyl cysteine was included as fixed modification. Methionine oxidation and N-terminal protein acetylation were included as variable modifications. Ten ppm mass accuracy was specified for precursor ions and 0.01 *m/z* for fragment ions. The false discovery rate for protein identification was set at 1% for peptide and protein identifications. No restriction was applied to the minimal peptide length for virtual expert mass spectrometrist search. Identified proteins were divided into evidence groups as previously defined.[18] Briefly, proteins with at least 1 peptide spectrum match (PSM) uniquely mapped are assigned to the highest evidence group 1. Proteins that share exactly the same set of PSMs are defined as semi-unique and assigned to evidence group 2. Proteins that are connected by shared PSMs but none of them are unique or semi-unique but needed in order to explain all of the identified PSMs are assigned evidence groups 3. Proteins with only PSMs that are already mapped to evidence group 1 to 3 are assigned to evidence group 4. Protein identifications resulting from PSMs that represent the second-best peptide match to the spectrum are assigned to evidence group 5 (only relevant for some search engines). In this study, only proteins in evidence groups 1 to 3 were maintained for further analysis.

To avoid sample-related biases in the biomarker analysis, we used reference proteomes of erythrocytes, platelets, and plasma[19] to check for potential biases.

Quantitative Analysis

Quantitative data from virtual expert mass spectrometrist were analyzed in the R statistical programming language. Intensity-based absolute quantification (iBAQ)[20] and protein spectral counts from the virtual expert mass spectrometrist were preprocessed by 3 approaches: (1) removing common MS contaminants followed by $\log_2(x+1)$ transformation; (2) removing common MS contaminants followed by $\log_2(x+1)$ transformation and quantile normalization (Figure S2); and (3) removing common MS contaminants followed by $\log_2(x+1)$ transformation, quantile normalization, and abundance filtering to optimize overall Gaussian distribution of the quantitative values. The function "normalize.quantiles" in the R package "preprocessCore" was applied for quantile normalization of the data for making the distributions' statistical properties more similar.[21] Quantitative values from technical replicas were averaged. Quantitative values from technical replicas were averaged. Because the majority of the iBAQ values were in the range of 10^4 to 10^9, adding a small value such as "one" [*to avoid* $log_2(0)$] to all observations would not importantly change these values and enable us to perform the $\log_2$ transform of the iBAQ values. The average values were not subjected to additional imputation to replace zero values. After confirming similar results from the 3 data processing approaches described above, quantitative data from iBAQ data were used for further statistical analysis using the R package "limma" after removing common MS contaminants, followed by $\log_2(x+1)$ transformation and quantile normalization (Figure S2). The distribution of the quantitative values among individual samples is depicted as a boxplot (Figure S2). Contrasts between "IS and ICH," "IS and SM," and "ICH and SM" were calculated. Correction for multiple testing was applied using the method of Benjamini and Hochberg.[22]

PLS Comparator Models

The R package "caret" was used to generate 6 supervised partial least squares (PLS) comparator models to test potential classification performance based on different input data.[23] Three models were built based on separating patients with IS from patients with ICH and 3 models were built for separating patients with IS from those with ICH and SM. For each comparison, models were based on clinical data only, the 10 most significant proteins based on training data, and all proteins. For both clinical models (differentiating IS from ICH and IS from ICH and SM) the following variables were used: vomiting, headache, epileptic seizure, decreased level of consciousness, diastolic blood pressure ≥110 mm Hg, systolic blood pressure ≥150 mm Hg, and history of stroke/transient ischemic attack. In the model for differentiating IS from ICH and SM, age

younger than 50 years, atrial fibrillation, history of epileptic seizures, isolated sensory deficit, and facial palsy were also included.

The data were randomly split for all models using 75% of the cases to train the models and 25% of the data as an unseen validation set. The data were preprocessed by removing zero variance cases, then scaled, centered, and $\log_2$ transformed. Statistical analysis based on the R package limma was used as a feature selection method based only on the training data.[24] Models were built aiming to minimize misidentification of patients with ICH as having IS due to safety reasons (eg, treating a patient with ICH with tPA).

PLS was chosen after testing for multiple classifiers with preliminary results showing high classification accuracy without significant differences compared with 7 other models (Figure S3).

PLS regression finds a set of new variables, referred to as latent variables or components, that correlate with the class labels and capture the majority of crucial information in the predictor variables. The initial predictor variables and the class labels are then combined to create these latent variables iteratively.

ROC curves were estimated by the R package "plotROC,"[25] which estimates exact CIs based on the Clopper and Pearson exact method.[26] The 95% CIs in the ROC curves were estimated at the point of maximum sensitivity where specificity was 100%. The "varImp" function from caret R package was applied calculation of variable importance scores based on the permutation method, which assesses the impact of permuting the values of each predictor variable on the model's performance. The resulting importance scores are scaled between 0 and 100, with higher scores indicating greater importance.

RESULTS

Participant Characteristics and Clinical Data

We included 90 patients, one-third within each subgroup. The median age was 71 years (interquartile range, 57–81 years), and 49 (54.4%) were women. Patients with ICH had higher systolic and diastolic blood pressure (P=0.008 and P=0.002, respectively), more frequently had headache (P<0.001), and less frequently had atrial fibrillation (P=0.013) than patients with IS. Patients with IS had a higher prevalence of hypertension (P=0.005), dyslipidemia (P=0.009), atrial fibrillation (P=0.002), smoking (P=0.010), and coronary artery disease (P=0.023), and higher NIHSS (P<0.001) compared with patients with SM. Patients with SM had a higher prevalence of history of seizures (P=0.005) and more frequently presented with isolated sensory symptoms (P=0.020), without facial palsy (P<0.001), and with acute seizures (P=0.010) than patients with IS (Table).

Half of the patients with IS presented with an identifiable intracranial vessel occlusion, 7 (23.3%) patients received isolated intravenous thrombolysis, 7 (23.3%) patients received bridging therapy, and 6 (20%) patients were treated with direct endovascular thrombectomy.

A favorable outcome was found in 14 (46.7%) patients with IS and 12 (40.0%) patients with ICH. For details on IS and ICH cause and SM final diagnosis, see Table S1.

Protein Identification

In total, 1606 protein isoforms were identified, which summed to 748 protein coding genes (Figure 1A). The total number of proteins was higher in patients with ICH compared with patients with IS or SM (Figure 1A). Of note, unique proteins identified within each group may be detected only in a subset of patients. Patients with IS and ICH shared more common protein identifications than they shared with patients with SM when analyzing the total identified proteins among all samples (Figure 1A). However, when analyzing the protein identifications at the individual liquid chromatography–MS run level among sample groups, it was not significant (Figure 1B).

Quality Control

Preliminary linear discriminant analysis based on the iBAQ values was able to provide strong separation between all sample groups (Figure S4). Linear discriminant analysis is a supervised classification method, and although all data were applied for training in this preliminary linear discriminant analysis, it suggested that reasonable classification into correct clinical sample groups was achievable based on MS data. No significant differences in the identification of bias proteins were observed among the 3 subgroups (Figure S5), supporting the hypothesis that the statistical differences observed between groups were not caused by blood collection or experimental artifacts.

Protein Quantitation

In quantitative analysis, the comparison between samples of patients with IS and ICH or IS and SM resulted in a higher number of significantly regulated proteins compared with the comparison between ICH and SM (Figure 2). Comparison between samples of patients with ICH and SM resulted in no significant differences after correction for multiple testing (Figure 2). Focusing on differentiating patients with IS from those with ICH and SM, we looked at the 5 most significantly dysregulated proteins. From the 2 comparisons, the same 5 proteins were identified—C3 (complement component

Table. Baseline Characteristics

	Total (N=90)	IS (n=30)	ICH (n=30)	SM (n=30)	*P* value	*P* value*	*P* value†
Demographics							
Age, y	71 (57–81)	73 (69–83)	70 (60–80)	62 (51–76)	0.120	1.000	0.127
Female sex	49 (54.4)	13 (43.3)	16 (53.3)	20 (66.7)	0.200	0.606	0.119
Vascular risk factors and medical history							
Hypertension	69 (76.7)	26 (86.7)	27 (90.0)	16 (53.3)	0.001	1.000	0.005
Diabetes	23 (25.6)	9 (30.0)	8 (26.7)	6 (20.0)	0.700	1.000	0.552
Dyslipidemia	35 (38.9)	18 (60.0)	9 (30.0)	8 (26.7)	0.014	0.037	0.009
Atrial fibrillation	16 (17.8)	11 (36.7)	2 (6.7)	3 (10.0)	0.004	0.013	0.002
Smoking	16 (17.8)	10 (33.3)	4 (13.3)	2 (6.7)	0.019	0.125	0.010
Coronary artery disease	11 (12.2)	7 (23.3)	3 (10.0)	1 (3.3)	0.072	0.299	0.023
Previous stroke/TIA	8 (8.9)	4 (13.3)	0 (0.0)	4 (13.3)	0.120	0.038	1.000
History of seizures	7 (7.8)	0 (0.0)	0 (0.0)	7 (23.3)	<0.001	1.000	0.005
Admission characteristics							
NIHSS	9 (3–17)	14 (6–19)	12 (9–22)	3 (2–5)	<0.001	0.443	<0.001
Presence of facial palsy	64 (71.1)	29 (96.7)	29 (96.7)	6 (20)	<0.001	1.000	<0.001
Isolated sensory symptoms	5 (5.6)	0 (0.0)	0 (0.0)	5 (16.7)	0.010	1.000	0.020
SBP, mm Hg	161 (134–188)	157 (146–177)	186 (160–220)	136 (126–163)	<0.001	0.008	0.319
SBP ≥150 mm Hg	55 (61.1)	20 (66.7)	25 (83.3)	10 (33.3)	<0.001	0.233	0.017
DBP, mm Hg	86 (74–100)	83 (70–94)	102 (88–117)	79 (73–86)	<0.001	0.002	1.000
DBP ≥110 mm Hg	15 (16.7)	2 (6.7)	13 (43.3)	0 (0.0)	<0.001	0.001	0.491
Decreased level of consciousness	13 (14.4)	4 (13.3)	9 (30.0)	0 (0.0)	0.003	0.209	0.112
Headache	9 (30.0)	0 (0.0)	12 (40.0)	3 (10.0)	<0.001	<0.001	0.237
Vomiting	14 (15.6)	3 (10.0)	9 (30.0)	2 (6.7)	0.052	0.104	1.000
Acute seizures	7 (7.8)	0 (0.0)	1 (3.3)	6 (20.0)	0.015	1.000	0.010
Onset to blood collection, min	90 (72–131)	97 (76–129)	77 (64–108)	112 (75–151)	0.119	0.151	1.000

Values are presented as median (interquartile range) or number (percentage). DBP indicates diastolic blood pressure; NIHSS, National Institutes of Health Stroke Scale; SBP, systolic blood pressure; and TIA, transient ischemic attack.

*Comparison between acute ischemic stroke (IS) and intracerebral hemorrhage (ICH).

†Comparison between IS and stroke mimic (SM).

3), ICAM-2 (intercellular adhesion molecule 2), PLGLA (plasminogen like A), STXBP5 (syntaxin-binding protein 5), and IGHV3-64 (immunoglobulin heavy variable 3-64) (Figure 3). C3, PLGLA, STXBP5, and IGHV3-64 were upregulated in patients with IS, and ICAM-2 was downregulated.

Classification

As a proof of concept, a PLS classifier was built to differentiate patients with IS from patients with ICH and SM. The first 3 models were designed to differentiate patients with IS and ICH. As a comparator, a first model using exclusively clinical data showed a 12% sensitivity and 53% negative predictive value (NPV) for diagnosis of IS (Figure S6).

In a second model, based on the top 10 regulated proteins in the training data set (Figure 4), sensitivity and NPV largely improved to 88% and 89%, respectively. ICAM-2, STXBP5, PLGLA, C3, and IGHV3-64 were, in this order, the most relevant proteins in the model.

To evaluate the full potential of the plasma proteome, a third model, using all proteins, showed improved results and additional potential useful biomarkers for stroke diagnosis (Figure S7).

To differentiate patients with IS from those with ICH and SM, 3 additional models were built, using the same 3 variable categories (Figure S8 through S10). Again, biomarker-based models showed better performance than the model using exclusively clinical data. In a model based on the top 10 regulated proteins in the training data (Figure S9), sensitivity and NPV were 75% and 95%, respectively. ICAM-2, PLGLA, C3, and IGHV3-64 were, in this order, the most important proteins in the model. The model using the full proteome showed perfect discrimination with 100% sensitivity and NPV (Figure S10).

To elucidate the complexity of the decision boundary of the top markers used in these models, data from all samples were visualized in a scatter plot with pairwise expression values. Figure 5 displays an example of a scatter plot of iBAQ intensity values, in this case for ICAM-2 and PLGLA. The results suggest that a

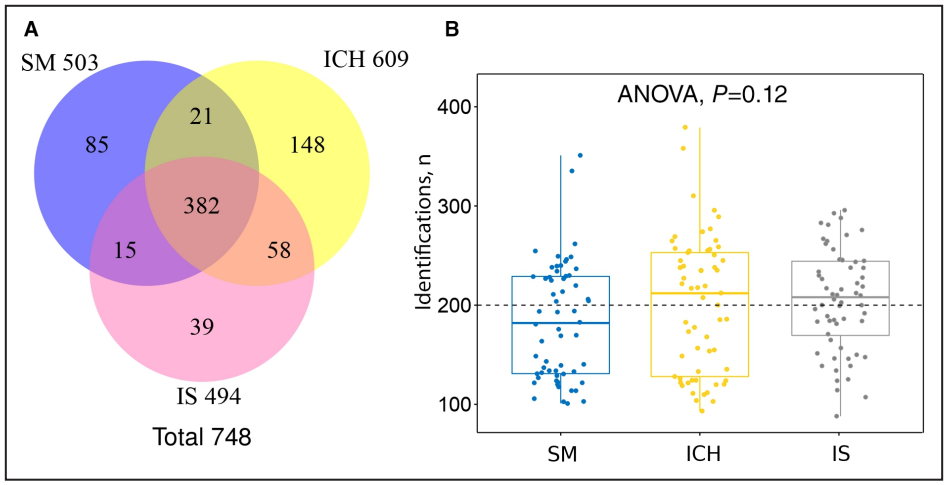

Figure 1. Total number of identifications across sample groups and in individual runs.
A, Venn diagram of overlapping proteins identified in IS, ICH, and SM. **B**, Boxplot of number of identifications in individual liquid chromatography–mass spectrometry runs vs sample groups. Boxplots indicate the median and quartiles, with whiskers indicating the 1.5 interquartile ranges. The dotted lines indicate the global medians. ICH indicates intracerebral hemorrhage; IS, ischemic stroke; and SM, stroke mimics.

simple classifier based on thresholds for 2 to 3 protein markers may provide a suitable separation of IS from ICH and SM. The linear decision boundaries provide additional evidence that the 2 plotted variables ICAM-2 and PLGLA have a relevant discriminatory power and influence the classification outcome.

DISCUSSION

In our pilot and exploratory study, using a proteomics analysis approach in patients with suspected acute stroke, we identified new potential targets for identifying patients with IS. Using a limited number of proteins, our models showed a sensitivity of 88% and an NPV of 89% for separating patients with IS and ICH, and a sensitivity of 75% and an NPV of 95% for differentiating patients with IS and ICH or SM. All models were created aiming to minimize misclassification of ICH as IS, which could potentially lead to harmful treatment (eg, giving tPA to a patient with ICH). To our knowledge, no previous blood-based biomarker analysis showed these high levels of sensitivity and specificity for identifying patients with acute IS.

Our models using the full proteome showed better results, including 100% sensitivity for differentiating

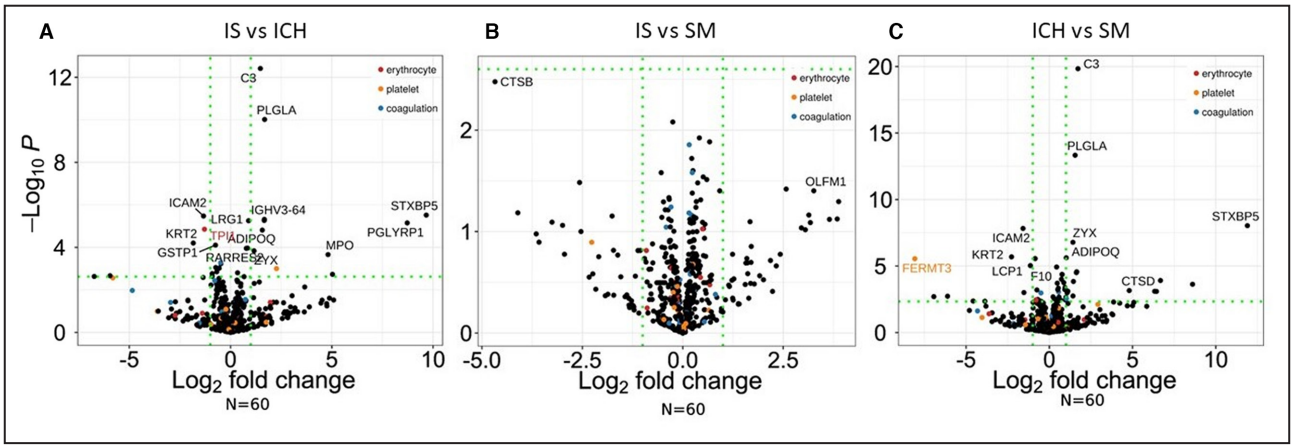

Figure 2. Volcano plots based on the 3 pairwise comparisons.
A, IS vs ICH, **B**, IS vs SM, and **C**, ICH vs SM. Dotted horizontal line indicates (green) an adjusted *P* value of 0.05—proteins above the horizontal line are significantly regulated. Dotted vertical lines correspond to a fold change of 1.5—proteins at the right side of the right vertical line are upregulated, proteins at the left side of the left vertical line are downregulated. Colored data points indicate potential bias proteins (blood coagulation, erythrocyte, and platelet proteins). ICH indicates intracerebral hemorrhage; IS, ischemic stroke; and SM, stroke mimics.

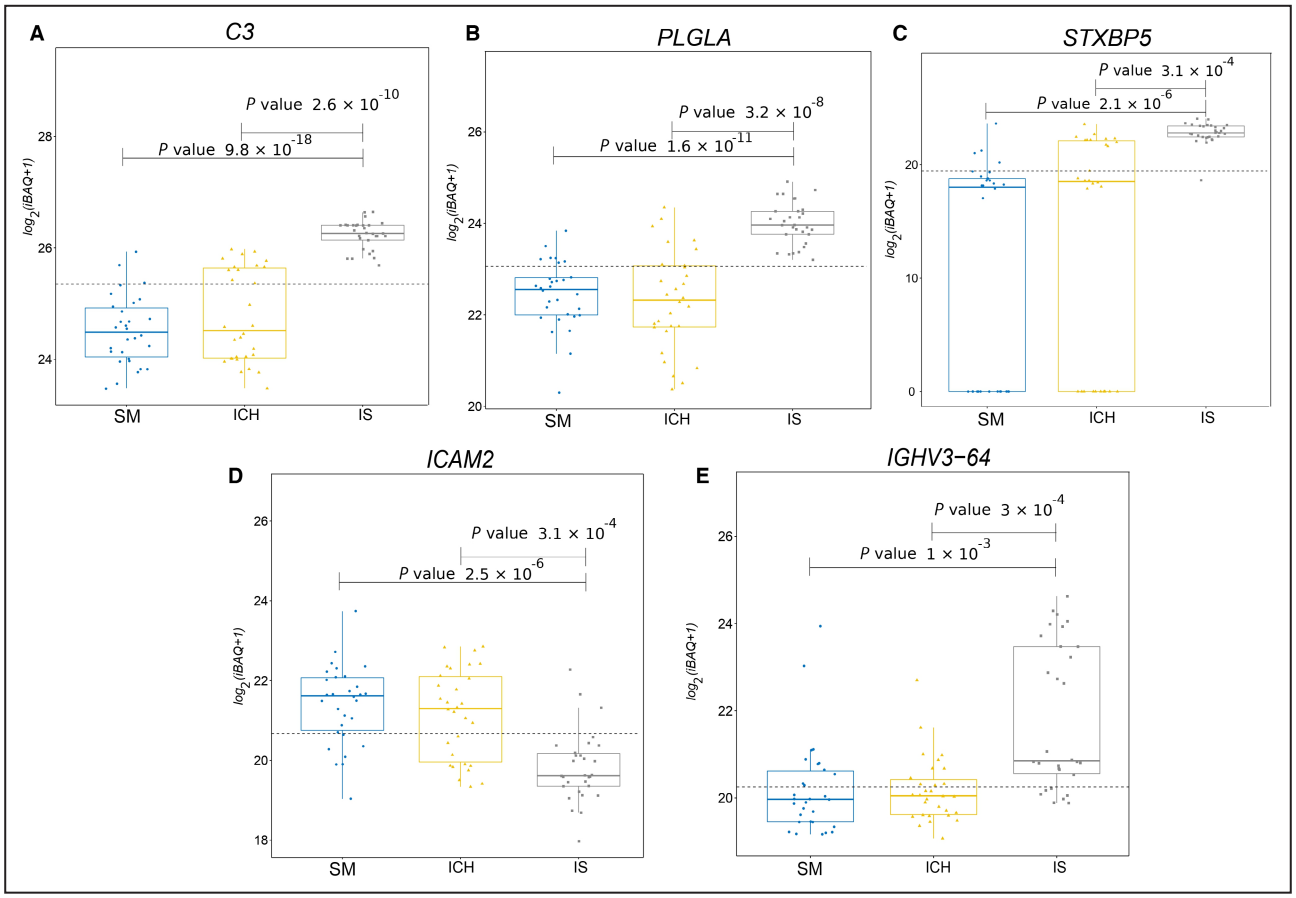

Figure 3. Boxplot of the top 5 most significantly dysregulated proteins obtained for the differentiating of IS from ICH and SM.

A, iBAQ values for C3 in SM, ICH, and IS; **B**, iBAQ values for PLGLA in SM, ICH, and IS; **C**, iBAQ values for STXBP5 in SM, ICH, and IS; **D**, iBAQ values for ICAM-2 in SM, ICH, and IS; **E**, iBAQ values for IGHV3-64 in SM, ICH, and IS. Zero-valued observations were not imputed for the statistical analysis, which explains the large spread of data points on (**C**). Data were transformed using a $\log_2$ function. To avoid missing values introduced by the log transform, a $\log_2(x+1)$ transformation was chosen. Different constants could affect the results of the linear models. To test the effect of the added constant on the linear models in the 3 pairwise comparisons, the comparisons were repeatedly calculated using 8 different constants ranging from 1 to 10^7 using a vector that follows an exponential growth pattern with a common ratio of 10. The 5 most significantly regulated proteins did not change rank when constants from 1 to 1000 were used. Constants in the range of 10^3 to 10^7 changed the ranking of the significantly regulated proteins, but the reported 5 significantly regulated proteins remained significantly regulated. C3 indicates complement component 3; iBAQ, Intensity-based absolute quantification; ICAM2, intercellular adhesion molecule 2; ICH, intracerebral hemorrhage; IGHV3-64, immunoglobulin heavy variable 3-64; IS, ischemic stroke; PLGLA, plasminogen like A; SM, stroke mimics; and STXBP5, syntaxin-binding protein 5.

patients with IS from patients with ICH or SM. Contrary to the models above, full proteome analysis would be more difficult to translate to clinical practice. Nevertheless, the results of these analyses can also suggest additional potential biomarkers for stroke diagnosis that can be further assessed in future studies.

Five proteins were identified as being particularly relevant in differentiating patients with IS from patients with ICH and SM, namely C3, ICAM-2, PLGLA, STXBP5, and IGHV3-64.

C3 is the most abundant protein of the complement system and is implicated in multiple physiological functions and pathophysiological processes. After cerebral ischemia, C3 activation is involved in the mechanism of brain injury, and inhibition of complement activation may reduce the volume of cerebral infarction.[27] Previous studies have found increased C3 levels after IS,[28] although there are conflicting data about the association between C3 levels and IS outcome.[29,30] In ICH, C3 is associated with inflammation and microglia activation, contributing to ICH-induced brain injury.[31] One study has assessed C3 levels in patients with ICH and found no association between C3 levels and ICH rebleeding or functional outcome.[32]

ICAM-2 is a transmembrane glycoprotein expressed on the surface of endothelial cells, platelets, and a variety of leucocyte subsets. It is expressed in the blood–brain barrier endothelium and contributes to both leucocyte diapedesis and extravasation from blood vessels into surrounding tissues.[33] ICAM-2 has

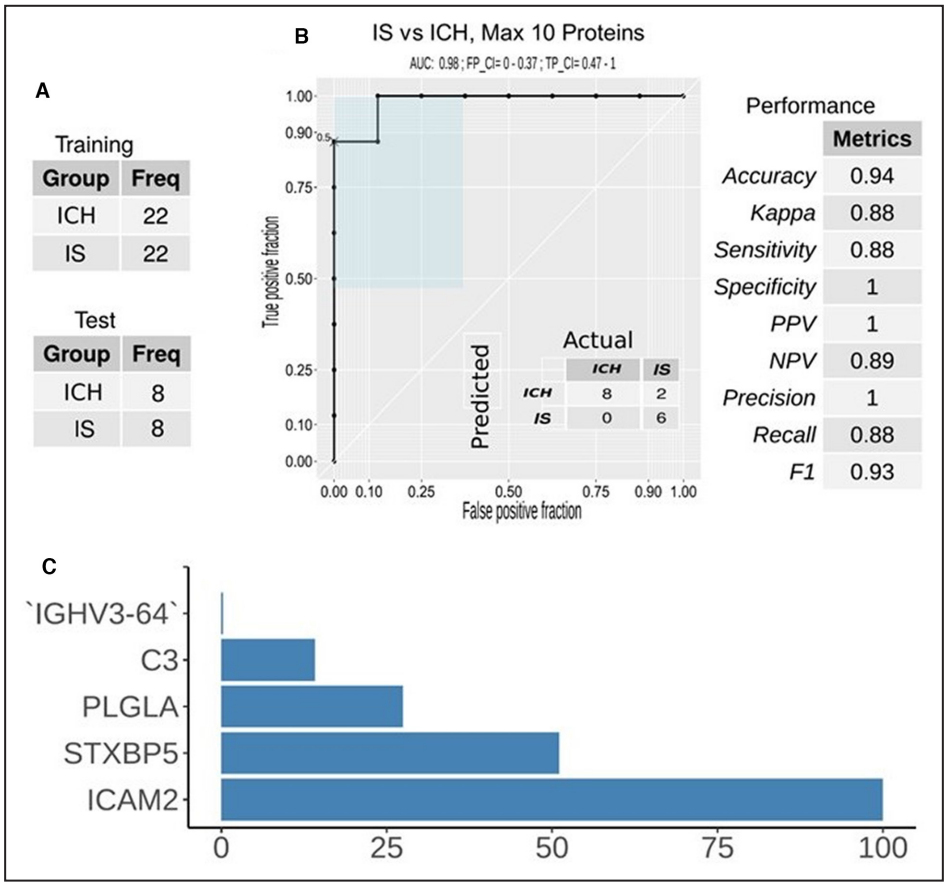

Figure 4. PLS classifier using a PLS model based on a maximum of 10 proteins for differentiating IS from ICH.

A, Tables indicating the number of samples used for training vs test data set. **B**, Receiver operating characteristics and AUC with confusion matrix embedded. Metrics of classification performance on the validation data set are provided on the right. The cross indicates the cut of threshold used to calculate the exact CI. **C**, The lower panel depicts the most important predictor variables for the PLS model. The x-axis indicates the importance score scaled from 0 to 100 provided by the R package "caret." AUC indicates area under the curve; C3, complement component 3; FP, false positive; ICAM-2, intercellular adhesion molecule 2; ICH, intracerebral hemorrhage; IGHV3-64, immunoglobulin heavy variable 3-64; IS, ischemic stroke; NPV, negative predictive value; PLGLA, plasminogen like A; PLS, partial least square; PPV, positive predictive value; STXBP5, syntaxin-binding protein 5; and TP, true positive.

overlapping characteristics and functions with ICAM-1, which, unlike ICAM-2, has been fairly well studied in patients with IS. Previous reports showed increased ICAM-1 levels in patients with IS,[34,35] with higher ICAM-1 levels associated with worse prognosis.[35] Due to the potential role of ICAM-1 in IS pathophysiology, a randomized clinical trial using ICAM-1 antibody was conducted but failed to show beneficial results.[36] In patients with ICH, there are conflicting data regarding the association between ICAM-1 levels and functional outcome.[29,37]

PLGLA functions have been rarely studied. Because nucleotide sequence comparisons showed a high degree of homology with plasminogen[38] it has been hypothesized that PLGLA is involved in the same functions as plasminogen. Besides its role in fibrinolysis, plasminogen also plays a role in synaptic plasticity and increased blood–brain barrier permeability.[39] In proteomic-based studies, higher plasminogen levels were described in patients with IS in comparison with ICH.[11]

STXBP5 is a protein involved in endothelial exocytosis and platelet secretion. In particular, STXBP5 regulates the secretion of von Willebrand factor (VWF), a glycoprotein with an important role in hemostasis, especially by promoting platelet adhesion.[40] STXBP5 mutations were associated with reduced plasma VWF levels and VWF activity,[40,41] while STXBP5 genetic variations may increase the risk of thromboembolic disease.[42] STXBP5 also seems to promote the release of tPA from endothelial cells.[42] To our knowledge, there is no study investigating the role of STXBP5 in patients with stroke or measuring its levels.

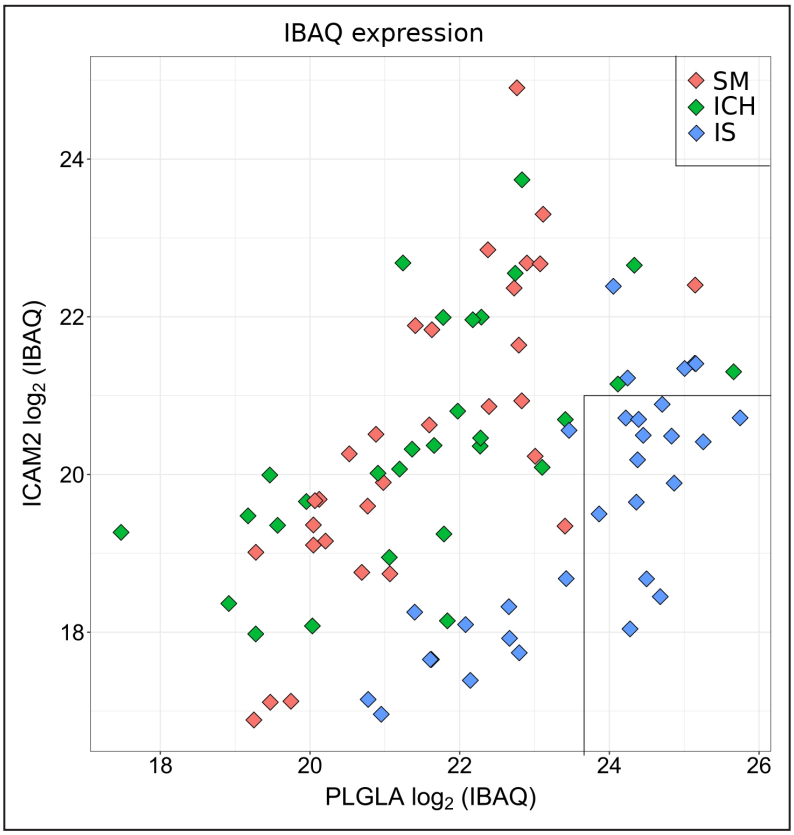

Figure 5. Scatterplot of iBAQ values for ICAM-2 and PLGLA.
Dashed lines indicate area containing exclusively IS samples. iBAQ indicates intensity-based absolute quantification; ICAM-2, intercellular adhesion molecule 2; ICH, intracerebral hemorrhage; IS, ischemic stroke; PLGLA, plasminogen like A; and SM, stroke mimics.

IGHV3-64 corresponds to a particular variable domain in a heavy-chain immunoglobulin, a polypeptide subunit of an antibody. Being part of an antibody structure, IGHV3-64 is involved in different functions such as antigen-binding activity, activation of the immune response, and integrating immunoglobulin complexes. However, its specific role, as well as that of other variable domains in heavy chain immunoglobulin, remains to be elucidated. Inflammation has a significant role in stroke, and different inflammatory proteins have been shown to be associated with stroke risk, diagnosis, cause, and prognosis.[8,11,28–32,34,35,43] In 1 proteomic study, other variable domains of heavy-chain immunoglobulin were found to be significantly increased in patients with stroke in comparison with controls.[11]

Previous studies, with other preselected biomarkers, have shown promising results for the identification of patients with IS.[7–10] In particular, a blood-based biomarker panel using GFAP, RBP4, and NT-proBNP showed a sensitivity rate of 51.5% at 100% specificity for IS diagnosis.[7] Despite its moderate sensitivity rate, this blood-based biomarker panel was studied in rapid point-of-care tests, making it closer to clinical application. Potentially, a combination of previously tested and newly disclosed biomarkers such as those presented in our study could contribute to better performance of blood-based biomarker panels, increasing the likelihood of their use in stroke care.

Only 1 previous study has used label-free proteomics for differentiating patients with IS and ICH, including a total of only 7 patients.[11] Nine candidate blood biomarkers were identified, none of which overlapped with our results or with previous studied biomarkers.[6–10] Time intervals of up to 15 days between symptom onset and blood collection are one of the possible explanations for these differences.

Proteomics has also been explored in the clot characterization of patients with IS, disclosing a different biomarker profile according to stroke cause.[44,45] While being different from our results, the described proteome profiles can work together with our findings by suggesting unexplored pathways implicated in stroke pathophysiology and potentially in stroke treatment.

As previously suggested,[7] if accurate blood-based biomarker panels were available for diagnosing patients with acute IS with feasible rapid point-of-care tests, this

could allow safe prehospital thrombolysis, in select patients, without the need of pre-hospital imaging. Albeit preliminary, our results can add to a potential future biomarker panel, that after validation in large cohort studies and subsequent testing in well-conducted trials, could be implemented with this goal. Our exploratory study revealed a new blood biomarker combination, that, while aiming to minimize false-positives, can correctly identify at least 4 of 5 patients with IS that could be immediately treated, with the remaining patients still having the opportunity to be treated after hospital admission.

In a hospital setting, after imaging confirms the absence of an intracranial hemorrhage, our results could help in guiding the decision to administer intravenous thrombolysis by assisting in the differentiation of IS from SM. This can be especially relevant in patients with risk factors for bleeding, in which the decision to administer thrombolysis becomes particularly challenging.

However, the application of biomarkers in this clinical setting is not devoid of challenges. First, because IS is a heterogeneous disease involving different pathophysiological processes, the goal of having a blood-based biomarker panel that integrates all dimensions of patients with IS is challenging. In addition, a different biomarker profile between patients with suspected stroke may result not from the acute event but from the differences in patients' medical conditions.[6,7] The presence of concomitant systemic medical conditions can also influence the biomarker measurement. Because a kinetic component of blood-based biomarkers cannot be excluded, there is the possibility that some biomarkers will only be useful for diagnostic purposes in a limited time window, as shown with GFAP.[46] The applicability of pre-hospital biomarkers will require trained paramedics and pre-hospital health workers, whose level of expertise is asymmetric among countries.

Although selecting patients for thrombolysis with a blood-based biomarker will be difficult in clinical practice, the applicability of blood-based biomarkers for diagnosing patients with stroke is not limited to this goal. In patients with severe neurological deficits, a blood-based biomarker could be used to better select patients for swift referral to a comprehensive stroke center in a mothership approach, increasing the probability of correctly selecting a candidate for endovascular treatment. As such, blood biomarkers could also help select patients for a direct transfer to an angiography suite approach in this subset of patients, a strategy with growing evidence and promising results.

Our study has limitations. Being an exploratory study designed to disclose new potential biomarkers, the full dimension of clinical heterogeneity as well as our biomarker analysis are limited by our small sample size. Although clinically relevant, the decision to build our models aiming to minimize misidentification of patients with ICH as having IS additionally limits our analysis. Due to the large CIs, we cannot exclude the possibility of misclassifications with our methodology.

Because alternative classifiers based on fewer than 10 markers provided similar but slightly lower performance, further validation should not only be restricted to the top 5 highlighted proteins. The data presented need to be replicated in different and larger cohorts and clinically validated using alternative measurement techniques. Furthermore, in a clinical and practical scenario, serum samples instead of plasma samples would need to be used for such biomarkers' detection. Likewise, MS analysis requires sample processing that could influence our analysis and contribute to different results compared with orthogonal methods. Also, because we cannot exclude that the identified biomarkers will have a particular kinetic profile over time, our results may not be applicable to patients with suspected stroke arriving after 6 hours.

Nonetheless, this is the first study, to our knowledge, to use MS on samples obtained in an acute setting to diagnose patients with suspected stroke. It can also add information on IS-related acute adaptation mechanisms that could potentially be applied to new neuroprotective targets. Our prospective and consecutive enrollment of patients with acutely suspected stroke and the robustness of our training and testing approach are important strengths of our approach.

In conclusion, in this proof-of-concept exploratory study, our proteomics analysis identified potential new biomarkers for diagnosing acute IS in patients with good to excellent sensitivity while aiming to avoid the misclassification of ICH as IS. Although our approach provided an internal validation, our results still need to be confirmed in other cohorts and with different measurement techniques.

ARTICLE INFORMATION

Received March 15, 2023; accepted October 11, 2023.

Affiliations

Department of Neurology, Hospital de Egas Moniz, Centro Hospitalar Lisboa Ocidental (J.P.M., M.S., B.M., M.F., F.S., G.C., R.V., A.S.-P., M.V.-B.) and Centro Clínico Académico de Lisboa (CCAL), NOVA Medical School (MNS), Lisbon, Portugal (J.P.M., M.V.-B.); iNOVA4Health, NOVA Medical School, Universidade NOVA de Lisboa, Lisbon, Portugal (A.S.C., I.G.M., M.M., H.L.V., R.M.); UCIBIO, Applied Molecular Biosciences Unit, NOVA School of Science and Technology (I.G.M., H.L.V.) and i4HB—Institute for Health and Bioeconomy, NOVA School of Science and Technology (I.G.M., H.L.V.), Universidade NOVA de Lisboa, Caparica, Portugal; Champalimaud Research and Clinical Centre, Champalimaud Foundation, Lisbon, Portugal (M.M.); and Department of Clinical Biochemistry, Odense University Hospital, Odense, Denmark (H.C.B.).

Acknowledgments

We thank all members from the Neurology and Nursing Emergency Department team for their support in sample collection. We thank Dr João Faro Viana and the Clinical Pathology Department team for their support in blood samples processing and storage.

Sources of Funding

This work was financed by Fundação para a Ciência e a Tecnologia (FCT) with PTDC/MEC-NEU/28750/2017 grant, LA/P/0140/2020 funding,

iNOVA4Health—UIDB/04462/2020, and UCIBIO-UID/Multi/04378/2019 units. A.S.C. is supported by FCT. R.M. is also supported by FCT (CEEC position, 2019–2025 investigator).

Disclosures

Marto JP declares consulting and speaker fees from Amicus Therapeutics and Boehringer Ingelheim; Educational/ travel fees from Boehringer Ingelheim, Bayer and Pfizer - Bristol Myers Squibb. Mendonça M has received honoraria from serving on a scientific advisory board for Bial, consulting fees from Pharmacademy and Evidenze and has received fees for speaking at conferences from AbbVie.

Supplemental Material

Figure S1–S10
Table S1

REFERENCES

1. GBD 2019 Stroke Collaborators. Global, regional, and national burden of stroke and its risk factors, 1990–2019: a systematic analysis for the global burden of disease study 2019. *Lancet Neurol*. 2021;20:795–820.
2. 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
3. Saver JL, Goyal M, van der Lugt A, Menon BK, Majoie CB, Dippel DW, Campbell BC, Nogueira RG, Demchuk AM, Tomasello A, et al. Time to treatment with endovascular thrombectomy and outcomes from ischemic stroke: a meta-analysis. *JAMA*. 2016;316:1279–1288. doi: 10.1001/jama.2016.13647
4. Li Q, Warren AD, Qureshi AI, Morotti A, Falcone GJ, Sheth KN, Shoamanesh A, Dowlatshahi D, Viswanathan A, Goldstein JN. Ultra-early blood pressure reduction attenuates hematoma growth and improves outcome in intracerebral hemorrhage. *Ann Neurol*. 2020;88:388–395. doi: 10.1002/ana.25793
5. Turc G, Hadziahmetovic M, Walter S, Churilov L, Larsen K, Grotta JC, Yamal JM, Bowry R, Katsanos AH, Zhao H, et al. Comparison of mobile stroke unit with usual care for acute ischemic stroke management: a systematic review and meta-analysis. *JAMA Neurol*. 2022;79:281–290. doi: 10.1001/jamaneurol.2021.5321
6. Perry LA, Lucarelli T, Penny-Dimri JC, McInnes MD, Mondello S, Bustamante A, Montaner J, Foerch C, Kwan P, Davis S, et al. Glial fibrillary acidic protein for the early diagnosis of intracerebral hemorrhage: systematic review and meta-analysis of diagnostic test accuracy. *Int J Stroke*. 2019;14:390–399. doi: 10.1177/1747493018806167
7. Bustamante A, Penalba A, Orset C, Azurmendi L, Llombart V, Simats A, Pecharroman E, Ventura O, Ribó M, Vivien D, et al. Blood biomarkers to differentiate ischemic and hemorrhagic strokes. *Neurology*. 2021;96:e1928–e1939. doi: 10.1212/WNL.0000000000011742
8. Bustamante A, López-Cancio E, Pich S, Penalba A, Giralt D, García-Berrocoso T, Ferrer-Costa C, Gasull T, Hernández-Pérez 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
9. Llombart V, García-Berrocoso T, Bustamante A, Giralt D, Rodriguez-Luna D, Muchada M, Penalba A, Boada C, Hernández-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
10. Montaner J, Mendioroz M, Ribó M, Delgado P, Quintana M, Penalba A, Chacón P, Molina C, Fernández-Cadenas I, Rosell A, et al. A panel of biomarkers including caspase-3 and D-dimer may differentiate acute stroke from stroke-mimicking conditions in the emergency department. *J Intern Med*. 2011;270:166–174. doi: 10.1111/j.1365-2796.2010.02329.x
11. Malicek D, Wittig I, Luger S, Foerch C. Proteomics-based approach to identify novel blood biomarker candidates for differentiating intracerebral hemorrhage from ischemic stroke-a pilot study. *Front Neurol*. 2021;12:713124. doi: 10.3389/fneur.2021.713124
12. Hatano S. Experience from a multicentre stroke register: a preliminary report. *Bull World Health Organ*. 1976;54:541–553.
13. Goyal N, Tsivgoulis G, Male S, Metter EJ, Iftikhar S, Kerro A, Chang JJ, Frey JL, Triantafyllou S, Papadimitropoulos G, et al. FABS: an intuitive tool for screening of stroke mimics in the emergency department. *Stroke*. 2016;47:2216–2220. doi: 10.1161/STROKEAHA.116.013842
14. Runchey S, McGee S. Does this patient have a hemorrhagic stroke?: clinical findings distinguishing hemorrhagic stroke from ischemic stroke. *JAMA*. 2010;303:2280–2286. doi: 10.1001/jama.2010.754
15. Geisler F, Ali SF, Ebinger M, Kunz A, Rozanski M, Waldschmidt C, Weber JE, Wendt M, Winter B, Schwamm LH, et al. Evaluation of a score for the prehospital distinction between cerebrovascular disease and stroke mimic patients. *Int J Stroke*. 2019;14:400–408. doi: 10.1177/1747493018806194
16. Wiśniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. *Nat Methods*. 2009;6:359–362. doi: 10.1038/nmeth.1322
17. Carvalho AS, Ribeiro H, Voabil P, Penque D, Jensen ON, Molina H, Matthiesen R. Global mass spectrometry and transcriptomics array-based drug profiling provides novel insight into glucosamine induced endoplasmic reticulum stress. *Mol Cell Proteomics*. 2014;13:3294–3307. doi: 10.1074/mcp.M113.034363
18. Matthiesen R, Prieto G, Amorim A, Aloria K, Fullaondo A, Carvalho AS, Arizmendi JM. SIR: deterministic protein inference from peptides assigned to MS data. *J Proteomics*. 2012;75:4176–4183. doi: 10.1016/j.jprot.2012.05.010
19. Geyer PE, Voytik E, Treit PV, Doll S, Kleinhempel A, Niu L, Müller JB, Buchholtz ML, Bader JM, Teupser D, et al. Plasma proteome profiling to detect and avoid sample-related biases in biomarker studies. *EMBO Mol Med*. 2019;11:e10427. doi: 10.15252/emmm.201910427
20. Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M. Global quantification of mammalian gene expression control. *Nature*. 2011;473:337–342. doi: 10.1038/nature10098
21. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*. 2003;19:185–193. doi: 10.1093/bioinformatics/19.2.185
22. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc Series B*. 1995;57:289–300.
23. Vinzi VE, Chin WW, Henseler J, Wang H. *Handbook of Partial Least Squares: Concepts, Methods and Applications*. Springer; 2010. doi: 10.1007/978-3-540-32827-8
24. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol*. 2004;3:1–25. doi: 10.2202/1544-6115.1027
25. Sachs MC. plotROC: a tool for plotting ROC curves. *J Stat Softw*. 2017;79:2. doi: 10.18637/jss.v079.c02
26. Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*. 1934;26:404–413. doi: 10.1093/biomet/26.4.404
27. Ma Y, Liu Y, Zhang Z, Yang GY. Significance of complement system in ischemic stroke: a comprehensive review. *Aging Dis*. 2019;10:429–462. doi: 10.14336/AD.2019.0119
28. Mocco J, Wilson DA, Komotar RJ, Sughrue ME, Coates K, Sacco RL, Elkind MS, Connolly ES Jr. Alterations in plasma complement levels after human ischemic stroke. *Neurosurgery*. 2006;59:28–33. doi: 10.1227/01.neu.0000243280.75920.f4
29. Zhong C, Zhu Z, Wang A, Xu T, Bu X, Peng H, Yang J, Han L, Chen J, Xu T, et al. Multiple biomarkers covering distinct pathways for predicting outcomes after ischemic stroke. *Neurology*. 2019;92:e295–e304. doi: 10.1212/WNL.0000000000006717
30. Yang P, Zhu Z, Zang Y, Bu X, Xu T, Zhong C, Wang A, Peng H, Guo D, Zheng X, et al. Increased serum complement C3 levels are associated with adverse clinical outcomes after ischemic stroke. *Stroke*. 2021;52:868–877. doi: 10.1161/STROKEAHA.120.031715
31. Wang M, Xia F, Wan S, Hua Y, Keep RF, Xi G. Role of complement component 3 in early erythrolysis in the hematoma after experimental intracerebral hemorrhage. *Stroke*. 2021;52:2649–2660. doi: 10.1161/STROKEAHA.121.034372
32. Wang KW, Cho CL, Chen HJ, Liang CL, Liliang PC, Tsai YD, Wang HK, Lu K. Molecular biomarker of inflammatory response is associated with rebleeding in spontaneous intracerebral hemorrhage. *Eur Neurol*. 2011;66:322–327. doi: 10.1159/000332027
33. Gorina R, Lyck R, Vestweber D, Engelhardt B. β2 integrin-mediated crawling on endothelial ICAM-1 and ICAM-2 is a prerequisite for

transcellular neutrophil diapedesis across the inflamed blood-brain barrier. *J Immunol*. 2014;192:324–337. doi: 10.4049/jimmunol.1300858
34. Hassan A, Hunt BJ, O'Sullivan M, Parmar K, Bamford JM, Briley D, Brown MM, Thomas DJ, Markus HS. Markers of endothelial dysfunction in lacunar infarction and ischaemic leukoaraiosis. *Brain*. 2003;126:424–432. doi: 10.1093/brain/awg040
35. Castellanos M, Castillo J, García MM, Leira R, Serena J, Chamorro A, Dávalos A. Inflammation-mediated damage in progressing lacunar infarctions: a potential therapeutic target. *Stroke*. 2002;33:982–987. doi: 10.1161/hs0402.105339
36. Enlimomab Acute Stroke Trial Investigators. Use of anti-ICAM-1 therapy in ischemic stroke: results of the enlimomab acute stroke trial. *Neurology*. 2001;57:1428–1434. doi: 10.1212/WNL.57.8.1428
37. Castillo J, Dávalos A, Alvarez-Sabín J, Pumar JM, Leira R, Silva Y, Montaner J, Kase CS. Molecular signatures of brain injury after intracerebral hemorrhage. *Neurology*. 2002;58:624–629. doi: 10.1212/WNL.58.4.624
38. Lewis VO, Gehrmann M, Weissbach L, Hyman JE, Rielly A, Jones DG, Llinás M, Schaller J. Homologous plasminogen N-terminal and plasminogen-related gene a and B peptides. Characterization of cDNAs and recombinant fusion proteins. *Eur J Biochem*. 1999;259:618–625. doi: 10.1046/j.1432-1327.1999.00055.x
39. Keragala CB, Medcalf RL. Plasminogen: an enigmatic zymogen. *Blood*. 2021;137:2881–2889. doi: 10.1182/blood.2020008951
40. Lind-Halldén C, Manderstedt E, Carlberg D, Lethagen S, Halldén C. Genetic variation in the syntaxin-binding protein STXBP5 in type 1 von Willebrand disease patients. *Thromb Haemost*. 2018;118:1382–1389. doi: 10.1055/s-0038-1661352
41. van Loon JE, Leebeek FW, Deckers JW, Dippel DW, Poldermans D, Strachan DP, Tang W, O'Donnell CJ, Smith NL, de Maat MP. Effect of genetic variations in syntaxin-binding protein-5 and syntaxin-2 on von Willebrand factor concentration and cardiovascular risk. *Circ Cardiovasc Genet*. 2010;3:507–512. doi: 10.1161/CIRCGENETICS.110.957407
42. Huang J, Huffman JE, Yamakuchi M, Trompet S, Asselbergs FW, Sabater-Lleal M, Trégouët DA, Chen WM, Smith NL, Kleber ME, et al. Genome-wide association study for circulating tissue plasminogen activator levels and functional follow-up implicates endothelial STXBP5 and STX2. *Arterioscler Thromb Vasc Biol*. 2014;34:1093–1101. doi: 10.1161/ATVBAHA.113.302088
43. Esenwa CC, Elkind MS. Inflammatory risk factors, biomarkers and associated therapy in ischaemic stroke. *Nat Rev Neurol*. 2016;12:594–604. doi: 10.1038/nrneurol.2016.125
44. Dargazanli C, Zub E, Deverdun J, Decourcelle M, de Bock F, Labreuche J, Lefèvre PH, Gascou G, Derraz I, Riquelme Bareiro C, et al. Machine learning analysis of the cerebrovascular thrombi proteome in human ischemic stroke: an exploratory study. *Front Neurol*. 2020;11:575376. doi: 10.3389/fneur.2020.575376
45. Rossi R, Mereuta OM, Barbachan E, Silva M, Molina Gil S, Douglas A, Pandit A, Gilvarry M, McCarthy R, O'Connell S, et al. Potential biomarkers of acute ischemic stroke etiology revealed by mass spectrometry-based proteomic characterization of formalin-fixed paraffin-embedded blood clots. *Front Neurol*. 2022;13:854846. doi: 10.3389/fneur.2022.854846
46. Dvorak F, Haberer I, Sitzer M, Foerch C. Characterisation of the diagnostic window of serum glial fibrillary acidic protein for the differentiation of intracerebral haemorrhage and ischaemic stroke. *Cerebrovasc Dis*. 2009;27:37–41. doi: 10.1159/000172632