Discovery and Longitudinal Evaluation of Candidate Biomarkers for Ischaemic Stroke by Mass Spectrometry-Based Proteomics

Marie Dagonnier[1], Ira Robin Cooke[2,3], Pierre Faou[2], Tara Kate Sidon[1], Helen Margaret Dewey[1], Geoffrey Alan Donnan[1] and David William Howells[1,4]

[1]Stroke Department, The Florey Institute of Neuroscience & Mental Health and Melbourne Brain Centre, Melbourne, VIC, Australia. [2]La Trobe Institute for Molecular Science, La Trobe University, Melbourne, VIC, Australia. [3]Life Sciences Computation Centre, Victorian Life Sciences Computation Initiative, La Trobe University, Melbourne, VIC, Australia. [4]School of Medicine, Faculty of Health, University of Tasmania, Hobart, TAS, Australia.

ABSTRACT: Application of acute therapies such as thrombolysis for ischaemic stroke (IS) is constrained because of diagnostic uncertainty and the dynamic nature of stroke biology. To investigate changes in blood proteins after stroke and as a result of thrombolysis treatment we performed label-free quantitative proteomics on serum samples using high-resolution mass spectrometry and long high-performance liquid chromatography gradient (5 hours) combined with a 50-cm column to optimise the peptide separation. We identified (false discovery rate [FDR]: 1%) and quantified a total of 574 protein groups from a total of 92 samples from 30 patients. Ten patients were treated by thrombolysis as part of a randomised placebo-controlled trial and up to 5 samples were collected from each individual at different time points after stroke. We identified 26 proteins differently expressed by treatment group (FDR: 5%) and significant changes of expression over time for 23 proteins (FDR: 10%). Molecules such as fibrinogen and C-reactive protein showed expression profiles with a high-potential clinical utility in the acute stroke setting. Protein expression profiles vary acutely in the blood after stroke and have the potential to allow the construction of a stroke clock and to have an impact on IS treatment decision making.

KEYWORDS: Stroke, biomarker, protein expression, mass spectrometry

Introduction

Stroke, of which ischaemic stroke (IS) is the most common form, is a leading cause of death and disability.[1] Thrombolysis with recombinant tissue plasminogen activator (rt-PA) given within the first 4.5 hours of onset is the most specific and biologically powerful treatment for acute IS.[2,3] Unfortunately, thrombolysis is only received by only a small proportion of patients with stroke. This is mainly because of uncertainty about diagnosis, time of onset, and the perceived risks of cerebral bleeding these entail.[4–8] Even in specialised centres, these limits mean that fewer than 15% of patients with ischaemic stroke are treated with thrombolysis.[7,9–11]

Neuroimaging biomarkers are widely used to aid in the diagnosis of stroke, to characterise its type, to assist in determining a likely cause, to estimate the severity, and to predict functional outcome and risk of recurrence. Computed tomography (CT) detection of bleeds is critical for the use of rt-PA,[2,12] and penumbral imaging by CT perfusion (CTp) and magnetic resonance imaging (MRI) are increasingly used to select patients with an ischaemic penumbra likely to respond well to thrombolysis.[13,14] Perfusion-weighted imaging (PWI) in combination with diffusion-weighted imaging (DWI) is thought to identify this ischaemic but potentially salvageable tissue in acute ischaemic stroke.[15,16] However, CTp and MRI do not provide all the information needed for patient care and their use is limited by availability, patient contraindications, and cost.

Early identification of salvageable brain tissue without the resource constraints of imaging would allow patients to be treated earlier (for example, in the ambulance) and improve stroke outcome. Likewise, identification of patients at risk of bleeding after thrombolysis treatment and identification of patients who would respond well to rt-tissue plasminogen activator (tPA) therapy would further improve patient outcomes. Biomarkers of disease evolution detected in blood could provide these improvements. Biomarkers acting as a stroke clock would help to determine stroke onset time and be particularly valuable for patients who wake up with stroke (which represents up to 25% of all IS) and who are currently ineligible for rt-PA therapy.[6,17,18]

To date, biochemical biomarker studies have focused on confirmation of stroke diagnosis and particularly on differentiation between ischaemic and haemorrhagic strokes[19–24] or on using biomarkers as tools for outcome prediction.[25–29] Although many candidate blood-based biomarkers for IS have been hypothesised, none are currently used in clinical practice.[30–32]

The use of mass spectrometry (MS) in the quest to identify biomarkers, particularly proteins or peptides, for clinical use has increased dramatically during the past decades.[33] Importantly, once identified, immunologically based assay systems (such as enzyme-linked immunosorbent assays [ELISAs]) offer the rapid detection needed for practical decision making.

A number of plasma or serum protein biomarkers have already been successfully introduced into clinical practice. These include the myocardium-specific structural protein troponins, B-type natriuretic peptide, and D-dimers for the exclusion or diagnosis of myocardial infarction, congestive heart failure, or pulmonary embolism, respectively.[34–36]

Proteomics studies have been performed specifically regarding study IS using a range of different body tissues including cerebrospinal fluid, microdialysates, brain tissue, or urine but more rarely in blood.[29,37–42]

Biomarkers that are detectable in blood are ideal due to the ease of collection and the fact that blood composition changes rapidly in response to disease.[43] However, proteomic analysis on blood-derived products such as serum is extremely challenging due to the high dynamic range of proteins present. Despite these challenges, the apolipoproteins, ApoC-I and ApoC-III, detected by MS in plasma have been reported as biomarkers able to differentiate IS from haemorrhagic strokes.[44,45] In addition, Zhang et al reported a panel of 13 markers for diagnosis of stroke identified by comparing plasma from patients with acute cerebral infarction and control subjects using surface-enhanced laser desorption/ionisation time of flight (SELDI-TOF) MS. This panel showed a sensitivity of 84.4% and a specificity of 95% to discriminate the 2 groups.[46] In a similar experiment, Huang et al[47] identified serum-free haemoglobin as a potential biomarker for the diagnosis of IS (sensitivity of 70.2% and specificity of 85.3%) and MS was also used to identify PBP (platelet basic protein) as a candidate serum biomarker for transient ischaemic attack.[48]

Here, we report the MS analysis of blood samples sequentially collected from a cohort of patients with stroke with imaging and functional follow-up and after rt-PA treatment. In this proof-of-principle study, we demonstrate that advanced proteomic technology can be used in the quest of stroke biomarkers in a clinical setting.

Methods

Sample collection and selection

Samples were obtained from patients with stroke who were recruited in the EXtending the time for Thrombolysis in Emergency Neurological Deficits (EXTEND) trial.

The EXTEND trial is an ongoing randomised, multicentre, double-blinded, placebo-controlled phase 3 study combining clinical and imaging criteria to select a population more likely to benefit from the advantages of tPA when given after the traditional 4.5-hour time window (in comparison with patients treated with standard stroke care). Trial protocol, inclusion, and exclusion criteria were previously published.[49] Patients with a known stroke onset time and randomised in the EXTEND treatment arms were selected for study.

For each patient, samples were planned to be collected immediately after imaging (baseline), immediately before treatment administration (T0), 12 to 24 hours after treatment administration (T12-24h) and at 3 days (T3d), 90 days (T90d), and 12 months (T12m) after stroke.

For biomarker analysis, 5 mL of venous blood were collected for each time point. The serum was separated, centrifuged at 1500*g* for 10 minutes and stored at –80°C. Samples were processed and frozen within 1 hour from the initial collection time.

Clinical data including age, sex, stroke type according to the trial of ORG 10172 in acute stroke treatment (TOAST) classification, presence of parenchymal haemorrhage (PH), and detailed past medical history were recorded. Imaging (by CTp or MRI) was performed on admission and PWI and DWI volumes were measured. Stroke severity was assessed using the National Institutes of Health Stroke Scale (NIHSS) score. Functional outcome was evaluated at 3 months after the event by means of the modified Rankin scale (mRS).

Written informed consent was obtained from all study participants or relatives. The study conformed to the principles of the Declaration of Helsinki and was approved by the ethics committees of the different involved institutions.

Sample preparation

In-solution trypsin digest for proteomics. About 10 µL of serum from each patient was first stabilised in 100 µL of 8 M urea, pH = 8.3, and stored at –80°C until used. For proteomic analysis, the stabilised samples were randomised by batches of 16 and processed as follows: 20 µL protein solution was added to 90 µL of 8 M urea, pH = 8.3, and reduced for 5 hours with 1 µL of 200 mM tris(2-carboxyethyl)phosphine. After this, samples were alkylated for 1 hour at 25°C in the dark with 4 µL of 1 M iodoacetamide. Digests were performed overnight (37°C) by addition of 1 µg of trypsin (Promega, Madison, WI, USA) and 900 µL of 50 mM Tris, pH = 8.3, followed by a second digestion step with 1 µg trypsin and an additional incubation of 4 hours at 37°C.

Solid-phase extraction cleanup of tryptic peptides. About 200 µL of the digested solution was collected and dried by SpeedVac centrifugation. The digested proteins were resuspended in 100 µL of 1% (v/v) formic acid and centrifuged at 14 000 rpm for 2 minutes. The solid-phase extraction was performed with Empore reversed-phase extraction disks (SDB-XC reversed-phase material, 3M, St. Paul, USA) according to Ishihama et al[50] with the following modifications: the membrane was conditioned with 50 µL of 80% (v/v) acetonitrile, 0.1% (w/v) trifluoroacetic acid, and then washed with 50 µL of 0.1% trifluoroacetic acid before the tryptic peptides were bound to the

membrane. The bound peptides were eluted by 50 μL 80% (v/v) acetonitrile, 0.1% (w/v) trifluoroacetic acid, and dried in a SpeedVac centrifuge.

Electrospray ionisation and liquid chromatography tandem MS conditions

Peptides (2 μL) reconstituted in a final volume of 10 μL 0.1% trifluoroacetic acid and 2 % acetonitrile (buffer A) were loaded and washed onto a trap column (C18 PepMap 100-μm i.d. × 2 cm trapping column; Thermo Fisher Scientific, Waltham, MA, USA) at 5 μL/min for 6 minutes before switching the precolumn in line with the analytical column (Easy-Spray 75-μm i.d. × 50 cm; Thermo Fisher Scientific). The separation of peptides was performed at 250 nL/min using a linear acetonitrile gradient of 2 buffers: buffer A (0.1% formic acid, 2% acetonitrile) and buffer B (0.1% formic acid, 80% acetonitrile), starting from 5% buffer B increasing to 60% over 300 minutes. Data were collected on an Orbitrap Elite (Thermo Fisher Scientific) in data-dependent acquisition mode using mass to charge ratio (m/z) 300 to 1500 as MS scan range. Collision-induced dissociation (CID) tandem MS (MS/MS) spectra were collected for the 20 most intense ions. The dynamic exclusion parameters used were repeat count 1, duration 90 seconds, and the exclusion list size was set at 500 with early expiration disabled. Other instrument parameters for the Orbitrap were as follows: MS scan at 120 000 resolution, maximum injection time 150 ms, automatic gain control (AGC) target 1×10^6, and CID at 35% energy for a maximum injection time of 150 ms with AGC target of 5000. The Orbitrap Elite was operated in dual analyser mode with the Orbitrap analyser being used for MS and the linear trap being used for MS/MS.

Protein identification and label-free quantitation

Identification and label-free quantitation of proteins across all 92 samples were performed using MaxQuant version 1.5.0.30.[51] All raw data and complete details of MaxQuant parameters and result files have been deposited in ProteomeXchange and are available with accession number PXD003218. Identification of peptides and proteins was performed internally by MaxQuant using the Andromeda[51] search engine to search against all reviewed and unreviewed human proteins in the Uniprot database (July 2014). Common contaminants and decoys were included automatically by Andromeda. Prior to searching, MS/MS spectra were filtered by retaining only the top 8 peaks per 100 Da. Precursor mass tolerance was set to 5 ppm (after precursor refinement) and MS/MS tolerance to 0.5 Da. Carbamidomethylation of cysteines was set as a fixed modification, and N-term acetylation and oxidation of methionine were included as variable modifications. Up to 2 missed cleavages were allowed, and peptides were required to be at least 7 amino acids in length. False discovery rate (FDR) cut-offs for both peptides and proteins in the database search were set to 1%. Both unique and razor peptides (peptides shared by different proteins of a group) were used for quantitation with a minimum of 2 peptides including at least 1 unique peptide required to calculate a protein quantitative value. The 'match between runs' setting in MaxQuant was used to transfer peptide identifications from one run to another on the basis of matching retention time and mass to charge ratio.

Statistical analysis

Protein group intensity values calculated by MaxQuant (in proteinGroups.txt) were used for all further statistical analysis in R.[52] Data were analysed using different statistical models to investigate whether relative protein abundance (1) changed with time after stroke onset, (2) differed between patients with different extents of stroke based on imaging results, (3) responded differently to treatment, and (4) differed between patients with different outcomes after stroke. Regardless of the statistical model used, the limma package was always used to fit the model and perform the following processing tasks.[53] Intensity values were always normalised using cyclic loess normalisation,[54] and the method of Ritchie et al[55] was used to calculate sample reproducibility and down-weight less reproducible samples. After initial model fitting, empirical Bayes was used to calculate moderated test statistics,[56] and Benjamini-Hochberg correction was applied to adjust P values for multiple testing. Missing values were simply excluded from the analysis with degrees of freedom adjusted accordingly.

For the time course analysis: 19, 15, 20, 19, and 10 patient samples were available at time points baseline, T0 (prior to treatment), 12 to 24 hours, 3, and 90 days after treatment, respectively. Although the T0 time point was always after baseline for any particular patient, the 2 time points were overlapped in terms of time since stroke onset. Patient was included as a blocking factor as the same patients were followed for the course of the study. Time was included as a fixed effect, and an adjusted P value was calculated for each protein based on a moderated F test calculated across all time points.

To determine the effect of treatment, we specifically focused on samples immediately prior to and after treatment for each patient. Patient was again used as a blocking factor and both an interaction between time (before or after) and treatment (A or B) was included as a fixed effect.

Results

In January 2014, 25 patients with a known stroke onset time had been randomised in Australia for the EXTEND trials. From these pre-selected 25 patients, 21 had a significant number of blood samples available and were therefore used for this discovery project. Regarding the low number of samples available of this time point (due to patient death, loss of follow-up,

Table 1. Patients demographic characteristics for the 21 patients recruited in the EXTEND randomised arm.

PATIENT	AGE	GENDER	BASELINE NIHSS	12-24 H NIHSS	DAY 3-7 NIHSS	DAY 90 NIHSS	PRESTROKE MRS	DAY 90 MRS	DEATH <3 MONTHS	TOAST	IMAGING TYPE	DWI ML	PWI ML	PH	HTA	AF	LIPID DISORDER	PREVIOUS STROKE	ISCHAEMIC HEART DISEASE	DIABETES	PERIPH. VASC. DISEASE
1	70	M	15	2	0	1	0	2	No	CE	CTp	2	133	No	Yes	Yes	Yes	No	Yes	No	Yes
2	49	M	4	0	0	0	0	0	No	SVO	CTp	0	55	No	Yes	No	Yes	No	No	No	No
3	82	M	8	0	0	0	1	1	No	und/inc	CTp	2	54	No	Yes	No	No	No	No	No	Yes
4	81	M	8	3	2	—	0	5	No	LAA	CTp	0	78	No	Yes	No	Yes	Yes	Yes	No	Yes
5	62	M	18	16	16	10	0	4	No	CE	CTp	4	364	No	Yes	No	Yes	No	No	No	No
6	91	M	6	6	9	14	0	4	No	LAA	CTp	0	27	No	Yes	Yes	No	No	Yes	No	No
7	82	F	16	20	18	17	0	5	No	LAA	CTp	0	73	No	Yes	No	No	No	No	No	No
8	82	F	15	15	14	—	0	6	Yes	CE	CTp	6	71	No	Yes	No	No	No	No	No	No
9	56	F	19	2	1	1	0	1	No	LAA	CTp	25	51	No	No	Yes	No	No	No	No	No
10	76	F	5	4	3	0	0	0	No	LAA	CTp	0	60	No	No	Yes	No	No	No	Yes	No
11	79	M	10	7	9	3	0	3	No	LAA	CTp	3	41	No	Yes	No	Yes	No	Yes	No	No
12	75	F	18	10	11	—	0	6	Yes	CE	CTp	0	129	No	Yes	Yes	No	No	Yes	Yes	No
13	78	M	14	3	3	0	0	0	No	CE	CTp	1	106	No	Yes	Yes	Yes	No	Yes	Yes	No
14	66	M	15	17	18	—	0	5	No	CE	CTp	24	127	Yes	Yes	No	No	No	No	Yes	No
15	58	F	14	14	13	5	0	4	No	CE	CTp	8	132	No	Yes	Yes	Yes	No	No	No	No
16	85	F	19	18	18	16	1	5	No	CE	MRI	13	56	Yes	Yes	Yes	Yes	No	No	Yes	No
17	87	F	11	6	8	2	0	3	No	CE	MRI	0	15	Yes	Yes	Yes	No	No	No	No	No
18	59	M	5	5	—	3	0	1	No	CE	MRI	15	31	No	No	Yes	No	No	No	No	No
19	85	F	12	15	15	—	0	6	Yes	LAA	MRI	68	124	No	Yes	Yes	Yes	No	No	No	No
20	86	F	21	19	17	—	0	6	Yes	CE	MRI	38	133	Yes	Yes	No	Yes	No	No	Yes	No
21	80	F	16	29	17	14	1	5	No	CE	MRI	44	152	Yes	Yes	Yes	Yes	No	No	No	No

Abbreviations: AF, atrial fibrillation; CE, cardioembolic; CTp, perfusion computed tomography; DWI, diffusion-weighted imaging; F, female; HTA, arterial hypertension; LAA, large artery atherosclerosis; M, male; MRI, magnetic resonance imaging; mRS, modified ranking scale; NIHSS, National Institutes of Health Stroke Scale; PH, parenchymal haemorrhage; Periph. vasc. disease, peripheral vascular disease; PWI, perfusion-weighted imaging; SVO, small vessel occlusion; TOAST, trial of ORG 10172 in acute stroke treatment; und/inc, undetermined/unclassified.
Missing data are represented by —.

Table 2. Numbers of samples from the 21 patients recruited in the randomised EXTEND arm analysed, at each time point.

	BASELINE	T0	T12-24H	T3D	T90D
Number of samples analysed	19	15	20	19	10

or collection time still to happen), the analysis of 12-month sampling was not included in this project. Patients' characteristics are reported in Table 1.

A total of 83 serum samples from 21 stroke patients were analysed in triplicate for the estimation of the relative abundance of proteins (numbers of samples analysed at each time point are reported in Table 2).

An additional 9 patients (Table 3) were selected from the EXTEND cohort (patients not meeting the inclusion criteria for the randomised trial arms) to enrich the imaging profiles (with extreme values of DWI and PWI volumes). For these 9 additional patients, only the baseline samples were collected and processed.

A total of 614 different protein groups were identified and quantified across our 92 samples at an FDR of 1%. This included 33 common contaminants and 6 decoys (automatically added by the Andromeda search engine), which have been excluded from all further analyses. An overview of the quantitative data is presented as a multidimensional scaling plot in Figure 1 and clearly shows that samples from the same patient tend to be more similar than between patients.

Time

Proteomic profiles over time were investigated using the 83 serum samples collected from the 21 patients randomised into EXTEND. Changes of protein expressions over time were evaluated using Degust, an interactive Web tool for visualising differential gene expression data (http://victorian-bioinformatics-consortium.github.io/degust/). Filtering with an FDR of 10% (FDR ≤ 0.1), 23 molecules were identified as significantly varying over time. A heatmap was generated, showing regulation of these 23 protein expressions over time (Figure 2) and highlighting the strong biphasic nature of some of these protein profiles.

Interesting patterns of proteins expression over time were highlighted in Figure 2 regarding the timing of their expression changes and their potential clinical utility. For example, FIBG and V9HVY1 showed an early significant upregulation (peaking at T0) then gradually decreasing and reaching a level of expression of –2 log of its baseline concentration at 3 months after stroke. The CRP, SAA1, and D3DQX7 presented an expression level gradually increasing after stroke and reaching a peak at 3 days after stroke before drastically decreasing. In contrast, the expression of APOA4 and E3UN46 decreased during the first 3 days after stroke before returning to their baseline value at 3 months after stroke (Figure 3).

The 23 proteins significantly varying over time were loaded into the Search Tool for the Retrieval of Interacting Genes/Proteins 10 (STRING10; http://string-db.org/). The STRING builds functional protein association networks based on compiled available experimental evidence.[57] From the initial 23 proteins varying significantly over time, only 10 were successfully matched to proteins within the STRING10 database. The 13 unmatched proteins represent those for which insufficient interaction data exist. An interaction network for the 10 proteins within the STRING10 database was highly connected with evidence available to connect all proteins with the exception of C09_HUMAN (represented by its gene name C9; Figure 4).

Imaging parameters

To investigate whether the protein profiles were associated with the patient imaging characteristics, volume of core and penumbra (defined as the difference between perfusion volume and core volume) was used as continuous variables. Baseline samples (or T0 when baseline was not available) from all 30 patients were analysed. No single protein was shown to be significantly associated with the core (defined as DWI volume) or the penumbra (defined as DWI-PWI volumes) volumes. No proteins were significantly associated with patients presenting with clinical diffusion mismatch (defined as the combination of NIHSS ≥ 8 and DWI < 25 mL; n = 14).

Finally, protein expressions were compared regarding the type of imaging used. Based on the analysis of the serum sample collected immediately after the completion of the imaging (baseline when CTp was used and T0 when MRI was used). No proteins were significantly different between the CTp and the MRI groups. Nevertheless, there was a clear difference in adjusted *P* value with the first 2 proteins and the rest of the data set. These 2 proteins (Uniprot IDs: Q8N5F4_HUMAN and Q6MZQ6_HUMAN) reaching an adjusted *P* value of <0.1 are currently uncharacterised proteins but both contain some immunoglobulin domains.
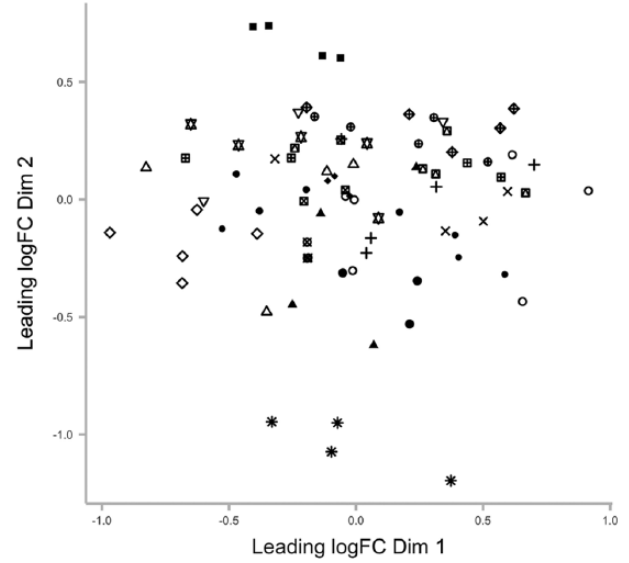
Treatment

Treatment group allocation remained blinded until completion of the protein analysis between treatment groups A and B. Treatment allocation (rt-PA versus placebo) is not presented in the patient characteristics summary table (Table 1) as EXTEND is still an ongoing trial.

Changes due to treatment administration were evaluated by comparing protein expression of baseline samples with T0 samples (or T12-24 if T0 was not available) in each patient. One

Table 3. Patients demographic characteristics for the 9 patients recruited in the EXTEND cohort to enriched the imaging profiles.

PATIENT	AGE	GENDER	BASELINE NIHSS	12-24 H NIHSS	DAY 3-7 NIHSS	DAY 90 NIHSS	PRESTROKE MRS	DAY 90 MRS	DEATH <3 MO	TOAST	IMAGING TYPE	DWI, ML	PWI, ML	PH	HTA	AF	LIPID DISORDER	PREVIOUS STROKE	ISCHAEMIC HEART DISEASE	DIABETES	PERIPH. VASC. DISEASE
22	85	M	5	5	4	3	1	1	No	Other	MRI	3	3	No	No	No	No	No	No	No	Yes
23	64	F	7	5	5	2	0	1	No	Other	MRI	4	3	No	Yes	No	No	No	No	No	No
24	47	F	3	3	3	1	0	2	No	SVO	MRI	9	0	No	No	Yes	No	No	No	No	No
25	68	M	8	7	7	1	1	1	No	CE	MRI	20	10	No	Yes	Yes	No	Yes	Yes	No	No
26	88	F	—	—	—	—	1	6	Yes	CE	MRI	144	176	Yes	Yes	Yes	Yes	No	Yes	Yes	No
27	64	F	23	38	—	—	0	6	Yes	CE	MRI	141	181	No	Yes	Yes	Yes	No	No	No	No
28	70	M	25	20	16	10	0	5	No	LAA	MRI	78	92	No	No	No	No	No	No	No	No
29	62	M	19	18	18	16	0	4	No	LAA	MRI	131	168	No	Yes	Yes	No	No	No	No	No
30	78	M	22	33	—	—	0	6	Yes	LAA	CTp	108	106	Yes	No	No	Yes	No	Yes	Yes	No

Abbreviations: AF, atrial fibrillation; CE, cardioembolic; CTp, perfusion computed tomography; DWI, diffusion-weighted imaging; F, female; HTA, arterial hypertension; LAA, large artery atherosclerosis; M, male; MRI, magnetic resonance imaging; mRS, modified ranking scale; NIHSS, National Institutes of Health Stroke Scale; PH, parenchymal haemorrhage; Periph. vasc. disease, peripheral vascular disease; PWI, perfusion-weighted imaging; SVO, small vessel occlusion; TOAST, trial of ORG 10172 in acute stroke treatment; und/inc, undetermined/unclassified. Missing data are represented by —.


Figure 1. Multidimensional scaling plot showing distances between samples based on top 50 most different proteins between samples. All patients measured at multiple time points are included and distinguished by symbols.
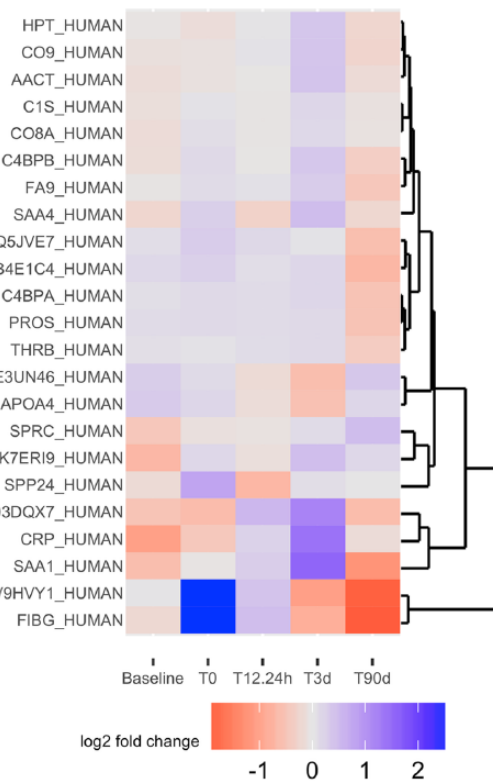
Figure 2. Heatmap of the 23 proteins significantly changing over time. Heatmap showing log fold change for proteins with a significant change in expression as a function of time (adjusted $P \leq 0.1$ based on F test). Relative expression values are shown as log 2 fold change relative to the average across all times.

patient was excluded for treatment effect analysis as its initial baseline sample was collected after treatment administration and this patient had therefore no real baseline available.
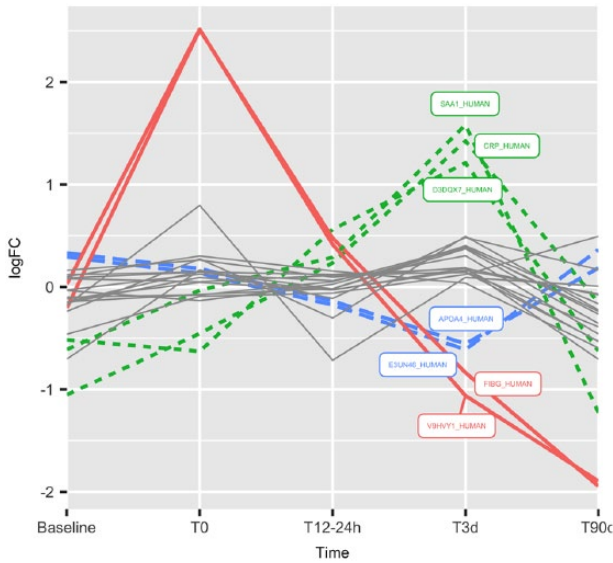
Figure 3. Expression profile for proteins showing significant (FDR = 0.1) changes over time. Expression is shown as log 2 fold change relative to the average across all time points. Seven proteins with 3 distinctly different expression profiles are highlighted with bold coloured lines. FDR indicates false discovery rate.

Figure 4. Interactions between the time-changing proteins. String-db plots shows interactions between proteins that showed significant change over time. A green line indicates neighbourhood evidence; a blue line – concurrence evidence; a purple line – experimental evidence; a yellow line – text mining evidence; a light blue line – database evidence; and a black line – expression evidence.

In total, 26 proteins were found to be significantly differently expressed after treatment between the 2 treatment groups (FDR ≤ 0.05). Their protein IDs, average expression levels, associated adjusted *P* value, and the direction of the change are presented in Table 4.

The 26 differentially expressed proteins by treatment administration were loaded into STRING10, and 13 were successfully matched showing an interaction network (Figure 5).

Outcome, severity, stroke subtypes, cardiovascular risk factors, and risk of PH

Functional outcome, clinical severity, stroke subtypes, cardiovascular risk factors, and risk of PH profiles were evaluated using samples collected from the 21 patients with stroke recruited in the randomised arm of EXTEND as multiple blood collection and clinical evaluation time points were available.

Potential biomarkers for good or bad outcome prediction were evaluated using the correlation of proteins expression on baseline sample and shift of mRS score between prestroke mRS and mRS at 3 months. Expression of von Willebrand factor (VWF) was significantly associated with an increase in mRS at 3 months (adjusted *P* = .04).

Protein expression levels were also regressed with stroke severity defined by NIHSS score at all collection time points. No significant changes in protein expression were identified in this context.

Using baseline sample (or T0 when baseline sample was not available), there were no significant changes in protein expression associated with strokes caused by large artery atherosclerosis or by cardioembolism.

Similarly, past medical history of hypertension, diabetes, atrial fibrillation, lipid disorder, ischaemic heart disease, and peripheral vascular disease were also each compared with protein levels measured in the baseline sample (or T0 when baseline sample was not available; n = 3-18). No significant differences in protein profiles were identified with any of the different cardiovascular risk factors investigated.
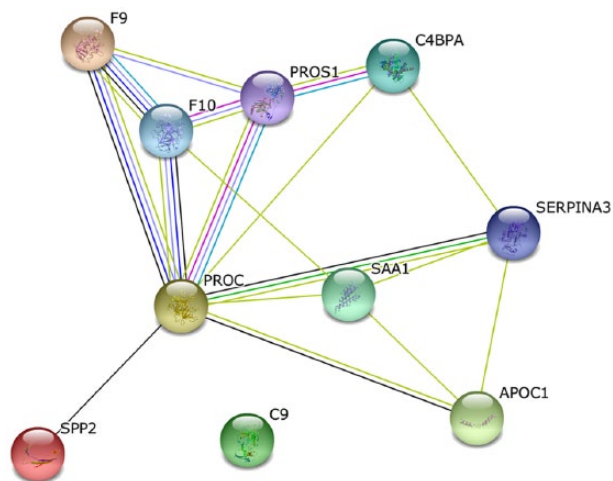
Finally, no specific proteins were significantly associated with the presence of a PH after stroke (by comparing baseline sample or T0 when baseline was not available and T12-24 between patients with PH and patients without PH; n = 5 and 16, respectively).

Discussion

The results of this study demonstrate the usefulness of the proteomic approach for identifying candidate serum biomarkers in the acute stroke setting. This clinical experiment is novel in attempting to identify biomarkers acting as a stroke clock and novel in examining the possibility of correlation between well-used imaging biomarkers and their potential bloodborne equivalents. In addition, this is to our knowledge the first time that MS technology has been used to achieve these aims.

Microarray studies have dominated the stroke biomarker research field in the past decade.[31] Following the lead of Frank Sharp's team, microarray studies have shown some promising results.[19,20,22,58] In an animal model, our team recently identified reproducible patterns of gene expression changing over time after IS and demonstrated that this change occurs hyperacutely after the ischaemic event (manuscript under review).

However, measurement of RNA with microarrays has several limitations with respect to translation of their results into

Table 4. Differentially expressed proteins before and after treatment administration.

UNIPROT ID	GENE NAME	LOGFC	AVERAGE EXPRESSION	ADJUSTED *P* VALUE
F5GXS5_HUMAN	APOF	−1.59	23.99	.014
A2AP_HUMAN	SERPINF2	−1.21	27.77	.03
KV119_HUMAN		−1.12	24.43	.03
KV308_HUMAN		−1.51	27.2	.03
KV302_HUMAN		−1.42	24.03	.032
KV121_HUMAN		−1.15	25.27	.032
KV303_HUMAN		−1.83	26.65	.032
B7Z549_HUMAN	ITIH1	−0.75	30.27	.032
KAIN_HUMAN	SERPINA4	−1.05	25.92	.032
S6B2B6_HUMAN		−1.68	23.37	.032
Q5SQS3_HUMAN	MBL2	−1.15	22.92	.032
KV122_HUMAN		−1.22	23.76	.032
A0N7J6_HUMAN		−1.08	23.93	.032
TTHY_HUMAN	TTR	−0.85	31.36	.032
HEMO_HUMAN	HPX	−0.61	33.15	.032
KV304_HUMAN		−1.19	22.65	.039
Q86U78_HUMAN	AGT	−1.01	27.95	.039
Q0ZCH9_HUMAN		−1.37	28.08	.039
CO4B_HUMAN	C4B	−0.61	32.71	.042
G3V0E5_HUMAN	TFRC	−0.95	21.85	.042
Q5JP69_HUMAN	C2	−0.99	24.55	.042
LBP_HUMAN	LBP	−1.32	25.19	.042
Q9UL92_HUMAN		−1.11	23.42	.044
Q9UL90_HUMAN		−0.85	25.83	.044
FIBA_HUMAN	FGA	2.58	29.25	.044
KV102_HUMAN		−2.01	22.58	.046

Negative logFC values mean than expression is higher in placebo than in recombinant tissue plasminogen activator. Average expression is given as log2 of the average intensity.

the clinical setting. First, the technology for rapid quantitation of RNA is in its infancy. Second, because not all RNAs code for proteins, discovery at the RNA level does not guarantee finding a matched protein that might be easier to measure. Finally, serum protein levels may differ from the expression patterns due to post-translational changes and regulatory feedback mechanisms and therefore gene expression may not be an accurate reflection of protein expression during cerebral ischaemia.[59–61]

Modern proteomic technology allows a high-throughput method to quantitatively analyse a large number of proteins present in a disease state. By taking an agnostic and data-driven approach rather than hypothesis-driven approach, an unbiased evaluation of the proteome can identify proteins or pathways not previously considered as potential candidate biomarkers.

Candidate biomarkers have rarely been selected in the context of looking for a stroke clock. Turck et al analysed blood samples taken at 2 different time points (within 3 hours and within 36 hours) after stroke but had to combine 2 separate cohorts of patients to achieve their aim.[62] The only study collecting sequential human blood samples in the same individual early after stroke was conducted by Tang et al[20] in 2006 but their attention focused on elaboration of a diagnostic tool capable of differentiating patients with IS from controls, and the earliest blood samples were collected 3 hours after the ischaemic event.
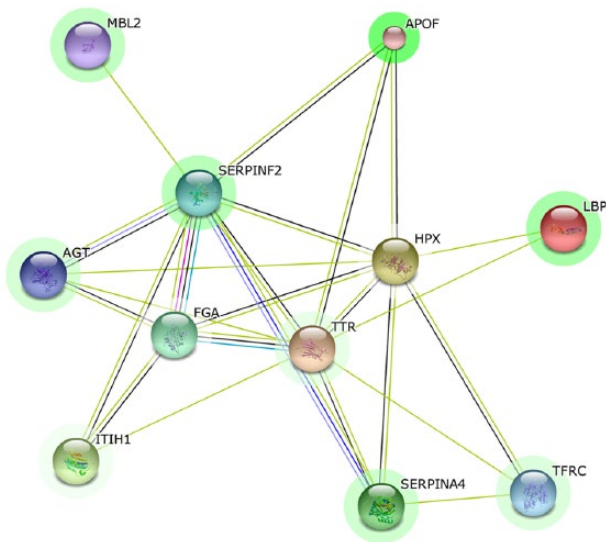
Figure 5. Interactions between the differentially expressed proteins before and after treatment analysis. String-db plots shows interactions between proteins that showed significant change between the 2 treatment groups. A green line indicates neighbourhood evidence; a blue line – concurrence evidence; a purple line – experimental evidence; a yellow line – text mining evidence; a light blue line – database evidence; a black line – expression evidence. Halos around protein names show logFC with green showing proteins that are higher in patients treated with rt-PA than in patients receiving placebo.

In this experiment, we identified 23 proteins whose expression significantly varied over time after IS. Proteins expression profiles showed a range of patterns including downregulation and upregulation and early or delayed peaks of expression. Some proteins (such as the apolipoprotein A-IV) showed complex expression profiles with more than one peak of expression and varied in opposite directions over time. This highlights the importance of accounting for time when searching for biomarkers in dynamic diseases such as stroke.

Other apolipoproteins (APOC-I and APOC-II) have been previously reported as plasma biomarkers capable of accurately distinguishing between ischaemic and haemorrhagic strokes.[44] Nevertheless, these molecules are yet to reach the clinical world.

Not surprisingly, pathway analysis revealed that a number of these proteins (such as vitamin K–dependent protein S [PROS], haptoglobin, prothrombin, fibrinogen, and coagulation factors IX and X) are involved in the coagulation cascade.

The fibrinogen profile appears to have a high potential for clinical utility in the acute stroke setting. Indeed, fibrinogen expression differed dramatically during the first 24 hours after IS. More interestingly, levels of expression even increased between the samples collected at baseline and T0. As our baseline sample was collected 4.5 hours after symptoms onset at the earliest, we postulate that changes of fibrinogen availability might progressively increase from stroke onset and reach a peak of expression during the acute phase of stroke. Comparison of 2 consecutive levels of fibrinogen might therefore be used to

estimate the time window from symptoms onset and help to decide if a patient is eligible for thrombolysis.

The epididymis secretory sperm-binding protein (gene V9HVY1_HUMAN) has a similar expression profile to the fibrinogen expression that fits with its role in the fibrinogen complex (http://www.uniprot.org/uniprot/V9HVY1).

In contrast, CRP expression peaked at 3 days after stroke. The CRP has extensively been studied as a molecule with the potential to differentiate between haemorrhagic stroke and IS.[30] Our data suggest that CRP might be more useful as a later time point than for early stroke diagnosis.

Three members of the serum amyloid protein A family were also present in the molecules significantly changing over time after IS. As serum amyloid proteins have rarely been associated with IS,[44,63,64] this result offers a new incentive to study this class of molecules more extensively and specifically.

Only one other study has investigated the changes of candidate biomarkers across time using proteomics technologies in blood of patients with IS.

In 2010, Ning et al reported differences of protein expression before and after treatment in thrombolysed patients with stroke. They demonstrated that rt-PA treatment changed the plasma degradomic profiles in patients with acute stroke. More specifically, rt-PA treatment induced changes in the expression of some molecules known to be involved in the thrombolysis pathway such as fibrinogen and α2 macroglobulin and in the IS cascade such as transthyretin and spectrin.[65] Our experiment confirmed Ning's results regarding fibrinogen and transthyretin changes after administration of rt-PA. However, brain-specific angiogenesis inhibitor (BAI-1) and utrophin (2 molecules newly identified by Ning et al as potential candidate IS biomarker) were not significantly differently expressed between our 2 treatment groups.

We also identified VWF as a marker of functional outcome after IS. Indeed, levels of VWF expression were associated with a shift of mRS score before and at 3 months after stroke. Known to be a key player in thrombus formation at sites of vascular damage, high plasma levels of VWF have previously been identified as a strong predictor of stroke as well as being associated with stroke subtypes and outcome.[66–68]

Recent studies, where blood samples were processed by microarrays or ELISA, have identified specific signatures for the different stroke causes.[69–71] Due to the low number of patients in our data set presenting with each of the TOAST entities, we did not identify any significant correlation between protein expression and stroke types.

Similarly, we did not identify any protein biomarker significantly associated with DWI or the penumbral volume. This is interesting despite some molecules having been previously described as significantly associated in some imaging-based selected populations. For example, Rodríguez-Yáñez et al[72] showed that serum levels of interleukin 10, tumour necrosis factor α, and glutamate were significantly higher in patients

with stroke presenting with clinical diffusion mismatch, whereas levels of neuron-specific enolase, interleukin 6, and MMP-9 were shown to be significantly lower.

These differences in discovery may reflect the difference in sensitivity of the different methods used here and in previous studies.

This might also explain why some molecules known to be associated with the stroke physiopathology such as chemokines and cytokines were not detected in this experiment. A total of 22 proteins comprise more than 90% of the total protein mass in human serum (and albumin alone accounts for more than 50%), and only 1% of the entire protein content of serum consists of proteins that are considered to be of low abundance and possibly of greater interest in the search of potential biomarkers.[73,74] The masking effect of these high-abundance proteins could be reduced using depletion strategies such as antibody-based affinity chromatography and immunoaffinity depletion columns.[74,75] Additional offline fractionation and longer liquid chromatography (LC) gradients might also be used for increased protein identification. Nevertheless, this study was performed with the hypothesis that it was a pre-requisite to check whether high-abundance proteins had a critical role in the IS pathophysiology and might serve as good biomarkers before studying lower abundant proteins. Combining a high-resolution MS instrument with high-performance LC and 5 hours elution gradient – which represented a total of 45 days of instrument running time – allowed us to identify a total of 574 different protein groups.

We are aware that candidate biomarker discovery with MS has rarely translated into clinical utility and this is mainly due to false discovery and lack of reproducibility.[76] The gap between biomarker discovery and biomarker validation remains a challenge. Our results highlighting major temporal changes illustrate that describing precisely when a sample is collected will have a significant impact on stroke biomarker discovery. They also suggest that it may be possible to generate a predictive stroke clock. We are also conscious that this article describes a discovery study with a relatively small group of patients. Validation by other measuring techniques such as ELISA or quantitative polymerase chain reaction as well as replication in a larger population and compared with control subjects will be important steps towards the use of such biomarkers in the clinic.

Conclusions

Our data suggest that serum biomarkers acting as a stroke clock will complement multimodal imaging in the future. They will improve the selection of patients for reperfusion therapies and therefore ultimately increase the number of patients presenting with acute IS that will benefit from rt-PA or newer therapies, including clot retrieval.

Author Contributions

MD conceived and designed the study, participated in the samples processing and in the statistical analysis and drafted the manuscript. IRC performed the statistical analysis and helped to draft the manuscript. PF performed the samples processing and helped to draft the manuscript. TKS helped to conceived the study and participated in the samples processing. HMD helped to draft the manuscript. GAD participated in the study design. DWH conceived and designed the study, supervised its execution, participated in the statistical analysis and helped to draft the manuscript.

REFERENCES

1. Roger VL, Go AS, Lloyd-Jones DM, et al. Executive summary: heart disease and stroke statistics – 2012 update: a report from the American Heart Association. *Circulation*. 2012;125:188–197.
2. Hacke W, Kaste M, Bluhmki E, et al. Thrombolysis with alteplase 3 to 4.5 hours after acute ischemic stroke. *N Engl J Med*. 2008;359:1317–1329.
3. Jauch EC, Saver JL, Adams HP, et al. Guidelines for the early management of patients with acute ischemic stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke*. 2013;44:870–947.
4. Howells DW, Donnan GA. Where will the next generation of stroke treatments come from? *PLoS Med*. 2010;7:e1000224.
5. Barber PA, Zhang J, Demchuk AM, Hill MD, Buchan AM. Why are stroke patients excluded from TPA therapy? an analysis of patient eligibility. *Neurology*. 2001;56:1015–1020.
6. Fink JN, Kumar S, Horkan C, et al. The stroke patient who woke up: clinical and radiological features, including diffusion and perfusion MRI. *Stroke*. 2002;33:988–993.
7. Faiz KW, Sundseth A, Thommessen B, Ronning OM. Reasons for low thrombolysis rate in a Norwegian ischemic stroke population. *Neurol Sci*. 2014; 35:1977–1982.
8. Albers GW, Amarenco P, Easton JD, Sacco RL, Teal P. Antithrombotic and thrombolytic therapy for ischemic stroke: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). *Chest*. 2008;133:630S–669S.
9. Fonarow GC, Smith EE, Saver JL, et al. Timeliness of tissue-type plasminogen activator therapy in acute ischemic stroke: patient characteristics, hospital factors, and outcomes associated with door-to-needle times within 60 minutes. *Circulation*. 2011;123:750–758.
10. Minnerup J, Wersching H, Ringelstein EB, et al. Impact of the extended thrombolysis time window on the proportion of recombinant tissue-type plasminogen activator-treated stroke patients and on door-to-needle time. *Stroke*. 2011;42:2838–2843.
11. Krogias C, Bartig D, Kitzrow M, Weber R, Eyding J. Trends of hospitalized acute stroke care in Germany from clinical trials to bedside. Comparison of nation-wide administrative data 2008-2012. *J Neurol Sci*. 2014;345:202–208.
12. The National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. Tissue plasminogen activator for acute ischemic stroke. *N Engl J Med*. 1995;333:1581–1587.
13. Davis S, Donnan GA. Time is Penumbra: imaging, selection and outcome. The Johann Jacob Wepfer Award 2014. *Cerebrovasc Dis*. 2014;38:59–72.
14. Wintermark M, Luby M, Bornstein NM, et al. International survey of acute stroke imaging used to make revascularization treatment decisions. *Int J Stroke*. 2015;10:759–762.
15. Muir KW, Buchan A, von Kummer R, Rother J, Baron J-C. Imaging of acute stroke. *Lancet Neurol*. 2006;5:755–768.
16. Kane I, Sandercock P, Wardlaw J. Magnetic resonance perfusion diffusion mismatch and thrombolysis in acute ischaemic stroke: a systematic review of the evidence to date. *J Neurol Neurosurg Psychiatry*. 2007;78:485–491.
17. Barreto AD, Martin-Schild S, Hallevi H, et al. Thrombolytic therapy for patients who wake-up with stroke. *Stroke*. 2009;40:827–832.
18. Kang DW, Kwon JY, Kwon SU, Kim JS. Wake-up or unclear-onset strokes: are they waking up to the world of thrombolysis therapy? *Int J Stroke*. 2012;7:311–320.
19. Moore DF, Li H, Jeffries N, et al. Using peripheral blood mononuclear cells to determine a gene expression profile of acute ischemic stroke: a pilot investigation. *Circulation*. 2005;111:212–221.
20. Tang Y, Xu H, Du X, et al. Gene expression in blood changes rapidly in neutrophils and monocytes after ischemic stroke in humans: a microarray study. *J Cereb Blood Flow Metab*. 2006;26:1089–1102.
21. Stamova B, Xu H, Jickling G, et al. Gene expression profiling of blood for the prediction of ischemic stroke. *Stroke*. 2010;41:2171–2177.
22. Barr TL, Conley Y, Ding J, et al. Genomic biomarkers and cellular pathways of ischemic stroke by RNA gene expression profiling. *Neurology*. 2010;75:1009–1014.

23. Baird AE. Blood genomics in human stroke. *Stroke*. 2007;38:694–698.

24. 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.

25. Wakisaka Y, Ago T, Kamouchi M, et al. Plasma S100A12 is associated with functional outcome after ischemic stroke: research for biomarkers in ischemic stroke. *J Neurol Sci*. 2014;340:75–79.

26. Yang XY, Gao S, Ding J, Chen Y, Zhou XS, Wang JE. Plasma D-dimer predicts short-term poor outcome after acute ischemic stroke. *PLoS ONE*. 2014;9:e89756.

27. Garcia-Berrocoso T, Giralt D, Bustamante A, et al. Role of beta-defensin 2 and interleukin-4 receptor as stroke outcome biomarkers. *J Neurochem*. 2014; 129:463–472.

28. Elkind MS, Luna JM, McClure LA, et al. C-reactive protein as a prognostic marker after lacunar stroke: levels of inflammatory markers in the treatment of stroke study. *Stroke*. 2014;45:707–716.

29. Garcia-Berrocoso T, Penalba A, Boada C, et al. From brain to blood: new biomarkers for ischemic stroke prognosis. *J Proteomics*. 2013;94:138–148.

30. Hasan N, McColgan P, Bentley P, Edwards RJ, Sharma P. Towards the identification of blood biomarkers for acute stroke in humans: a comprehensive systematic review. *Br J Clin Pharmacol*. 2012;74:230–240.

31. Sharp FR, Jickling GC. Whole genome expression of cellular response to stroke. *Stroke*. 2013;44:S23–S25.

32. Whiteley W, Chong WL, Sengupta A, Sandercock P. Blood markers for the prognosis of ischemic stroke: a systematic review. *Stroke*. 2009;40:e380–e389.

33. Boja ES, Rodriguez H. Mass spectrometry-based targeted quantitative proteomics: achieving sensitive and reproducible detection of proteins. *Proteomics*. 2012;12:1093–1110.

34. Quinn DA, Fogel RB, Smith CD, et al. D-dimers in the diagnosis of pulmonary embolism. *Am J Respir Crit Care Med*. 1999;159:1445–1449.

35. Adams J, Bodor G, Dávila-Román V, et al. Cardiac troponin I. A marker with high specificity for cardiac injury. *Circulation*. 1993;88:101–106.

36. Shapiro B, Chen H, Burnett JJ, Redfield M. Use of plasma brain natriuretic peptide concentration to aid in the diagnosis of heart failure. *Mayo Clin Proc*. 2003;78:481–486.

37. Cuadrado E, Rosell A, Colomé N, et al. The proteome of human brain after ischemic stroke. *J Neuropathol Exp Neurol*. 2010;69:1105–1115.

38. Cuadrado E, Rosell A, Penalba A, et al. Vascular MMP-9/TIMP-2 and neuronal MMP-10 up-regulation in human brain after stroke: a combined laser microdissection and protein array study. *J Proteome Res*. 2009;8:3191–3197.

39. Dawson J, Walters M, Delles C, Mischak H, Mullen W. Urinary proteomics to support diagnosis of stroke. *PLoS ONE*. 2012;7:e35879.

40. Dayon L, Turck N, García-Berrocoso T, et al. Brain extracellular fluid protein changes in acute stroke patients. *J Proteome Res*. 2011;10:1043–1051.

41. Dayon L, Hainard A, Licker V, et al. Relative quantification of proteins in human cerebrospinal fluids by MS/MS using 6-plex isobaric tags. *Anal Chem*. 2008;80:2921–2931.

42. Maurer M, Haux D, Unterberg A, Sakowitz O. Proteomics of human cerebral microdialysate: from detection of biomarkers to clinical application. *Proteomics Clin Appl*. 2008;2:437–443.

43. Ray S, Reddy P, Jain R, Gollapalli K, Moiyadi A, Srivastava S. Proteomic technologies for the identification of disease biomarkers in serum: advances and challenges ahead. *Proteomics Clin Appl*. 2011;11:2139–2161.

44. Allard L, Lescuyer P, Burgess J, et al. ApoC-I and ApoC-III as potential plasmatic markers to distinguish between ischemic and hemorrhagic stroke. *Proteomics*. 2004;4:2242–2251.

45. Lopez MF, Sarracino DA, Prakash A, et al. Discrimination of ischemic and hemorrhagic strokes using a multiplexed, mass spectrometry-based assay for serum apolipoproteins coupled to multi-marker ROC algorithm. *Proteomics Clin Appl*. 2012;6:190–200.

46. Zhang X, Guo T, Wang H, et al. Potential biomarkers of acute cerebral infarction detected by SELDI-TOF-MS. *Am J Clin Pathol*. 2008;130:299–304.

47. Huang P, Lo LH, Chen YC, Lin RT, Shiea J, Liu CK. Serum free hemoglobin as a novel potential biomarker for acute ischemic stroke. *J Neurol*. 2009;256:625–631.

48. George PM, Mlynash M, Adams CM, Kuo CJ, Albers GW, Olivot JM. Novel TIA biomarkers identified by mass spectrometry-based proteomics. *Int J Stroke*. 2015;10:1204–1211.

49. Ma H, Parsons MW, Christensen S, et al. A multicentre, randomized, double-blinded, placebo-controlled Phase III study to investigate EXtending the time for Thrombolysis in Emergency Neurological Deficits (EXTEND). *Int J Stroke*. 2012;7:74–80.

50. Ishihama Y, Rappsilber J, Mann M. Modular stop and go extraction tips with stacked disks for parallel and multidimensional peptide fractionation in proteomics. *J Proteome Res*. 2006;5:988–994.

51. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol*. 2008;26:1367–1372.

52. R Core Team. R: A language and environment for statistical computing. http://www.R-project.org/. Published 2015.

53. Ritchie M, Phipson B, Wu D, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43:e47.

54. Bolstad B, Irizarry R, Astrand M, Speed T. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*. 2003;19:185–193.

55. Ritchie M, Diyagama D, Neilson J, et al. Empirical array quality weights in the analysis of microarray data. *BMC Bioinformatics*. 2006;7:261.

56. Smyth G. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol*. 2004;3:Article3.

57. Jensen LJ, Kuhn M, Stark M, et al. STRING 8 – a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res*. 2009;37:D412–D416.

58. Tang Y, Lu A, Aronow BJ, Sharp FR. Blood genomic responses differ after stroke, seizures, hypoglycemia, and hypoxia: blood genomic fingerprints of disease. *Ann Neurol*. 2001;50:699–707.

59. Ansong C, Purvine SO, Adkins JN, Lipton MS, Smith RD. Proteogenomics: needs and roles to be filled by proteomics in genome annotation. *Brief Funct Genomic Proteomic*. 2008;7:50–62.

60. Humphrery-Smith I, Cordwell S, Blackstock W. Proteome research: complementarity and limitations with respect to the RNA and DNA worlds. *Electrophoresis*. 1997;18:1217–1242.

61. Arrell D, Neverova I, Van Eyk J. Cardiovascular proteomics: evolution and potential. *Circ Res*. 2001;88:763–773.

62. Turck N, Robin X, Walter N, et al. Blood glutathione S-transferase-π as a time indicator of stroke onset. *PLoS ONE*. 2012;7:e43830.

63. Sezer S, Uçar F, Ulusoy EK, et al. Serum amyloid A, fetuin-A, and pentraxin-3 levels in patients with ischemic stroke: novel prognostic biomarkers? *Turk J Med Sci*. 2014;44:16–23.

64. Brea D, Sobrino T, Blanco M, et al. Usefulness of haptoglobin and serum amyloid A proteins as biomarkers for atherothrombotic ischemic stroke diagnosis confirmation. *Atherosclerosis*. 2009;205:561–567.

65. Ning M, Sarracino DA, Buonanno FS, et al. Proteomic protease substrate profiling of tPA treatment in acute ischemic stroke patients: a step toward individualizing thrombolytic therapy at the bedside. *Transl Stroke Res*. 2010;1:268–275.

66. Wieberdink R, van Schie M, Koudstaal P, et al. High von Willebrand factor levels increase the risk of stroke: the Rotterdam study. *Stroke*. 2010;41:2151–2156.

67. Bath P, Blann A, Smith N, Butterworth R. Von Willebrand factor, P-selectin and fibrinogen levels in patients with acute ischaemic and haemorrhagic stroke, and their relationship with stroke sub-type and functional outcome. *Platelets*. 1998;9:155–159.

68. De Meyer SF, Stoll G, Wagner DD, Kleinschnitz C. Von Willebrand factor: an emerging target in stroke therapy. *Stroke*. 2012;43:599–606.

69. Xu H, Tang Y, Liu DZ, et al. Gene expression in peripheral blood differs after cardioembolic compared with large-vessel atherosclerotic stroke: biomarkers for the etiology of ischemic stroke. *J Cereb Blood Flow Metab*. 2008;28:1320–1328.

70. Jickling GC, Xu H, Stamova B, et al. Signatures of cardioembolic and large-vessel ischemic stroke. *Ann Neurol*. 2010;68:681–692.

71. Hanson E, Jood K, Karlsson S, Nilsson S, Blomstrand C, Jern C. Plasma levels of von Willebrand factor in the etiologic subtypes of ischemic stroke. *J Thromb Haemost*. 2011;9:275–281.

72. Rodríguez-Yáñez M, Sobrino T, Arias S, et al. Early biomarkers of clinical-diffusion mismatch in acute ischemic stroke. *Stroke*. 2011;42:2813–2818.

73. Tirumalai R, Chan K, Prieto D, Issaq H, Conrads T, Veenstra T. Characterization of the low molecular weight human serum proteome. *Mol Cell Proteomics*. 2003;2:1096–1103.

74. Haudenschild DR, Eldridge A, Lein PJ, Chromy BA. High abundant protein removal from rodent blood for biomarker discovery. *Biochem Biophys Res Commun*. 2014;455:84–89.

75. Bellei E, Bergamini S, Monari E, et al. High-abundance proteins depletion for serum proteomic analysis: concomitant removal of non-targeted proteins. *Amino Acids*. 2011;40:145–156.

76. Rifai N, Gillette MA, Carr SA. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat Biotechnol*. 2006;24:971–983.