Circulating Metabolites Differentiate Acute Ischemic Stroke from Stroke Mimics

Steffen Tiedt, MD, PhD,[1,2] Stefan Brandmaier, PhD,[3,4] Hanna Kollmeier,[1] Marco Duering, MD,[1,2] Anna Artati, PhD,[5] Jerzy Adamski, PhD,[5,6,7,8] Matthias Klein, MD,[9] Thomas Liebig, MD,[10] Lesca M. Holdt, MD, PhD,[11] Daniel Teupser, MD,[11] Rui Wang-Sattler, PhD,[3,4,7] Edzard Schwedhelm, MD,[12,13] Christian Gieger, PhD,[3,4,7] and Martin Dichgans, MD[1,2]

Objective: Early discrimination of patients with ischemic stroke (IS) from stroke mimics (SMs) poses a diagnostic challenge. The circulating metabolome might reflect pathophysiological events related to acute IS. Here, we investigated the utility of early metabolic changes for differentiating IS from SM.
Methods: We performed untargeted metabolomics on serum samples obtained from patients with IS (N = 508) and SM (N = 349; defined by absence of a diffusion weighted imaging [DWI] positive lesion on magnetic resonance imaging [MRI]) who presented to the hospital within 24 hours after symptom onset (median time from symptom onset to blood sampling = 3.3 hours; interquartile range [IQR] = 1.6–6.7 hours) and from neurologically normal controls (NCs; N = 112). We compared diagnostic groups in a discovery-validation approach by applying multivariable linear regression models, machine learning techniques, and propensity score matching. We further performed a targeted look-up of published metabolite sets.
Results: Levels of 41 metabolites were significantly associated with IS compared to NCs. The top metabolites showing the highest value in separating IS from SMs were asymmetrical and symmetrical dimethylarginine, pregnenolone sulfate, and adenosine. Together, these 4 metabolites differentiated patients with IS from SMs with an area under the curve (AUC) of 0.90 in the replication sample, which was superior to multimodal cranial computed tomography (CT; AUC = 0.80) obtained for routine diagnostics. They were further superior to previously published metabolite sets detected in our samples. All 4 metabolites returned to control levels by day 90.
Interpretation: A set of 4 metabolites with known biological effects relevant to stroke pathophysiology shows unprecedented utility to identify patients with IS upon hospital arrival, thus encouraging further investigation, including multicenter studies.

ANN NEUROL 2020;88:736–746

Stroke is the leading cause of adult disability[1] with the majority caused by brain infarction. Discriminating patients with ischemic stroke (IS) from stroke mimics (SMs) is important, as misclassification may result in inappropriate interventions, in withholding treatments with proven efficacy, and in inefficient allocation of resources, such as stroke unit care. Routine diagnostics with multimodal computed tomography (CT) has limited

View this article online at wileyonlinelibrary.com. DOI: 10.1002/ana.25859

Received Dec 4, 2019, and in revised form Jul 29, 2020. Accepted for publication Jul 30, 2020

[Correction added on 23 February, 2021, after first online publication: Copyright line was changed and Projekt DEAL funding statement has been added]..

Address correspondence to Dr Dichgans, Institute for Stroke and Dementia Research, University Hospital, LMU Munich, Feodor-Lynen-Straße 17, 81377 Munich, Germany. E-mail: martin.dichgans@med.uni-muenchen.de

From the [1]Institute for Stroke and Dementia Research, University Hospital, LMU Munich, Munich, Germany; [2]Munich Cluster for Systems Neurology (SyNergy), Munich, Germany; [3]Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, Neuherberg, Germany; [4]Institute of Epidemiology, Helmholtz Zentrum München, Neuherberg, Germany; [5]Research Unit Molecular Endocrinology and Metabolism, Genome Analysis Center, Helmholtz Zentrum München, Neuherberg, Germany; [6]Institute of Experimental Genetics, Technical University of Munich, Freising, Germany; [7]German Center for Diabetes Research (DZD), Munich, Germany; [8]Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore; [9]Department of Neurology, University Hospital, LMU Munich, Munich, Germany; [10]Institute of Neuroradiology, University Hospital, LMU Munich, Munich, Germany; [11]Institute of Laboratory Medicine, University Hospital, LMU Munich, Munich, Germany; [12]Institute of Clinical Pharmacology and Toxicology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; and [13]German Centre for Cardiovascular Research (DZHK), Partner Site Hamburg / Kiel / Lübeck, Hamburg, Germany

Additional supporting information can be found in the online version of this article.

capacity for differentiating IS from SMs within the first hours after the event.[2–4] Hence, there is great demand for novel diagnostic markers. The circulating metabolome sensitively captures pathophysiological events in multiple organs and might thus reflect events related to acute IS, such as cerebral injury, systemic responses (eg, stress),[5] and processes within the vasculature.[6] Here, we used untargeted metabolomics in a discovery-validation approach to identify metabolites that would allow differentiating IS from SMs. We further performed a systematic literature review to identify circulating metabolites reportedly associated with IS in the first 24 hours after symptom onset and assessed their diagnostic utility in our dataset in relation to the metabolite set identified here.

Methods

Patient Enrolment

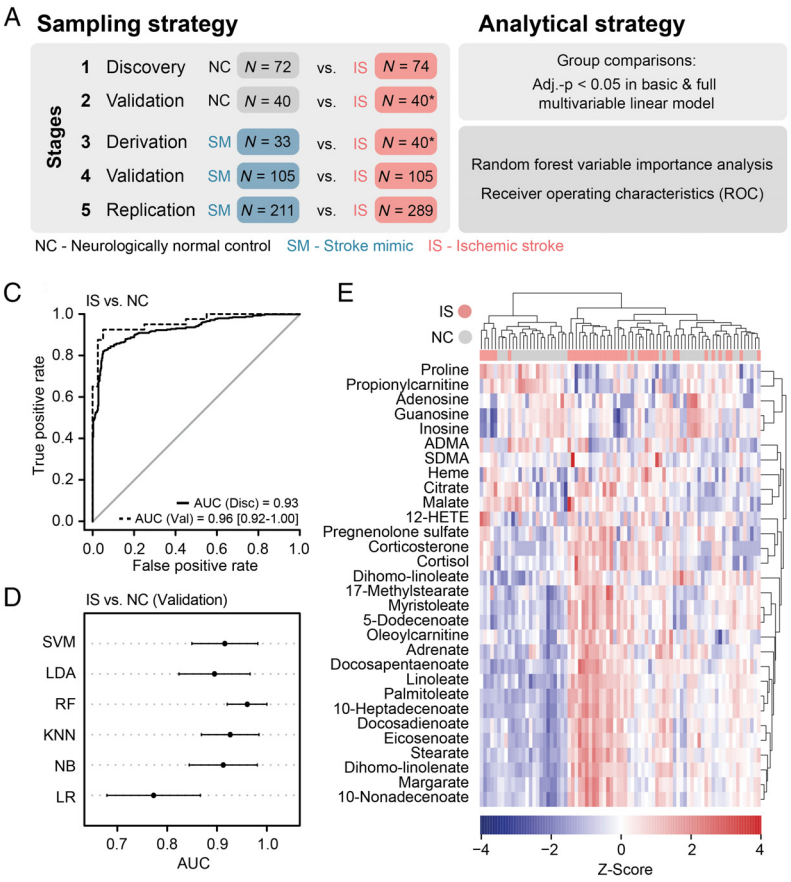
We recruited patients with rapidly developing clinical signs suggestive of stroke who presented to the emergency department of the medical center at LMU, Munich, Germany, a tertiary level center, within 24 hours of symptom onset. Blood sampling was performed upon hospital arrival and prior to routine diagnostic multimodal CT ("reference standard"). The final diagnosis of IS was based on the presence of a diffusion weighted imaging (DWI)-positive lesion on magnetic resonance imaging (MRI) or a new lesion on a delayed CT scan.[7] Patients without a DWI-positive lesion on MRI were classified as having SMs.[7] For all cases, the final decision about group allocation (IS or SM) was made after blood sampling. Stroke etiology was classified according to the TOAST system. Neurologically normal controls (NCs) were recruited through the stroke prevention unit at LMU. Subjects were excluded if they suffered from hemorrhagic stroke, active malignant disease, inflammatory or infectious disease, or had undergone surgery within the last 3 months. The discovery (stage 1) and validation (stage 2) samples (Fig 1A) were recruited between February 2014 and August 2015 and between August 2015 and November 2016, respectively. Patients with IS and NCs from the validation sample were matched with respect to age, sex, hypertension, smoking history, hypercholesterolemia, obesity, diabetes mellitus, family history of cardiovascular disease, and use of medication. The derivation sample (stage 3) included the IS group from stage 2 and patients with SMs recruited between February 2014 and November 2016. The validation sample (stage 4; see Fig 1A) was recruited between April 2015 and October 2018 and used propensity scores to match patients with IS and patients with SMs for age, sex, hypertension, diabetes mellitus, major cardiovascular events, and time from symptom onset to blood sampling (Table 1). These samples were convenience samples. The replication sample (stage 5; see Fig. 1A) was recruited between October 2018 and February 2020 and formed a consecutive series with hemorrhagic stroke as the only exclusion criterion. SM in the replication sample also included patients who received CT as the only imaging modality. Written informed consent was obtained from all subjects in accord with ethical approval. The study was conducted in accordance with the Declaration of Helsinki and is reported according to the Standards for Reporting of Diagnostic Accuracy Studies (STARD) guidelines.[8]

Neuroimaging

Multimodal CT was obtained as part of clinical routine using a standard protocol that included noncontrast CT, CT angiography, and CT perfusion. CT examinations were performed on SOMATOM Definition scanners (Siemens Healthcare, Forchheim, Germany). CT perfusion was obtained with 0.6-mm collimation and 100-mm scan coverage in the z-axis using adaptive spiral scanning. The datasets were acquired continuously over 48 seconds (32 cycles, 1 sweep every 1.5 seconds). CT perfusion data were processed using the manufacturer's software (Syngo Neuro Perfusion CT; Siemens Healthineers, Forchheim, Germany) to generate perfusion maps. MRI scanning was performed between 1 and 8 days after symptom onset on 1.5 Tesla or 3 Tesla scanners with 3 or 5 mm slice gaps. Infarct volumes were quantified as previously described.[9,10]

Metabolome Analyses

Untargeted metabolite profiling ("index test") was planned after data collection and performed by experimenters blinded to group allocations at the Genome Analysis Center (Helmholtz-Zentrum München) and by Metabolon (Durham, NC).[11] Serum samples from the discovery sample were analyzed with a Linear Ion Trap tandem mass spectrometry (LTQ-XL MS/MS; Thermo Scientific) coupled with Ultra Performance Liquid Chromatography (UPLC; Acquity-UPLC, Waters) and consisted of reverse phase (RP)/ UPLC-MS/MS analysis in positive and negative electrospray ionization (ESI) mode. All other samples were analyzed with a Q-Exactive high resolution/accurate mass spectrometry (MS) interfaced with a heated ESI source and Orbitrap mass analyzer (Thermo Scientific) coupled with UPLC and consisted of 2 separate (RP)/ UPLC-MS/MS methods with positive ESI mode, one with negative ESI mode, and (hydrophilic interaction liquid chromatography [HILIC])/UPLC-MS/MS with negative ESI mode. Curation was performed by Metabolon. Because of the identification of dimethylarginine in stage 1, we performed targeted MS-based analyses of methylated arginine derivatives evidencing significantly elevated levels of



B

Metabolite	MC	β	Adj.-p
Corticosterone	L	1.36	4.4×10^{-7}
Cortisol	L	1.20	1.2×10^{-5}
Myristoleate (14:1n5)	L	1.07	1.1×10^{-4}
Stearate (18:0)	L	1.07	1.1×10^{-4}
Palmitoleate (16:1n7)	L	1.05	1.1×10^{-4}
17-Methylstearate (i19:0)	L	0.98	1.7×10^{-4}
10-Heptadecenoate (17:1n7)	L	0.97	2.8×10^{-4}
Docosadienoate (22:2n6)	L	0.94	3.3×10^{-4}
Dihomo-linoleate (20:2n6)	L	0.89	0.001
Propionylcarnitine	L	-0.79	0.001
10-Nonadecenoate (19:1n9)	L	0.87	0.001
12-HETE	L	0.89	0.001
Docosapentaenoate (22:5n3)	L	0.91	0.001
Linoleate (18:2n6)	L	0.84	0.001
5-Dodecenoate (12:1n7)	L	0.87	0.001
Eicosenoate (20:1)	L	0.83	0.001
Margarate (17:0)	L	0.84	0.001
Pregnenolone sulfate	L	0.82	0.001
Dihomo-linolenate (20:3n3 or n6)	L	0.70	0.003
Adrenate (22:4n6)	L	0.70	0.005
Oleoylcarnitine (C18:1)	L	0.52	0.018
N-stearoyltaurine	L	0.22	0.136†
Glycerophosphorylcholine (GPC)	L	-0.15	0.169†
1-Stearoyl-GPE (18:0)	L	-0.13	0.183†
Citrate	E	0.67	0.005
Malate	E	0.49	0.018
Proline	A	-0.40	0.046
Dimethylarginine (SDMA/ADMA)	A	-0.04	0.245†
Heme	CV	0.86	0.001
Guanosine	N	-1.05	1.7×10^{-4}
Inosine	N	-0.99	3.2×10^{-4}
Adenosine	N	-0.82	0.002
Xanthosine	N	0.29	0.089†
2-Furoylglycine	X	-1.05	1.7×10^{-4}
3-Hydroxypyridine sulfate	X	-0.81	0.001
1,2,3-Benzenetriol sulfate	X	-0.73	0.003
Catechol sulfate	X	-0.61	0.009
4-Hydroxyhippurate	X	-0.48	0.020
Methyl-4-hydroxybenzoate sulfate	X	-0.35	0.070†
X-18307	U	1.00	1.7×10^{-4}
X-11470	U	0.73	0.001
X-17167	U	-0.71	0.005
X-12230	U	-0.63	0.006
X-11407	U	0.66	0.007
X-15220	U	0.60	0.012
X-13431	U	-0.53	0.018
X-17185	U	-0.49	0.023
X-17137	U	-0.19	0.155†
X-11442	U	0.08	0.224†

FIGURE 1: A discovery-validation approach identifies 30 metabolites associated with ischemic stroke (IS). (A) Sampling and analytical strategy employing a discovery-validation approach comparing patients with ischemic stroke (IS, N = 508) with neurologically normal controls (NCs, N = 112; stage 1 and 2) and with stroke mimics (SMs; N = 349; stages 3, 4, and 5). Diagnostic groups were compared by multivariable linear regression models and random forest regression. The diagnostic utility of metabolites was determined by receiver operating characteristic (ROC) analyses. (B) Forty-nine metabolites showed significantly altered levels in patients with IS compared to NCs in the discovery sample. Of these, 41 metabolites met the criteria for statistical significance also in the validation sample. Shown are results from the full model. (C and D) Thirty endogenous metabolites enabled differentiating patients with IS from NCs in a random forest classification (C) and in 5 other classification models (D). (E) Heatmap of the z-scores of the 30 endogenous metabolites. Metabolites and subjects were clustered using hierarchical clustering. The color gradient from red to blue reflects z-scores from −4 to 4. *The group of 40 patients with IS in stage 2 corresponded to the group of patients with IS in stage 3. †Metabolites with a false discovery rate-adjusted *p* value > 0.05, thus not meeting the criteria for statistical significance in the validation sample. 12-HETE = 12-Hydroxyeicosatetraenoic acid; A = amino acid; ADMA = asymmetrical dimethylarginine; adj.-*p* = *p* value adjusted for multiple testing; AUC = area under the curve; β = standardized β from the full model; CV = cofactor/vitamin; Disc = discovery sample; E = energy; GPC = glycerophospholipid; GPE = glycerophosphoethanolamine; KNN = *k*-nearest neighbors; L = lipid; LDA = linear discriminant analysis; LR = logistic regression; MC = metabolite class; N = nucleotide; NB = naïve Bayes; RF = random forest; SDMA = symmetrical dimethylarginine; SVM = support vector machines; U = unknown metabolite; Val = validation sample; X = xenobiotic; X-metabolites = structurally unnamed biochemicals, which have been identified by their recurrent chromatographic and mass spectral data.

asymmetrical dimethylarginine (ADMA) and symmetrical dimethylarginine (SDMA) in patients with IS compared with NCs, and thus considered both metabolites in stages 3, 4, and 5.

Search Strategy and Selection Criteria for the Systematic Review

We performed a comprehensive search of PubMed from database inception to July 10, 2020, using the following search inquiry: (*metabolomics* or *metabonomics* or *metabolome* or (*metabolic* and (*profiling* or *signature* or *biomarker* or *profile*)) or *metabolite*) and (*stroke* or (*cerebrovascular* and (*attack* or *accident* or *disease*)) or (*cerebral* and *ischemia*)). We considered publications for inclusion if they reported original data of circulating metabolite levels from patients with IS in comparison to NCs or SMs. We excluded studies with blood sampling beyond 24 hours of symptom onset. There was no language restriction.

Statistical Analysis

Metabolites were selected if detectable in at least 50% of the samples. After setting missing values as 50% of the

TABLE 1. Demographic and Clinical Characteristics of the Study Samples

	Stage 1		Stage 2 Stage 3			Stage 4		Stage 5	
Characteristics	**NC N = 72**	**IS 74**	**NC 40**	**IS 40**	**SM 33**	**IS 105**	**SM 105**	**IS 289**	**SM 211**
Demographic characteristics									
Age, mean, years	67	73[a]	72	73	71	71	71	74	74
Female, %	61	44[a]	40	38	73	47	47	44	51
Vascular risk factors, %									
Hypertension	46	77[a]	50	65	61	71	65	76	64[a]
Smoking history	42	57	35	45	39	50	44	30	33
Hypercholesterolemia	25	37	28	28	42	38	42	28	32
Obesity	14	19	13	13	14	15	16	15	19
Diabetes mellitus	3	20[a]	8	13	18	15	15	21	21
Previous TIA/stroke/MI	4	27[a]	10	30[a]	52	39	36	25	28
Family history[b]	13	15	13	13	18	8	13	15	16
ΔT, mean, hours	n/a	7.0	n/a	4.5	6.5	4.7	4.6	4.8	4.1
NIHSS[c], median	n/a	4	n/a	5	n/a	6	n/a	6	n/a
IV, mean, ml	n/a	17	n/a	16	n/a	30	n/a	27	n/a
Stroke etiology, %									
LAA	n/a	22	n/a	40	n/a	11	n/a	13	n/a
Cardioembolism	n/a	28	n/a	30	n/a	49	n/a	42	n/a
Small-artery occlusion	n/a	10	n/a	3	n/a	6	n/a	3	n/a
Other determined	n/a	4	n/a	0	n/a	7	n/a	5	n/a
Undetermined	n/a	37	n/a	28	n/a	29	n/a	37	n/a

[a] $p < 0.05$ for comparison of diagnostic groups in all 5 stages.
[b] Family history for cardiovascular disease.
[c] NIHSS upon admission.
ΔT = time from symptom onset until hospital arrival; IS = ischemic stroke; IV = infarct volume; LAA = large-artery atherosclerosis; n/a = not available; NC = neurologically normal control; NIHSS = National Institutes of Health Stroke Scale; MI = myocardial infarction; SM = stroke mimic; TIA = transient ischemic attack.

lowest observed value, ion counts were log-normalized and scaled. Metabolites were assessed using linear regression in a basic (age and sex) and full model (age, sex, hypertension, history of smoking, body mass index, major cardiovascular events, glucose levels, high-density lipoprotein cholesterol levels, and glomerular filtration rate). The *p* values were false discovery rate-adjusted using the 2-stage step-up procedure ($\alpha = 0.05$)[12] and considered statistically significant if $p < 0.05$ in both models. Receiver operating characteristics were constructed using random forest classification with 10,000 trees, linear discriminant analysis, logistic regression, *k*-nearest neighbors, naïve Bayes, and support vector machines.[13] For the training set, we used 5 repeats of 10-fold cross-validation. Conditional variable importance was determined by random forest regression across 200 runs. Statistical analyses were performed in R software, version 3.5.0.

Results

The sampling strategy and baseline characteristics of our study samples are presented in Figure 1A, Table 1, and Supplementary Table S1. CT findings and final diagnoses of patients from the discovery, derivation, and validation

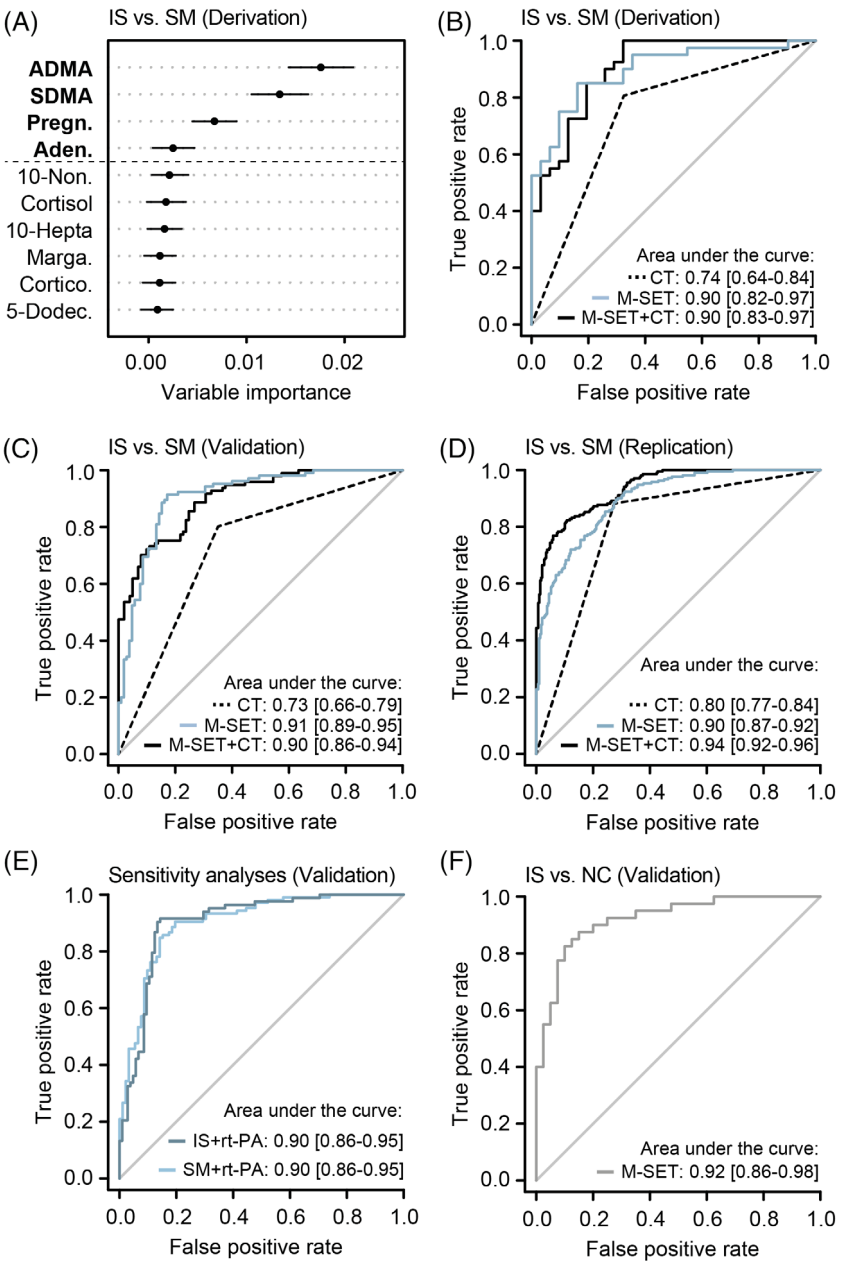


FIGURE 2: 2A set of 4 metabolites differentiates IS and SM. (A) Random forest variable importance analysis for the 30 validated endogenous metabolites to differentiate patients with IS from patients with SM (stage 3). Shown are the 10 metabolites that ranked highest. (B to D) ROC curves for differentiating patients with IS from patients with SMs in the derivation (stage 3; B), validation (stage 4; C), and replication sample (stage 5; D). (E) ROC curves for differentiating patients with IS from patients with SMs in the validation sample when excluding patients with SM that received rt-PA (N = 13) and when excluding patients with IS that were recruited after rt-PA administration (N = 22). Shown are the results for the combined set of 4 metabolites (M-SET: ADMA, SDMA, pregnenolone sulfate, and adenosine), multimodal CT, and the combination of both. (F) ROC curve for the comparison of IS versus NCs using the same set of metabolites in the validation sample (stage 2). 5-Dodec. = 5-Dodecenoate (12:1n7); 10-Hepta. = 10-Heptadecenoate (17:1n7); 10-Non. = 10-Nonadecenoate (19:1n9); ADMA = asymmetrical dimethylarginine; Cortico. = corticosterone; CT = computed tomography; IS = ischemic stroke; Marga. = margarate (17:0); NC = neurologically normal control; Pregn. = pregnenolone sulfate; ROC = receiver operating characteristic; SDMA = symmetrical dimethylarginine; SM = stroke mimic; rt-PA = recombinant tissue-type plasminogen activator.

sample are provided in Supplementary Table S2. Our sample of patients with IS covered a broad range of infarct volumes (0.035 to 385 ml) and was representative of the general stroke population admitted to our center with regard to age, sex, and stroke etiologies (Supplementary Table S3).

To establish a set of circulating metabolites that differentiates patients with IS from NCs and patients with

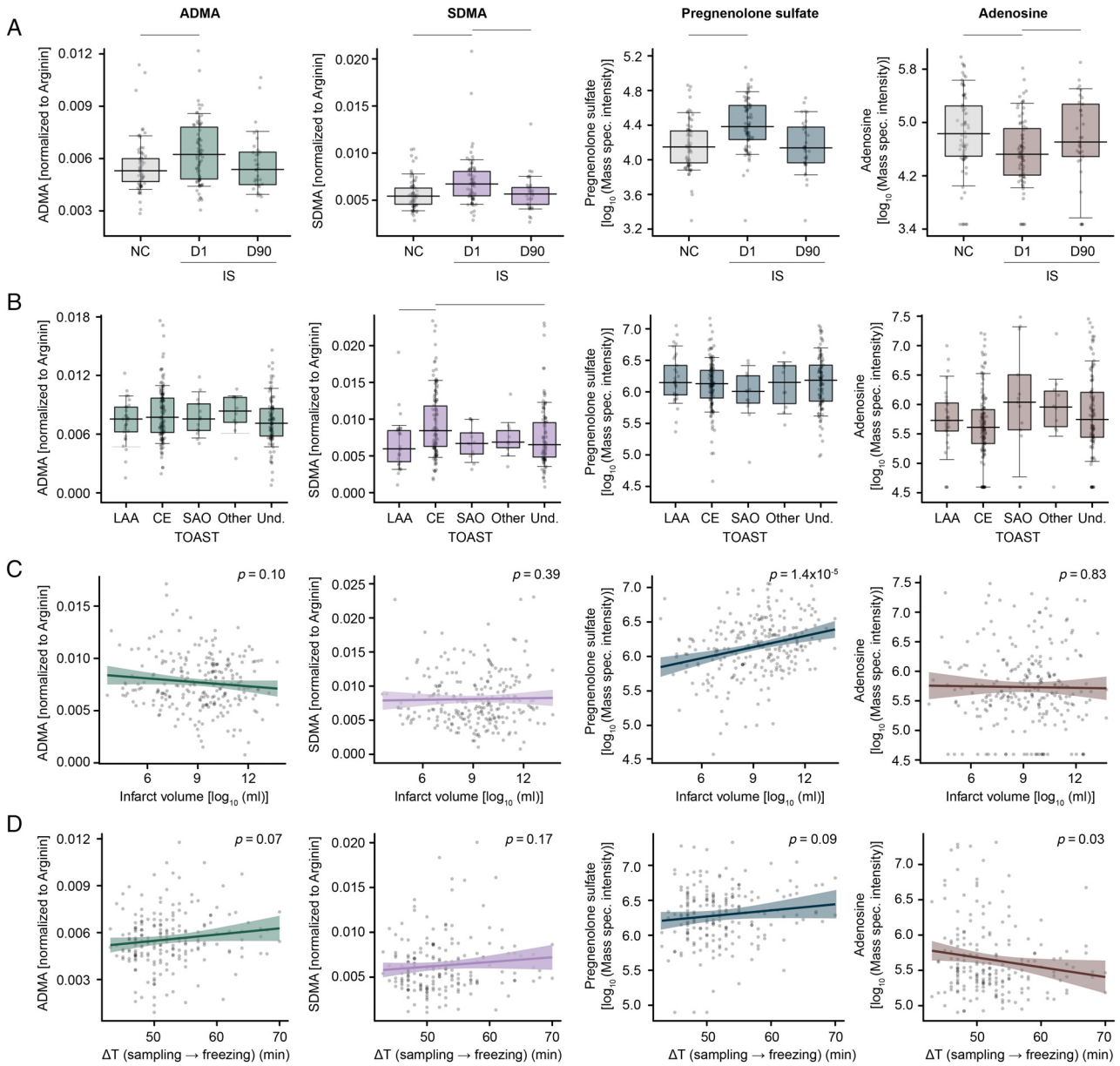


FIGURE 3: The 4 metabolites are linked to the acute phase of IS. (A) Metabolite levels of ADMA, SDMA, pregnenolone sulfate, and adenosine returned to control levels at day 90 post-stroke in the discovery sample (stage 1). (B) Levels of SDMA upon admission were higher in patients with cardioembolic stroke compared with patients with stroke due to LAA and undetermined etiology. (C) Levels of pregnenolone sulfate were positively associated with infarct volumes. (D) Adenosine levels negatively correlated with the processing time from blood sampling to freezing. There were 94.2% of samples that were processed between 40 and 70 minutes. A and B: Bars indicate $p < 0.05$. Kruskal–Wallis test followed by Dunn multiple comparison test. B and C: Data from stage 5. D: Data from stage 4. ΔT = time difference; ADMA = asymmetric dimethylarginine; CE = cardioembolism; LAA = large-artery atherosclerosis; mass spec. = mass spectrometry; min = minutes; other = other determined etiology; SAO = small-artery occlusion; SDMA = symmetric dimethylarginine; Und. = undetermined etiology.

SMs, we performed metabolomic profiling on serum samples obtained upon hospital arrival (IS: N = 508; SM: N = 349; median time from symptom onset = 3.3 hours; interquartile range [IQR]: 1.6–6.7 hours) and from NCs (N = 112; see Fig. 1A, Table 1). Among 689 measured metabolites in the discovery sample (stage 1), 180 metabolites were detected in less than 50% of the samples leaving 509 metabolites for analysis. Comparing patients with IS with NCs there were 49 metabolites that met the prespecified criteria for statistical significance in the discovery sample (stage 1; see Fig. 1A, B). Forty-one of them also met these criteria in the validation sample (stage 2), including 16 fatty acids, the steroids cortisol and pregnenolone sulfate, the tricarboxylic acid (TCA) cycle intermediates citrate and malate (all elevated), the amino acid proline, and the nucleosides adenosine, guanosine, and

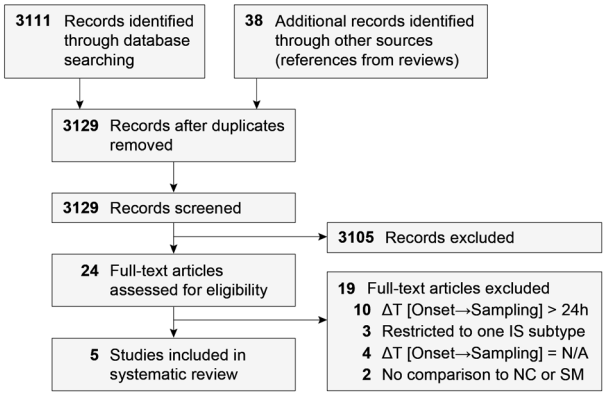


FIGURE 4: PRISMA Flow Diagram of Study Selection. Record identification, screening, eligibility review, and selection for this systematic review. ΔT = time difference; h = hours; IS = ischemic stroke; N/A = not available; NC = neurologically normal control; PRISMA = Preferred Reporting Items for Systematic Reviews and Meta-analyses; SM = stroke mimic.

inosine (all reduced; see Fig. 1B). Excluding xenobiotic and unknown metabolites, the 30 endogenous metabolites differentiated patients with IS from NCs with an area under the curve (AUC) of 0.961 (Fig. 1C) in a random forest classification. Alternative classification models showed similar AUC values (Fig. 1D). Hierarchical clustering showed considerable collinearity particularly among fatty acids (Fig. 1E), suggesting redundancies in the classifier. To identify the most informative metabolites for a reduced classifier that would distinguish IS from SMs while accounting for collinearity, we assessed conditional variable importance and found ADMA, SDMA, pregnenolone sulfate, and adenosine to rank highest (stage 3; Figs. 1A and 2A). These 4 metabolites showed a sensitivity of 82.5% and specificity of 84.9% resulting in a higher AUC compared to multimodal CT performed as part of

TABLE 2. Metabolite sets previously reported to be associated with IS within 24 hours from symptom onset and their diagnostic accuracy in our dataset

Study	Groups (N)	Platf. (Met.)	Metabolites associated with IS in original study	Diagnostic accuracy (stages) 1/2[a]	3	4	5
Hu et al	IS vs NC (129 vs 98)	T (49)	↑ Alanine, Citrulline, 3-OH-IVC[b]	0.63	0.54	0.57	0.51
Zhang et al	IS vs NC (129 vs 65)	UT	↑ Arginine, IVC, Ornithine, Aspartic acid ↓ Asparagine[b], AC, DEC, **Proline,** Tyrosine, Behenicarnitine[b], **Prop-C,** Valine, Carnitine, Butyryl-carnitine, Glycine[b], ODC[b], Tryptophan, Leucine, 3-OH-BC	0.80	0.51	0.64	0.61
Liu et al	IS vs NC (66 vs 63)	UT (1082)	↑ Betaine, LysoPE (18:2) ↓ Serine, Isoleucine, PC(5:0/5:0)[b]	0.62	0.60	0.51	0.53
Sun et al	IS vs HA or VT (38 vs 46)	T (58)	↑ VC[b], Arginine, Palm-C, 3-OH-BC	0.61	0.61	0.51	0.58
Jiang et al	IS vs NC (67 vs 62)	UT	↑ Cysteine[b], S-AHC[b], Ox-Glutathione[b], **HETE,** HODE ↓ Folic acid[b], THF[b], **Adenosine,** Aldosterone[b], DOC[b], S-6P[b], Betanin[b]	0.70	0.57	0.53	0.53

[a]Using the discovery sample (stage 1) as a training set and the validation sample (stage 2) as a test set for differentiating patients with IS from NCs.
[b]Metabolite not detected in the discovery sample.
[a]Bold text indicates metabolites meeting pre-specified criteria for statistical significance in the discovery and validation sample (stage 1 and 2).
3-OH-BC = 3-Hydroxylbutyrylcarnitine; 3-OH-IVC = 3-Hydroxyisovalerylcarnitine; AC = Acetylcarnitine; DEC = decanoylcarnitine; DOC = deoxocathasterone; HA = headache; HETE = hydroxyeicosatetraenoic acid; HODE = hydroxy octadecadienoic acid; IS = ischemic stroke; IVC = isovalerylcarnitine; LysoPE = lyso-phosphatidylethanolamine; Met. = number of metabolites; NC = neurologically normal control; ODC = octadecanoylcarnitine; Palm-C = palmitoylcarnitine; PC = phosphatidylcholine; Platf. = platform; Prop-C = propionylcarnitine; S-6P = Sucrose 6-phosphate; S-AHC = S-adenosyl-homocysteine; T = targeted; THF = tetrahydrofolate; UT = untargeted; VC = vaccenylcarnitine; VT = vertigo.

clinical diagnostic routine in the derivation sample (0.90 vs 0.74, stage 3; Fig. 2B). Similar results were obtained in the validation sample (stage 4, sensitivity: 92.4%, specificity: 82.9%; Fig. 2C), which was selected by propensity score matching to further control for potential confounders, such as demographics, vascular risk factors, and medication (N = 210; see Fig. 1A and Table 1), and in a consecutively recruited replication sample (stage 5, sensitivity: 88.6%, specificity: 72.7%, AUC = 0.90, N = 500; Fig. 2D). We further performed 2 sensitivity analyses: first, excluding patients with SMs who received recombinant tissue-type plasminogen activator (rt-PA) and might thus represent aborted strokes and second, excluding patients with IS with blood sampling after rt-PA administration. Both analyses showed similar results (Fig. 2E). The selected set of 4 metabolites also differentiated between patients with IS and NCs (AUC = 0.92; Fig. 2F).

Levels of the 4 diagnostic metabolites were largely normalized by day 90 (D90) after stroke (Fig. 3A) thus demonstrating a link between altered metabolite levels on day 1 (D1) and the acute event. Levels of SDMA upon admission (D1) were significantly higher in patients with cardioembolic stroke compared to patients with large-artery atherosclerotic stroke and stroke of undetermined etiology (Fig. 3B). Levels of pregnenolone sulfate significantly associated with infarct volume ($\beta = 0.66$, $p < 0.0001$; Fig. 3C) and the National Institutes of Health Stroke Scale scores upon admission ($\beta = 0.47$, $p = 0.001$). Assessing the influence of pre-analytical variability, we found adenosine levels to negatively correlate with pre-analytical sample processing time (Fig. 3D), which importantly did not differ between patients with IS and patients with SMs (51 vs 51 minutes, $p = 0.81$).

To relate our findings to metabolite sets previously reported to be associated with acute IS, we conducted a systematic review searching for studies that provided data on circulating metabolites in patients with IS and that were limited to blood sampling in the first 24 hours after symptom onset. Of the 3,149 records identified, 5 studies met the selection criteria (Fig. 4, Table 2, Supplementary Table S4).[14–18] All of them were single-center, hospital-based studies comparing patients with IS with either NCs (N = 4 studies) or patients with headache or vertigo (N = 1 study). The majority (N = 4 studies) lacked validation in independent samples. Collectively, these studies reported on 41 different metabolites associated with IS with only 1 metabolite overlapping between studies. Twenty-four of the 41 metabolites were represented in our dataset. A targeted look-up in the current dataset revealed 11 metabolites that met the pre-specified criteria for statistical significance in the basic model in the discovery sample (stage 1) and 4 metabolites (adenosine, proline, propionylcarnitine, and 12-Hydroxyeicosatetraenoic acid [HETE]) that also met the criteria in the full model, all with the same directionality. These 4 metabolites also met the criteria in the validation sample (see Supplementary Table S4). The AUC values obtained with the published metabolite sets to differentiate patients with IS and NCs in our discovery and validation samples (stage 1 and 2) ranged from 0.61 to 0.80 (see Table 2), which was considerably lower compared to the metabolite set identified in the current sample (0.93 and 0.96 in the discovery and validation sample, respectively; see Fig. 1C). Applying published metabolite sets to differentiate patients with IS and those with SMs in our dataset revealed AUC values below 0.7 (stages 3–5; see Table 2). Conversely, of the 30 abnormal endogenous metabolites identified and validated in the current study (stage 1 and 2), 5 (12-HETE,[16] adenosine,[16] proline,[17] oleoylcarnitine,[18] and propionylcarnitine[14,18]) were assessed in previous studies. Notably, 4 (80%) of them were abnormal and showed the same directionality in previous datasets (Supplementary Table S5).

Discussion

Improvements in diagnostic accuracy to discriminate patients with IS from patients with SM have been shown to translate into gains in patient safety,[19] clinical outcome,[19] and cost-effectiveness.[20] Multimodal CT represents the current diagnostic standard upon admission for detecting intracranial hemorrhage, vessel occlusion, and a tissue window for systemic thrombolysis and endovascular treatment,[21–23] but has limited utility for the distinction between IS and SM within the first 6 hours after symptom onset[2–4] with AUC values of approximately 0.8 as was also reflected by the AUC in the current study. Importantly, the discriminative power of ADMA, SDMA, pregnenolone sulfate, and adenosine was markedly higher in the present study and to our knowledge unprecedented for blood-based biomarkers.[24] Thus, our metabolite set could complement CT imaging in emergency settings where MRI is unavailable. Although untargeted metabolomics remain time- and cost-intensive, targeted metabolite panels now allow quantification within less than 10 minutes[25] and at the point of care[26,27] moving clinical applications into reach.

Notably, all 4 metabolites have been linked to stroke pathophysiology. ADMA and SDMA are endogenous inhibitors of nitric oxide synthase (NOS) and have been associated with vascular risk.[28] We found a return to control levels 90 days post-stroke, indicating an association with the acute event rather than marking a population at

risk. Elevated levels of ADMA and SDMA promote vasoconstriction through NOS inhibition, thus potentially promoting cerebral ischemia, although this has not been investigated in detail. Adenosine acts as a vasodilator, inhibits inflammation, and might be neuroprotective.[29] We found circulating adenosine levels to be reduced in the acute phase of IS, which conceivably could have unfavorable effects on stroke outcome, although this would need to be examined in larger patient samples. We further found a negative correlation with pre-analytical processing time, potentially reflecting an increased clearance, which has been described to occur ex vivo.[30] Pregnenolone sulfate is mostly synthesized in the adrenal glands, but also locally in the brain[31] and can freely penetrate the blood–brain barrier. Notably, we found a significant association between the extent of ischemic brain injury and the levels of pregnenolone sulfate upon admission compatible with its potential release from the brain. Intracranial injection of pregnenolone sulfate elicits seizures, which has been explained by its modulatory effect on $GABA_A$ and NMDA receptors.[32] Whether elevated levels of pregnenolone sulfate after IS contribute to the occurrence of seizures after stroke remains to be been investigated. Together, the 4 identified metabolites might reflect both local and systemic pathophysiological events, including processes within the vasculature, cerebral injury, and systemic responses (eg, stress and inflammation). Studies in experimental stroke models may provide further insights.

Compared to other metabolite sets previously reported to be associated with IS within 24 hours after symptom onset,[14–18] our set showed the highest accuracy for separating patients with IS from patients with SMs and NCs in our samples. Differences in diagnostic accuracies and in the sets of circulating metabolites may be routed in (1) different recruitment strategies resulting in sample differences with regard to age, sex, comorbidities, and medication, which are known to influence circulating metabolite levels[33]; (2) different metabolomic profiling approaches (untargeted vs targeted) and technologies (eg, nuclear magnetic resonance spectroscopy vs mass spectrometry); (3) small sample sizes resulting in insufficient statistical power to capture true differences of hundreds of measured metabolites; (4) lack of a validation step; and (5) lack of statistical adjustments for potential confounders and multiple testing. Among 11 metabolites that overlapped between previously published studies and our discovery analysis after adjusting for age and sex ("basic model," stage 1), only 4 also met the prespecified criteria for statistical significance after adjustment for additional confounders ("full model") and were validated in an independent sample (stage 2). Although full metabolite datasets from previous studies were unavailable, thus precluding precise determination of the overlap between studies, this observation highlights the requirement to consider confounders in metabolomic studies and to validate findings, including ours, in independent samples.

Strengths of our study include the discovery-validation approach with application of machine-learning techniques, the comparison with DWI MRI-confirmed SM as a clinically relevant differential diagnosis, the use of untargeted metabolomics with a leading platform[11] followed by targeted analyses, and a sample size at least 4 times the size of previous studies on the 24-hour time window. Limitations include the recruitment from a single center. In addition, our metabolomic assessment did not capture some metabolite subclasses, such as short-chain fatty acids. We cannot exclude that metabolite levels had already been altered before stroke onset. However, we consider this unlikely given the return to control levels 90 days post-stroke. Our study did not include patients with DWI-negative stroke, as we used imaging-defined stroke cases to establish the diagnostic accuracy of the metabolite set. Finally, we cannot exclude that the diagnostic signature identified here overlaps with the signature of other acute diseases, such as myocardial infarction. However, studies focusing on coronary artery disease and myocardial infarction have identified metabolic signatures[6,34] different to ours. Future studies may further explore whether the metabolite set identified here also discriminates between patients with ischemic and hemorrhagic stroke in the prehospital setting and performs equally well across specific subtypes of ischemic stroke.

In conclusion, a set of 4 metabolites shows high accuracy in differentiating patients with IS from patients with SMs, thus encouraging external validation and studies in multicenter settings to determine their potential in guiding clinical decision making.

Acknowledgments

Financial support

This study was supported by grants from the Deutsche Forschungsgemeinschaft (CRC 1123 [B3] and Munich Cluster for Systems Neurology [SyNergy]; EXC 2145 SyNergy – ID 390857198, M.D.), the German Federal Ministry of Education and Research (BMBF, e:Med program e:AtheroSysMed, M.D.), the FP7 / 2007-2103 European Union project CVgenes@target (grant agreement No. Health-F2-2013-601456, M.D.), the European Union Horizon2020 projects SVDs@target (grant agreement No. 66688, M.D.), and CoSTREAM (grant agreement No. 667375, M.D.), the Fondation Leducq (Transatlantic Network of Excellence on the Pathogenesis

of Small Vessel Disease of the Brain, M.D.), the Vascular Dementia Research Foundation, and the Jackstaedt Foundation. Support was also received by a grant from the Deutsche Forschungsgemeinschaft (Munich Cluster for Systems Neurology [SyNergy], S.T.). Support was also received by the European Union FP7 / HEALTH-2013-2.4.2-1/602936 (Project CarTarDis, S.B. and R.W.-S.). Dr. Werner Jackstädt-Stiftung (Grant Numbers: S134-10.065); ASCRS Research Foundation; Fondation Leducq (Transatlantic Network of Excellence on the Pathogenesis of Small Vessel Disease of the Brain); Deutsche Forschungsgemeinschaft and Munich Cluster for Systems Neurology [SyNergy]; EXC 2145 SyNergy – ID 390857198; Bundesministerium für Bildung und Forschung (Grant Numbers: e:Med program e:AtheroSysMed); H2020 Health (Grant Numbers: Health-F2-2013-601456); European Commission (Grant Numbers: FP7); Vascular Dementia Research Foundation. Open access funding enabled and organized by Projekt DEAL.

Author Contributions

S.T., C.G., and M. Dichgans were responsible for the conception and design of the study. S.T., S.B., H.K., M. Duering, A.A., J.A., M.K., T.L., L.M.H., D.T., R.W.-S., E.S., C.G., and M. Dichgans were responsible for acquisition and analysis of data. S.T. and M. Dichgans drafted a significant portion of the manuscript or figures.

Potential Conflict of Interest

The authors declared no conflict of interest.

References

1. Mortality GBD, Causes of Death Collaboators. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the global burden of disease study 2015. Lancet 2016;388:1459–1544.
2. Wardlaw JM, Mielke O. Early signs of brain infarction at CT: observer reliability and outcome after thrombolytic treatment--systematic review. Radiology 2005;235:444–453.
3. Hopyan J, Ciarallo A, Dowlatshahi D, et al. Certainty of stroke diagnosis: incremental benefit with CT perfusion over noncontrast CT and CT angiography. Radiology 2010;255:142–153.
4. Campbell BC, Weir L, Desmond PM, et al. CT perfusion improves diagnostic accuracy and confidence in acute ischaemic stroke. JNNP 2013;84:613–618.
5. Feibel JH, Hardy PM, Campbell RG, et al. Prognostic value of the stress response following stroke. JAMA 1977;238:1374–1376.
6. Fan Y, Li Y, Chen Y, et al. Comprehensive metabolomic characterization of coronary artery diseases. JACC 2016;68:1281–1293.
7. Sacco RL, Kasner SE, Broderick JP, et al.; American Heart Association Stroke Council, Council on Cardiovascular Surgery and Anesthesia., Council on Cardiovascular Radiology and Intervention., Council on Cardiovascular and Stroke Nursing., Council on Epidemiology and Prevention., Council on Peripheral Vascular Disease., Council on Nutrition, Physical Activity and MetabolismAn updated definition of stroke for the 21st century: a statement for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2013;44:2064–2089.
8. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ 2015;351:h5527.
9. Tiedt S, Duering M, Barro C, et al. Serum neurofilament light: a biomarker of neuroaxonal injury after ischemic stroke. Neurology 2018;91:e1338–e1347.
10. Tiedt S, Prestel M, Malik R, et al. RNA-Seq identifies circulating miR-125a-5p, miR-125b-5p, and miR-143-3p as potential biomarkers for acute ischemic stroke. Circ Res 2017;121:970–980.
11. Evans AM, DeHaven CD, Barrett T, et al. Integrated, nontargeted ultrahigh performance liquid chromatography/electrospray ionization tandem mass spectrometry platform for the identification and relative quantification of the small-molecule complement of biological systems. Anal Chem 2009;81:6656–6667.
12. Benjamini Y, Krieger AM, Yekutieli D. Adaptive linear step-up procedures that control the false discovery rate. Biometrika 2006;93:491–507.
13. Kuhn, M. Building Predictive Models in R Using the caret Package. J Stat Softw 2008;28:1-26.
14. Zhang X, Li Y, Liang Y, et al. Distinguishing intracerebral hemorrhage from acute cerebral infarction through metabolomics. Rev Invest Clin 2017;69:319–328.
15. Hu Z, Zhu Z, Cao Y, et al. Rapid and sensitive differentiating ischemic and hemorrhagic strokes by dried blood spot based direct injection mass spectrometry metabolomics analysis. J Clin Lab Anal 2016;30:823–830.
16. Jiang Z, Sun J, Liang Q, et al. A metabonomic approach applied to predict patients with cerebral infarction. Talanta 2011;84:298–304.
17. Liu P, Li R, Antonov AA, et al. Discovery of metabolite biomarkers for acute ischemic stroke progression. J Proteome Res 2017;16:773–779.
18. Sun R, Li Y, Cai M, et al. Discovery of a new biomarker pattern for differential diagnosis of acute ischemic stroke using targeted metabolomics. Front Neurol 2019;10:1011.
19. Schellinger PD, Thomalla G, Fiehler J, et al. MRI-based and CT-based thrombolytic therapy in acute stroke within and beyond established time windows: an analysis of 1210 patients. Stroke 2007;38:2640–2645.
20. Liberman AL, Choi HJ, French DD, Prabhakaran S. Is the cost-effectiveness of stroke thrombolysis affected by proportion of stroke mimics? Stroke 2019;50:463–468.
21. Hill MD, Goyal M, Demchuk AM, Fisher M. Ischemic stroke tissue-window in the new era of endovascular treatment. Stroke 2015;46:2332–2334.
22. Nogueira RG, Jadhav AP, Haussen DC, et al. Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct. N Engl J Med 2018;378:11–21.
23. Campbell BCV, Ma H, Ringleb PA, et al. Extending thrombolysis to 4.5-9 h and wake-up stroke using perfusion imaging: a systematic review and meta-analysis of individual patient data. Lancet 2019;394:139–147.
24. An SA, Kim J, Kim OJ, et al. Limited clinical value of multiple blood markers in the diagnosis of ischemic stroke. Clin Biochem 2013;46:710–715.
25. Kennedy J, Shanks KG, Van Natta K, et al. Rapid screening and identification of novel psychoactive substances using PaperSpray interfaced to high resolution mass spectrometry. Clin Mass Spectrom 2016;1:3–10.
26. Yu Q, Xue L, Hiblot J, et al. Semisynthetic sensor proteins enable metabolic assays at the point of care. Science 2018;361:1122–1126.

27. Yu Q, Pourmandi N, Xue L, et al. A biosensor for measuring NAD+ levels at the point of care. Nat Metab 2019;1:1219–1225.

28. Zobel EH, von Scholten BJ, Reinhard H, et al. Symmetric and asymmetric dimethylarginine as risk markers of cardiovascular disease, all-cause mortality and deterioration in kidney function in persons with type 2 diabetes and microalbuminuria. Cardiovasc Diabetol 2017; 16:88.

29. Chen JF, Lee CF, Chern Y. Adenosine receptor neurobiology: overview. Int Rev Neurobiol 2014;119:1–49.

30. Lofgren L, Pehrsson S, Hagglund G, et al. Accurate measurement of endogenous adenosine in human blood. PloS one 2018;13: e0205707.

31. Shibuya K, Takata N, Hojo Y, et al. Hippocampal cytochrome P450s synthesize brain neurosteroids which are paracrine neuromodulators of synaptic signal transduction. Biochim Biophys Acta 2003;1619: 301–316.

32. Harteneck C. Pregnenolone sulfate: from steroid metabolite to TRP channel ligand. Molecules 2013;18:12012–12028.

33. Beuchel C, Becker S, Dittrich J, et al. Clinical and lifestyle related factors influencing whole blood metabolite levels - a comparative analysis of three large cohorts. Mol Metab 2019;29:76–85.

34. Shah SH, Kraus WE, Newgard CB. Metabolomic profiling for the identification of novel biomarkers and mechanisms related to common cardiovascular diseases: form and function. Circulation 2012;126:1110–1120.