



Universiteit
Leiden
The Netherlands

Triage of stroke patients in the chain of acute stroke care

Nguyen, T.T.M.

Citation

Nguyen, T. T. M. (2022, May 25). *Triage of stroke patients in the chain of acute stroke care*. Retrieved from <https://hdl.handle.net/1887/3304543>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3304543>

Note: To cite this publication please use the final published version (if applicable).

Chapter 5

Circulating tRNA fragments as a novel biomarker class to distinguish acute stroke subtypes

International Journal of Molecular Sciences, 2021

T. Truc My Nguyen^{*}; M. Leontien van der Bent^{*}; Marieke J.H. Wermer; Ido R. van den Wijngaard; Erik W. van Zwet; Bas de Groot; Paul H.A. Quax; Nyika D. Kruyt[#]; A. Yaël Nossent[#]

^{*},[#] These authors contributed equally to this work

ABSTRACT

Background and purpose

Early blood biomarkers to diagnose acute stroke could drastically reduce treatment delays. We investigated whether circulating small non-coding RNAs can serve as biomarkers to distinguish between acute ischemic stroke (IS), intracerebral hemorrhage (ICH) and stroke mimics (SM).

Methods

In an ongoing observational cohort study, we performed small RNA-sequencing in plasma obtained from a discovery cohort of 26 patients (9 IS, 8 ICH and 9 SM) presented to the emergency department within 6 hours of symptom onset. We validated our results in an independent data set of 20 IS patients and 20 healthy controls.

Results

ICH plasma had the highest abundance of ribosomal and tRNA-derived fragments, while microRNAs were most abundant in plasma of IS patients. Combinations of four to five tRNAs yielded diagnostic accuracies (areas under the receiver operating characteristics curve) up to 0.986 (ICH vs. IS and SM) in the discovery cohort. Validation of the IS and SM models in the independent data set yielded diagnostic accuracies of 0.870 and 0.885 to distinguish IS from healthy controls.

Conclusion

We identified tRNA-derived fragments as a promising novel class of biomarkers to distinguish between acute IS, ICH and SM, as well as healthy controls.

INTRODUCTION

The diagnosis of acute stroke currently relies on clinical assessment and neuroimaging. While neuroimaging can rule out intracerebral hemorrhage (ICH), acute ischemic stroke (IS) can often not be confirmed and, therefore, diagnostic uncertainty can persist. Moreover, about 40% of suspected stroke patients presented to the emergency department as possible reperfusion candidates (so-called 'stroke codes') appear to have symptoms based on a different pathology than stroke: stroke mimics (SM).¹ It is critical to establish the diagnosis as quickly as possible, as the different types of stroke require completely different interventions. Particularly in acute IS speed is of the essence, as efficacy of reperfusion therapies sharply declines with time². An early biomarker that reliably distinguishes stroke subtypes (IS or ICH) and SMs would therefore be of substantial clinical value.

Circulating small non-coding RNAs (sncRNAs) potentially fit this profile and especially microRNAs have been studied intensively in this context.³ Recently, sncRNAs derived from transfer RNAs (tRNAs) have been described as an emerging biomarker class.⁴ tRNAs are encoded in the human genome codes by more than 400 genes, which can be divided into 48 *isodecoders*: tRNAs with the same anticodon. Although tRNAs are best known for their role in protein synthesis, tRNAs and tRNA-derived fragments (tRFs) also play a role in regulatory processes such as gene expression and translational control.⁵ Importantly, tRFs are found in liquid biopsies such as plasma and serum and show promise as biomarkers for several other pathologies, such as cancer and epilepsy.^{6,7}

We investigated whether sncRNAs in plasma, especially tRFs, can be used as early biomarkers to differentiate between acute IS, ICH and SM.

MATERIAL AND METHODS

Study design

This is a preplanned preliminary analysis of the MicroRNA in Acute Stroke (MIRAS) study, an ongoing prospective observational cohort study. Enrollment started on November 2018. This study is registered at AsPredicted.org with unique identifier #44134. Reporting was done in accordance with the STROBE Statement (see Supplementary Material).⁸ The study was conducted in accordance with the Declaration of Helsinki, the principles of Good Clinical Practice, the Dutch Agreement on Medical Treatment Act (WGBO) and the European General Data Protection Regulation and was approved by the Medical-Ethical Committee Leiden-Den Haag-Delft (P18.030; NL63060.058.17, 18 May 2018). Written informed consent was obtained from all patients or their legal guardian.

Patient selection

Consecutive stroke code patients aged 18 years or older who presented primarily at the emergency department of the Leiden University Medical Center within 6 h of symptom onset were included during office hours. Stroke was categorized according the American Heart Association/American Stroke Association.⁹ Stroke mimics (SM) were defined as stroke code patients for whom a stroke as the underlying pathology was excluded. Additional inclusion criteria for stroke patients were admittance to the neurology ward and a NIHSS score of ≥ 4 . Patients were excluded if they had an active malignant disease or used heparin within 24 h prior to symptom onset. In case of doubt about eligibility, blood samples were processed according to protocol. For all patients, the study coordinators (TTMN and MLvdB) verified that patients complied with the in- and exclusion criteria using the electronic patient record. Samples and clinical data of non-eligible patients or patients that did not provide informed consent were destroyed. For this discovery analysis, we selected 27 patients ($n = 8$ –10 per group) from a total of 68 eligible patients who were included between November 2018 and August 2019. Criteria for selection included technical quality of the sample (i.e., sufficient amount and absence of hemolysis) and a confirmed final diagnosis of IS, ICH or SM. One SM sample was excluded post hoc from our discovery cohort because of an uncertain onset-to-door time, leading to a final sample size of 26 (9 IS, 8 ICH and 9 SM). According to the standard stroke protocol, all patients underwent clinical assessment at presentation at the emergency department followed by non-contrast CT with CT-

angiography if indicated. Final diagnosis was established three months after discharge by an experienced stroke neurologist (NDK) based on all available clinical and neuro-imaging data and blinded to blood biomarker outcome.

Clinical data

Clinical data, neuro-imaging data and diagnosis at admission, discharge and after three months were retrieved from electronic patient records. Clinical data included demographic characteristics, medical history, medication use and stroke severity as assessed with the NIHSS score. In case the NIHSS score was not available, this was reconstructed from neurological examination at admission by NIHSS certified research members with a validated algorithm as described previously.¹⁰ In-hospital performance metrics included onset-to-door time. In case of a wake-up stroke, we included patients if the time of last seen well was within 6 hours.

Sample collection

Non-fasting venous blood was collected at the earliest possibility upon admission to the emergency department, prior to the administration of therapy or medication, using a 18G infusion needle in two 9 mL VACUETTE 9NC Coagulation sodium citrate tubes (Greiner Bio-One, Alphen aan den Rijn, the Netherlands). To prevent platelet activation and cell lysis, agitation of the blood collection tubes was kept to a minimum, tubes were kept vertical at all times after inverting carefully to mix the blood with the citrate buffer and they were delivered to the laboratory by one of the researchers, rather than through standard pneumatic tube delivery.¹¹ Blood was processed at room temperature within 60 min. Sodium citrate tubes were centrifuged for 15 min at 2500× g, after which plasma was transferred to a clean plastic 15 mL conical bottom centrifuge tube (Greiner) and centrifuged once more for 15 min at 2500× g. After each centrifugation step, plasma was pipetted off carefully from the top and at least 200 µL of plasma was left untouched in order to isolate platelet-poor plasma. Plasma aliquots were snap frozen in liquid nitrogen, then stored at -80 °C. Hemolysis was monitored visually and by spectrophotometry,¹² as well as by RT-qPCR of the erythrocyte-specific microRNA miR-451a relative to the stable miR-23a-3p, prior to sequencing.¹³

Ultracentrifugation and RNA-sequencing

To deplete plasma of extracellular vesicles (EVs), 2.5 mL of plasma was thawed and diluted once with phosphate-buffer saline (PBS). The plasma was then depleted of EVs by two rounds of ultracentrifugation

using a swinging-bucket rotor in a Beckman Coulter XE-90 instrument (Beckman Coulter, Woerden, The Netherlands). To remove large vesicles, samples were first centrifuged for 70 min at 4 °C at 17,500× g, using maximum acceleration (0) and slow deceleration (5). The upper 4.5 mL was transferred to a new tube and filled up to 5 mL with PBS, so that the final dilution of the plasma was approximately 2.2-fold. In the second step, samples were centrifuged for 70 min at 4 °C at 166,400× g to remove small vesicles, using the same acceleration and deceleration settings. Three fractions were separated: The top 0.5 mL, the middle 3.5 mL and the bottom 1 mL. These fractions, as well as the bottom 0.5 mL that remained after the first round of ultracentrifugation, were aliquoted and snap-frozen on dry ice, then stored at -80 °C. The upper 0.5 mL of EV-depleted plasma was sent to QIAGEN (Venlo, The Netherlands) for small RNA-sequencing (sRNA-seq). Briefly, RNA was isolated using the miRNeasy Serum/Plasma Kit (QIAGEN) according to the manufacturer's instructions. Five microliters of total RNA per sample were used for library preparation using the QIAseq miRNA Library Kit (QIAGEN). Adapters containing unique molecular identifiers (UMIs) were ligated to the RNA, followed by reverse transcription. cDNA was amplified by 22 cycles of PCR and indices were added during the PCR. PCR samples were purified, and libraries were pooled in equimolar ratios. These pools were quantified using qPCR and sequenced on a NextSeq500 instrument (Illumina, San Diego, CA, USA) according to the manufacturer's instructions. Raw data were de-multiplexed and FASTQ files were generated using the bcl2fastq software. Data quality was assessed using the FastQC tool.¹⁴

Publicly available sRNA-seq data of an independent study cohort were retrieved from the NCBI Sequence Read Archive (project identifier SRP133275)¹⁵. In this cohort, the authors included 20 IS patients (≤ 24 h of symptom onset; mean 3.9 h \pm 3.6 h) and 20 matched healthy controls. Briefly, the authors collected nonfasting blood samples at the emergency department in EDTA tubes using a 18G needle and tourniquet, after which platelet-poor plasma was isolated by two rounds of centrifugation (10 min at 2000× g followed by 15 min at 2500× g). RNA was isolated from the plasma using the miRCURY RNA Isolation Kit-Biofluids (Exiqon, Aarhus, Denmark); sRNA-seq libraries were prepared using the TruSeq Small RNA sample prepkit v2 (Illumina) and sequenced on an Illumina HighScan-SQ sequencer.

Data analysis

Clinical data were analyzed using SPSS (v24.0, IBM, New York, NY, USA). Continuous data were presented as mean \pm standard deviation (SD) if normally distributed, and as median with interquartile range (IQR) if skewed. Categorical data were presented as number and percentage.

Continuous data were compared using One-way ANOVA or Kruskal–Wallis tests, as appropriate. Descriptive categorical data were analyzed using Fisher's exact test.

Both sRNA-seq datasets were analyzed using the virtual machine of the sRNAtoolbox.¹⁶ Briefly, the sRNAbench module was used to (i) trim the Illumina adapter sequence from the raw reads in the fastq files, (ii) identify and deduplicate reads based on UMIs, and (iii) map the reads to the human genome using bowtie. Several libraries were used for annotation, including miRbase, gtRNAdb, RNACentral, and NCBI ncRNA and cDNA libraries. Then, the sRNAde module was used to investigate differential expression of various classes of small RNAs and variants. Data were visualized using Prism 8 (GraphPad) and R version 3.5.1.5.¹⁷

After mapping the sequencing reads, we calculated the levels of tRFs for each isodecoder by taking the sum of the multi-mapping adjusted total RPM (Reads Per Million, normalized to the total number of genome mapped reads). Isodecoder levels were then compared between the three groups by one-way ANOVA, followed by Tukey's post hoc tests for pairwise comparisons. Least absolute shrinkage and selection operator (LASSO) regression analysis was performed to select optimal statistical models for the prediction of pathology (IS, ICH or SM), using the glmnet package in R.¹⁸ Only isodecoders with a minimum expression of 1 read count in each sample and at least 10 RPM in at least 8 samples in our own dataset (resulting in 46 predictors) or at least 5 RPM in at least 20 samples in the independent dataset (resulting in 24 predictors) were considered in the LASSO analysis. Due to the small sample size per group in our own dataset, each group was compared to the other two groups using binomial analysis. The optimal model was defined as the model with the log (lambda) value that corresponded to the lowest binomial deviance. Generalized linear models were calculated for the optimal combination of predictors using the R stats package. Finally, using the pROC package in R, receiver operating characteristic (ROC) curves were generated to assess diagnostic accuracy for each separate predictor, as well as for the generalized linear models of the combinations of predictors.¹⁹

RESULTS

Patient characteristics

Our discovery cohort included 9 IS, 8 ICH and 9 SM patients, 58% of whom were men. The mean age was 71 ± 14 years, median onset-to-door-time was 78 min [IQR 57-116] and the median National Institutes of Health Stroke Scale (NIHSS) score of stroke patients was 8 [IQR 4-18] (Table 1). Patients with intracerebral hemorrhage more often had diabetes ($n = 4$; 50%) and hypertension ($n = 7$; 89%) compared to the ischemic stroke and stroke mimic group ($p < 0.05$).

Small RNAs in plasma

We detected RNAs from various sncRNA classes, including microRNA, ribosomal RNA and tRNA (Figure 1A). The distribution of the fraction of sequencing reads from these sncRNA classes differed between the groups. Notably, plasma of ICH patients contained the highest relative abundance of reads deriving from ribosomal RNAs and tRNAs (ICH: 22.89% and 1.65%; IS: 17.08% and 1.12%; SM: 18.39% and 1.46%, respectively), while microRNAs were relatively most abundant in IS plasma (IS: 44.13%; ICH: 32.52%; SM: 38.17%). tRFs showed the largest differences in abundance between the groups (ICH vs. IS: 1.47-fold; ICH vs. SM: 1.13-fold; SM vs. IS: 1.30-fold).

tRFs as biomarkers

In addition to overall differences in tRF abundance, we found that several isodecoders were differentially expressed between the groups ($p < 0.05$; Figure 1B and Supplementary Table S1). LASSO analysis yielded models consisting of four to five isodecoders. By combining these different isodecoders in a single model, diagnostic accuracy was markedly increased over each of the individual isodecoders, reaching an area under the ROC curve (AUC) of 0.986 (95% CI, 0.953–1.000) for ICH, 0.915 (95% CI, 0.809–1.000) for IS and 0.928 (95% CI, 0.832–1.000) for SM vs. the other two groups (Figure 1C).

Table 1.
Patient baseline characteristics of the discovery cohort

	Total No. Patients (n=26)	Ischemic Stroke (n= 9)	Intracerebral Hemorrhage[*] (n=8)	Stroke Mimics[†] (n=9)	p-value
Age, mean (SD), y	71 (14)	76 (8)	71 (14)	65 (16)	0.25
Male sex, n (%)	15 (58)	6 (67)	3 (38)	6 (67)	0.43
Medical history, n (%)					
Ischemic stroke/TIA	8 (31)	3 (33)	3 (38)	2 (22)	0.87
Intracerebral hemorrhage	2 (8)	0 (0)	2 (25)	0 (0)	0.09
Atrial fibrillation	5 (19)	1 (11)	2 (25)	2 (22)	0.84
Diabetes Mellitus	5 (19)	1 (11)	4 (50)	0 (0)	0.02
Hyperlipidemia	12 (46)	6 (67)	3 (38)	3 (33)	0.38
Hypertension	15 (58)	6 (67)	7 (89)	2 (22)	0.02
Medication use, n (%)					
Oral anticoagulation	5 (19)	1 (11)	3 (38)	1 (11)	0.38
Antiplatelets	8 (31)	3 (33)	2 (25)	3 (33)	1.00
Hospital admission					
ODT, median (IQR), min	78 (57-116)	72 (45-144)	95 (63-203)	68 (60-102)	0.76
NIHSS [‡] , median (IQR)	8 (4-18)	5 (4-19)	11 (6-18)	-	-
LAVO [§] , n (%)		4 (44)	-	-	-
Neuro-imaging					
Lesion confirmed, n (%)		8 (89)	8 (100)	-	
Reperfusion therapy					
IVT, n (%)	-	7 (78%)	-	-	-
EVT, n (%)	-	4 (44%)	-	-	-
DNT, median (IQR), min	-	20 (17-40)	-	-	-
DGT, median (IQR), min	-	56 (51-65)	-	-	-

TIA: transient ischemic attack; ODT: onset-to-door time; NIHSS: National Institute of Health Stroke Scale; LAVO: large anterior vessel occlusion; IVT: Intravenous thrombolysis; EVT: Endovascular treatment; DNT: door-to-needle time in IVT-treated patients; DGT: door-to-groin puncture time in EVT-treated patients.

* Location of hemorrhage: 4 (50%) basal ganglia, 4 (50%) lobar.

† Stroke mimics: 2 (22%) epilepsy, 5 (56%) acute vestibular syndrome (all with positive head impulse test), 1 (11%) intoxication, 1 (11%) partial oculomotor nerve palsy.

‡ Stroke only (i.e. ischemic stroke and intracerebral hemorrhage).

§ Occlusion of the intracranial carotid artery(-tandem) (ICA/ICA-T), middle cerebral artery (M1/M2), or anterior cerebral artery (A1/A2).

|| Computed tomography, -angiography/-perfusion, magnetic resonance imaging.

Validation of tRF models

We validated these models using data from an independent cohort of IS patients and healthy controls (hereafter referred to as the validation cohort).¹⁵ Our models to distinguish IS and SM yielded comparable AUCs of 0.870 (95% CI 0.756–0.984) and 0.885 (95% CI 0.781–0.989) (Figure 2A,B and Supplementary Table S2). The validation cohort did not include ICH patients and accordingly, our ICH model had much lower diagnostic accuracy in this dataset (AUC: 0.728; 95% CI 0.561–0.894; Figure 2C).

Common tRF model

When an optimal LASSO model to distinguish IS from healthy controls was determined in the validation cohort, three out of nine isodecoders overlapped with our IS and SM models: tRNA-TyrGTA, tRNA-ThrCGT and tRNA-ValCAC (Figure 3A). The combination of these three isodecoders yielded an AUC of 0.875 (95% CI, 0.759–0.991) to distinguish IS from healthy controls in the validation cohort (Figure 3B and Supplementary Table S2). This common tRF model showed poor diagnostic accuracy in our discovery cohort to distinguish between ICH and SM (AUC 0.653; 95% CI, 0.359–0.947), but high diagnostic accuracy for distinguishing between IS and ICH (AUC 0.847; 95% CI, 0.656–1.000) and between IS and SM (AUC 1.000) (Figure 3C and Supplementary Table S2).

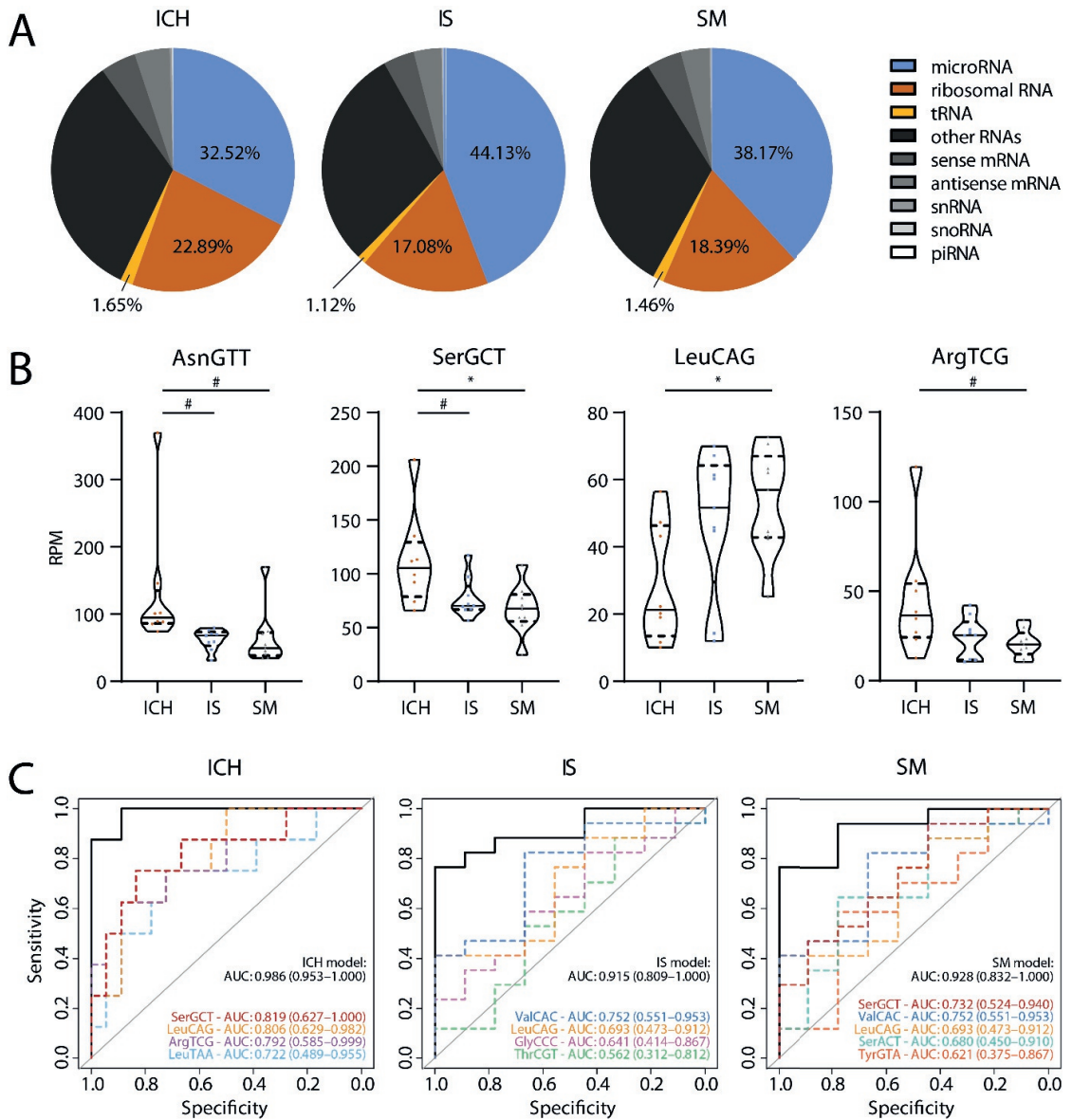


Figure 1.
sRNA-seq results in the discovery cohort

(A) Distribution of reads assigned to various small non-coding RNAs (sncRNA) classes. (B) Summed tRNA-derived fragments (tRF) abundance of differentially expressed isodecoders. Tukey's post-hoc test: # $p < 0.10$; * $p < 0.05$. (C) ROC analysis of diagnostic accuracy for each of the three groups from the other two groups. Dashed colored curves: Separate isodecoders in each optimal least absolute shrinkage and selection operator (LASSO) model. Black curves: Combination of these isodecoders. $n = 8-9$ per group. RPM: Reads per million; AUC: Area under the curve.

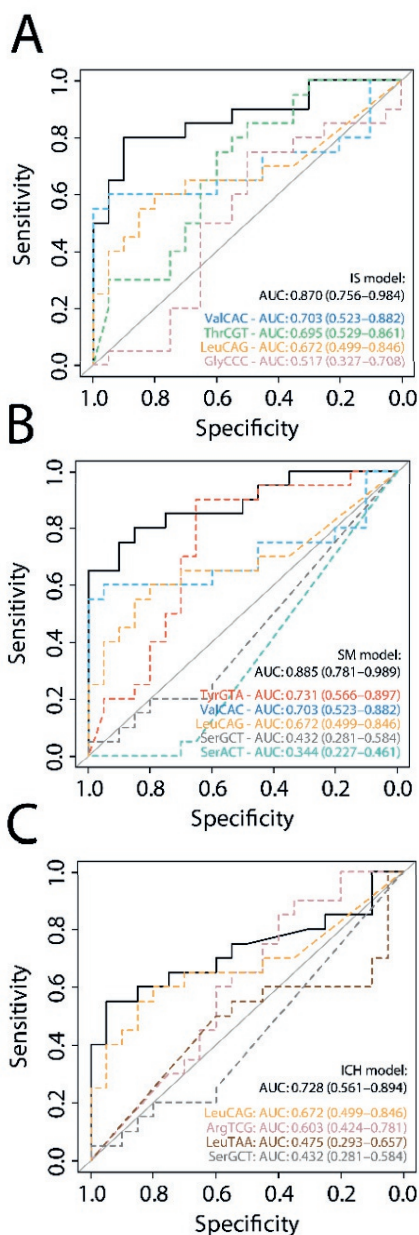


Figure 2.
 ROC analyses on the validation cohort

(A) Ischemic stroke (IS) model, (B) stroke mimic (SM) model and (C) intracerebral hemorrhage (ICH) model, as determined in our discovery cohort, applied to the sRNA-seq dataset from the validation cohort (IS vs. healthy controls; $n = 20$ per group). Dashed colored lines show the diagnostic potential of the separate isodecoders; black curves show the diagnostic potential of the combination of these. AUC: Area under the curve. HC: Healthy controls.

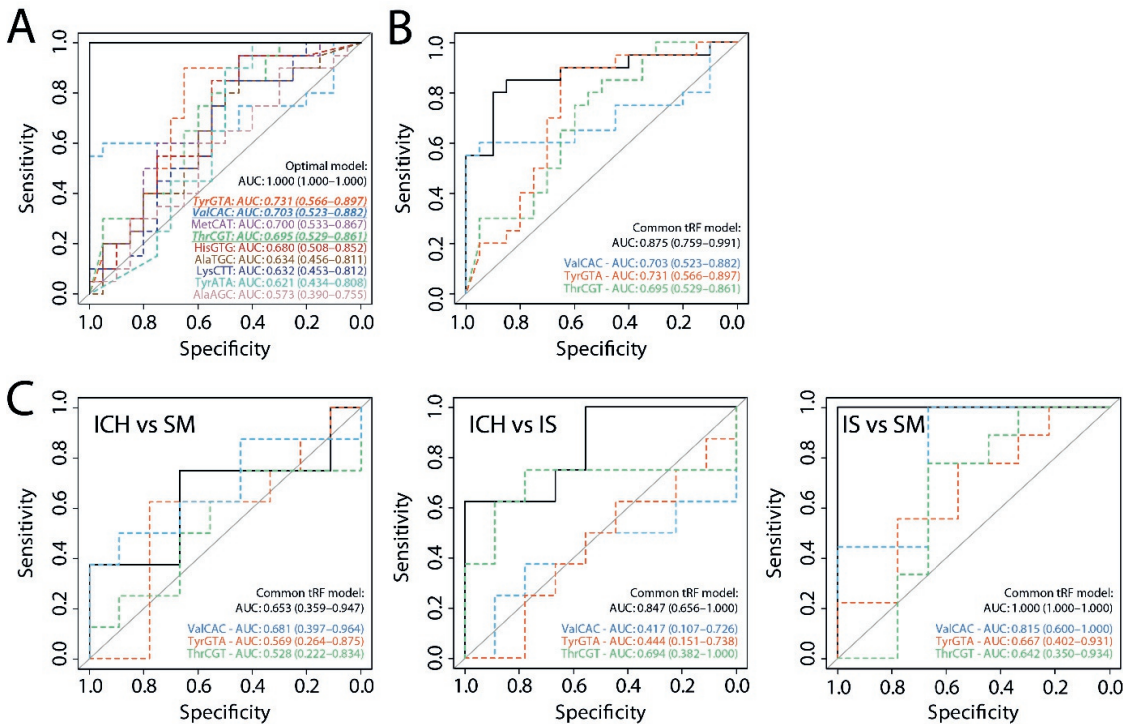


Figure 3.
Common tRF model

(A) Optimal LASSO model determined in the validation cohort; the three tRF isodecoders that were identified in both the discovery and the validation cohort are italicized and underlined. (B) ROC analysis of these three common tRFs in the validation cohort. (C) Common tRF model applied to the discovery cohort ($n = 8\text{--}9$ per group) to distinguish between ICH and SM (left), ICH and IS (middle), and IS and SM (right). Dashed colored lines show the diagnostic potential of the separate isodecoders; black curves show the diagnostic potential of the combination of these. AUC: Area under the curve. HC: Healthy controls.

DISCUSSION

We show that tRFs hold potential as biomarkers for the early diagnosis of acute stroke. Importantly, we found that tRF abundance from specific isodecoders can be used to (1) differentiate acute ICH from IS and SM, and (2) distinguish acute IS from SM and healthy controls. In all cases, combining different isodecoders in a single model yielded substantially higher diagnostic accuracies than individual isodecoders.

By combining the tRFs that were identified in both the discovery and the validation cohort, we formulated a model consisting of three tRFs that had good diagnostic potential for differentiating IS from healthy controls, SM and ICH. As there were no ICH patients in the validation cohort, it was not surprising that the common tRF model showed only poor diagnostic accuracy for the differentiation between ICH and SM. In fact, this confirmed the validity of the data obtained from the discovery cohort of the MIRAS study in distinguishing between IS and ICH.

To our knowledge, this is the first study on sncRNAs that includes all acute stroke codes, thereby reflecting clinical practice. Furthermore, we demonstrate the diagnostic potential of tRFs in stroke patients, which, to our knowledge, has only been reported in *in vitro* and *in vivo* models before.²⁰⁻²²

For stroke biomarkers to be of clinical use, rapid detection is crucial. A recent paper demonstrated that rapid detection of low concentrations of tRFs in plasma is indeed feasible.²³ We found that tRFs were also detectable in plasma depleted of extracellular vesicles, suggesting that circulating tRFs are mainly present outside vesicles. Therefore, we hypothesize that they can also be directly detected in whole blood, without needing to disrupt cellular or vesicular membranes. Importantly, these characteristics will absolve the need for labor- and time-intensive sample processing steps such as lysis, ultracentrifugation, RNA isolation and PCR, which is essential for point-of-care test development. Such a point-of-care test could potentially be used in the pre-hospital setting, for example in the ambulance or at the general practitioner. Thereby, the time to treatment could be drastically reduced, especially for ischemic strokes, which require reperfusion treatment as quickly as possible. Future studies need to confirm the feasibility, as well as the added diagnostic value of tRFs on top of clinical assessment (i.e., based on medical history, symptoms and neuroimaging), although the latter is known to be unreliable.²⁴

The prognostic potential of tRNA-derivatives in general has been shown by Ishida et al. in plasma collected from IS patients after seven days and has also been suggested in other diseases.^{4,25} Apart from their potential as diagnostic biomarkers, it will therefore also be highly

interesting to determine whether specific circulating tRFs have prognostic value for clinical outcome. Moreover, elucidating how tRFs are released into the circulation will be crucial to understand the pathophysiological link with stroke subtypes. Interestingly, a recent paper showed that tRNA-derived small RNAs in the brain may have a functional effect after experimental ICH induction in rats.²⁶

Our study is limited by the explorative design and small sample size. These limitations should be taken into account when interpreting the data. Due to the small sample size, we did not adjust for the fact that the groups were not completely matched for the vascular risk factors DM-II and hypertension, which may have confounded the results. Our results will therefore be validated in a larger cohort from the ongoing MIRAS study, allowing for multivariate analysis to adjust for vascular risk factors. Furthermore, the addition of tRFs as biomarkers to a clinical score could be investigated, for example using diagnostic measures such as the Net Reclassification Index or using likelihood-based methods.^{27,28} Finally, our results have not yet been validated using an independent detection method such as RT-qPCR. Nonetheless, the IS and SM models yielded comparable AUCs in the validation cohort. As expected, the ICH model had much lower predictive value in the validation cohort, as this cohort did not include ICH patients. These findings provide additional support for our results and underline the specificity of the models for each stroke subtype.

In conclusion, we identified tRFs as a promising novel biomarker class to distinguish between acute IS, ICH and SM. New detection technologies may facilitate swift translation into a point-of-care test that can eventually be implemented in clinical practice to reduce treatment delays and improve patient outcomes.

REFERENCES

-
1. Gioia LC, Zewude RT, Kate MP, et al. Prehospital systolic blood pressure is higher in acute stroke compared with stroke mimics. *Neurology*. Jun 7 2016;86(23):2146-53. doi:10.1212/wnl.0000000000002747
 2. Goyal M, Menon BK, van Zwam WH, et al. Endovascular thrombectomy after large-vessel ischaemic stroke: a meta-analysis of individual patient data from five randomised trials. *Lancet (London, England)*. Apr 23 2016;387(10029):1723-31. doi:10.1016/s0140-6736(16)00163-x
 3. Tiedt S, Dichgans M. Role of Non-Coding RNAs in Stroke. *Stroke*. Dec 2018;49(12):3098-3106. doi:10.1161/strokeaha.118.021010
 4. Jia Y, Tan W, Zhou Y. Transfer RNA-derived small RNAs: potential applications as novel biomarkers for disease diagnosis and prognosis. *Ann Transl Med*. Sep 2020;8(17):1092. doi:10.21037/atm-20-2797
 5. Li S, Xu Z, Sheng J. tRNA-Derived Small RNA: A Novel Regulatory Small Non-Coding RNA. *Genes (Basel)*. May 10 2018;9(5)doi:10.3390/genes9050246
 6. Dhahbi JM, Spindler SR, Atamna H, Boffelli D, Martin DI. Deep Sequencing of Serum Small RNAs Identifies Patterns of 5' tRNA Half and YRNA Fragment Expression Associated with Breast Cancer. *Biomark Cancer*. 2014;6:37-47. doi:10.4137/bic.S20764
 7. Hogg MC, Raof R, El Naggar H, et al. Elevation in plasma tRNA fragments precede seizures in human epilepsy. *J Clin Invest*. Apr 30 2019;129(7):2946-2951. doi:10.1172/jci126346
 8. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet (London, England)*. Oct 20 2007;370(9596):1453-7. doi:10.1016/s0140-6736(07)61602-x
 9. Powers WJ, Rabinstein AA, Ackerson T, et al. 2018 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*. Mar 2018;49(3):e46-e110. doi:10.1161/str.0000000000000158
 10. Williams LS, Yilmaz EY, Lopez-Yunez AM. Retrospective assessment of initial stroke severity with the NIH Stroke Scale. *Stroke*. Apr 2000;31(4):858-62. doi:10.1161/01.str.31.4.858
 11. Coumans FAW, Brisson AR, Buzas EI, et al. Methodological Guidelines to Study Extracellular Vesicles. *Circulation research*. May 12 2017;120(10):1632-1648. doi:10.1161/circresaha.117.309417

12. Appierto V, Callari M, Cavadini E, Morelli D, Daidone MG, Tiberio P. A lipemia-independent NanoDrop[®]-based score to identify hemolysis in plasma and serum samples. *Bioanalysis*. May 2014;6(9):1215-26. doi:10.4155/bio.13.344
13. Blondal T, Jensby Nielsen S, Baker A, et al. Assessing sample and miRNA profile quality in serum and plasma or other biofluids. *Methods*. Jan 2013;59(1):S1-6. doi:10.1016/j.ymeth.2012.09.015
14. S A. FastQC: A Quality Control Tool for High Throughput Sequence Data. Accessed 24 November 2020. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
15. 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. *Circulation research*. Sep 29 2017;121(8):970-980. doi:10.1161/circresaha.117.311572
16. Aparicio-Puerta E, Lebrón R, Rueda A, et al. sRNAbench and sRNAtoolbox 2019: intuitive fast small RNA profiling and differential expression. *Nucleic Acids Res*. Jul 2 2019;47(W1):W530-w535. doi:10.1093/nar/gkz415
17. Team RC. R: A Language and Environment for Statistical Computing; R Foundation for Statistical Computing. Accessed 24 November 2020, <http://www.R-project.org/>
18. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw*. 2010;33(1):1-22.
19. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. Mar 17 2011;12:77. doi:10.1186/1471-2105-12-77
20. Li Q, Hu B, Hu GW, et al. tRNA-Derived Small Non-Coding RNAs in Response to Ischemia Inhibit Angiogenesis. *Sci Rep*. Feb 11 2016;6:20850. doi:10.1038/srep20850
21. Elkordy A, Rashad S, Shehabeldeen H, et al. tiRNAs as a novel biomarker for cell damage assessment in in vitro ischemia-reperfusion model in rat neuronal PC12 cells. *Brain Res*. Jul 1 2019;1714:8-17. doi:10.1016/j.brainres.2019.02.019
22. Sato K, Rashad S, Niizuma K, Tominaga T. Stress Induced tRNA Halves (tiRNAs) as Biomarkers for Stroke and Stroke Therapy; Pre-clinical Study. *Neuroscience*. May 10 2020;434:44-54. doi:10.1016/j.neuroscience.2020.03.018
23. McArdle H, Hogg MC, Bauer S, et al. Quantification of tRNA fragments by electrochemical direct detection in small volume biofluid samples. *Sci Rep*. May 5 2020;10(1):7516. doi:10.1038/s41598-020-64485-4
24. Besson G, Robert C, Hommel M, Perret J. Is it clinically possible to distinguish nonhemorrhagic infarct from hemorrhagic stroke? *Stroke*. Jul 1995;26(7):1205-9. doi:10.1161/01.str.26.7.1205
25. Ishida T, Inoue T, Niizuma K, et al. Prediction of Functional Outcome in Patients with Acute Stroke by Measuring tRNA Derivatives. *Cerebrovasc Dis*. 2020;49(6):639-646. doi:10.1159/000511627
26. Li PF, Guo SC, Liu T, et al. Integrative analysis of transcriptomes highlights potential functions of transfer-RNA-derived small RNAs in experimental intracerebral hemorrhage. *Aging (Albany NY)*. Nov 16 2020;12(22):22794-22813. doi:10.18632/aging.103938

27. Burch PM, Glaab WE, Holder DJ, Phillips JA, Sauer JM, Walker EG. Net Reclassification Index and Integrated Discrimination Index Are Not Appropriate for Testing Whether a Biomarker Improves Predictive Performance. *Toxicol Sci.* Mar 1 2017;156(1):11-13. doi:10.1093/toxsci/kfw225
28. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med.* Jan 30 2008;27(2):157-72; discussion 207-12. doi:10.1002/sim.2929

SUPPLEMENTAL

SUPPLEMENTAL 1.**STROBE Statement. Checklist of items that should be included in reports of cohort studies**

	Item No	Recommendation	Page No
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	1
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	1
Introduction			
Background/ rationale	2	Explain the scientific background and rationale for the investigation being reported	2
Objectives	3	State specific objectives, including any prespecified hypotheses	2
Methods			
Study design	4	Present key elements of study design early in the paper	10
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	10-11
Participants	6	(a) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	10
		(b) For matched studies, give matching criteria and number of exposed and unexposed	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	11
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	11
Bias	9	Describe any efforts to address potential sources of bias	9

Study size	10	Explain how the study size was arrived at	10
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	13-15
Statistical methods	12	<p>Recommendation</p> <p>(a) Describe all statistical methods, including those used to control for confounding</p> <p>(b) Describe any methods used to examine subgroups and interactions</p> <p>(c) Explain how missing data were addressed</p> <p>(d) If applicable, explain how loss to follow-up was addressed</p> <p>(e) Describe any sensitivity analyses</p>	13-14
Results Participants	13*	<p>(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed</p> <p>(b) Give reasons for non-participation at each stage</p>	8
Descriptive data	14*	<p>(c) Consider use of a flow diagram</p> <p>(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders</p> <p>(b) Indicate number of participants with missing data for each variable of interest</p> <p>(c) Summarise follow-up time (eg, average and total amount)</p>	3
Outcome data	15*	Report numbers of outcome events or summary measures over time	3s

Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	3-7
		(b) Report category boundaries when continuous variables were categorized	
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done – eg analyses of subgroups and interactions, and sensitivity analyses	N.A
Discussion			
Key results	18	Summarise key results with reference to study objectives	8
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	9
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	8-9
Generalisability	21	Discuss the generalisability (external validity) of the study results	8
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	15

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at <http://www.strobe-statement.org>.

Table S1.

Sum of tRF levels for each isodecoder

Isodecoder	Mean	SD	Mean	SD	Mean	SD	p Value
SerGCT	112.2	44.0	77.4	18.7	67.1	23.0	0.0138
LeuCAG	28.8	17.6	47.5	21.3	53.5	15.6	0.0301
AsnGTT	132.0	98.5	62.2	15.3	64.0	42.3	0.0438
ArgTCG	45.2	33.1	24.4	11.3	21.0	7.6	0.0440
ValCAC	145.0	56.7	128.9	30.2	179.0	47.3	0.0800
ProAGG	77.6	83.3	30.1	9.8	32.9	18.7	0.0959
IleAAT	204.3	106.5	140.3	35.8	145.7	48.0	0.1302
LeuTAA	48.4	20.0	32.9	11.1	37.7	15.3	0.1385
ThrTGT	47.3	26.7	33.0	14.5	30.1	13.3	0.1567
GlyCCC	698.2	516.9	392.1	109.8	701.9	450.1	0.1894
HisGTG	49.4	21.3	53.5	27.3	68.7	22.1	0.2270
GlyTCC	220.4	434.8	41.2	19.0	44.3	24.1	0.2420
SerACT	7.1	4.9	9.9	8.8	13.6	8.9	0.2544
ProCGG	37.1	41.3	18.0	5.4	21.3	14.3	0.2551
GlyGCC	1016.6	744.5	535.6	239.0	1262.2	1376.3	0.2560
ArgTCT	162.1	23.9	134.9	46.3	136.6	52.0	0.3693
ValAAC	77.2	48.5	51.8	18.8	69.6	43.0	0.3877
TyrATA	6.4	3.2	8.4	3.5	6.7	2.8	0.3945
ArgCCG	21.3	12.9	15.4	6.4	18.1	7.3	0.4208
AlaAGC	156.7	55.4	136.2	22.5	136.3	26.1	0.4361
CysGCA	454.3	475.0	271.7	120.0	352.8	196.5	0.4566
AspGTC	932.5	325.4	783.0	108.1	888.2	286.6	0.4681
LeuAAG	192.8	165.6	231.8	122.2	272.2	101.9	0.4706
ThrCGT	27.2	14.4	20.8	6.4	24.7	10.5	0.4723
AlaTGC	152.5	68.1	168.4	39.8	185.9	60.0	0.4894
Undet???	6.2	3.5	8.2	3.7	7.7	3.8	0.5230
LeuTAG	74.2	48.2	79.1	42.0	95.8	36.9	0.5441
ProTGG	259.5	281.2	178.8	78.6	184.2	59.3	0.5487
TyrGTA	90.8	23.1	84.8	21.5	97.3	26.9	0.5525
LysCTT	464.4	436.5	343.7	158.2	342.3	57.2	0.5563
LysTTT	337.0	162.9	270.8	84.8	308.4	124.0	0.5620
SerCGA	61.4	51.6	65.7	47.7	45.6	17.3	0.5639
GluTTC	185.4	201.8	172.6	79.5	236.3	103.5	0.5829
SupTTA	18.5	10.4	14.1	3.4	17.0	11.0	0.5928
MetCAT	117.4	66.7	99.8	27.3	98.9	15.4	0.5930
TrpCCA	8.7	6.2	12.1	10.5	9.2	4.1	0.6025
GlnTTG	21.0	38.2	10.2	6.8	15.2	9.6	0.6130
SerAGA	96.2	37.5	102.4	21.7	90.7	26.3	0.6956
AlaCGC	51.8	44.8	41.8	8.1	50.4	18.1	0.7138
ValTAC	304.8	151.4	263.2	83.0	280.2	94.4	0.7465
GluCTC	278.0	110.0	278.1	115.2	249.1	44.7	0.7627
LeuCAA	59.1	36.1	50.6	14.6	55.8	19.6	0.7705
GlnCTG	11.6	8.5	9.7	5.8	11.3	4.6	0.8017
PheGAA	110.6	38.2	102.4	32.8	101.7	38.3	0.8605
ArgACG	20.7	27.2	16.5	10.1	16.8	10.6	0.8634
SerTGA	72.0	12.0	75.4	16.1	72.8	21.2	0.9083
SeC(e)TCA	31.9	23.1	31.1	23.6	34.1	21.3	0.9586
ArgCCT	28.5	10.6	29.1	7.9	28.0	8.0	0.9662

Mean and standard deviation of the sum of the multi-mapping adjusted total RPM. One-way ANOVA was used to test for differences between group means.

Table S2.

Diagnostic parameters determined by ROC analysis on generalized linear models of combinations of isodecoders

	Sensitivity	Specificity	AUC (95% CI)
Optimal LASSO models (discovery cohort; Figure 1C)			
ICH vs other groups (ICH model)*	1.000	0.889	0.986 (0.953-1.000)
IS vs other groups (IS model)†	0.889	0.941	0.967 (0.907-1.000)
SM vs other groups (SM model)‡	0.765	1.000	0.928 (0.832-1.000)
Previous models (validation cohort; Figure 2)			
ICH model*	0.550	0.950	0.728 (0.561-0.894)
IS model†	0.800	0.900	0.870 (0.756-0.984)
SM model‡	0.800	0.850	0.885 (0.781-0.989)
New models (validation cohort; Figure 3A and 3B)			
Optimal LASSO model§	1.000	1.000	1.000 (1.000-1.000)
Common tRF model	0.800	0.900	0.875 (0.759-0.991)
Common tRF model (discovery cohort; Figure 3C)			
ICH vs SM	0.750	0.667	0.653 (0.359-0.947)
ICH vs IS	0.625	1.000	0.847 (0.656-1.000)
IS vs SM	1.000	1.000	1.000 (1.000-1.000)

*LeuCAG, ArgTCG, LeuTAA and SerGCT; †ValCAC, ThrCGT, LeuCAG and GlyCCC; ‡TyrGTA, ValCAC, LeuCAG, SerGCT and SerACT; §TyrGTA, ValCAC, MetCAT, ThrCGT, HisGTG, AlaTGC, LysCTT, TyrATA and AlaAGC; ||ValCAC, TyrGTA and ThrCGT.

