

Longitudinal Metabolomics and High-Density EEG Signatures of One-Year Recurrent Ischemic Stroke: Exploratory Multimodal Classification and Internal Evaluation

Shiyang Liu

The Second Affiliated Hospital of Anhui University of Traditional Chinese Medicine

Meixiang Sun

The Second Affiliated Hospital of Anhui University of Traditional Chinese Medicine

Hui Liu

The Second Affiliated Hospital of Anhui University of Traditional Chinese Medicine

Yu Wang

The Second Affiliated Hospital of Anhui University of Traditional Chinese Medicine

Fang Zhang

The Second Affiliated Hospital of Anhui University of Traditional Chinese Medicine

Nan Li

The Second Affiliated Hospital of Anhui University of Traditional Chinese Medicine

Tingting Qian

The Second Affiliated Hospital of Anhui University of Traditional Chinese Medicine

Ling Zou

Yangzhou Hospital of TCM

Peiyang Sun

`sunpeiyang2008@aliyun.com`

The Second Affiliated Hospital of Anhui University of Traditional Chinese Medicine

Research Article

Keywords: recurrent ischemic stroke, longitudinal metabolomics, high-density electroencephalography, multimodal classification, Transformer, graph neural network, internal evaluation

Posted Date: October 6th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11131457/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Recurrent ischemic stroke remains a major source of preventable disability after an index event. Conventional risk assessment relies mainly on clinical variables measured at isolated time points and may not capture evolving systemic metabolism or brain-network organisation. We evaluated whether longitudinal metabolomic information and high-density electroencephalographic (HD-EEG) features could be integrated to distinguish patients with and without one-year recurrence and to determine how much of the observed signal was retained after model simplification.

Methods

This single-centre observational cohort included 428 adults with acute ischemic stroke enrolled between January 2020 and December 2022 and followed for one year; 51 participants had recurrent ischemic stroke and 377 remained recurrence-free. Longitudinal plasma UPLC-MS/MS metabolomic features and 128-channel resting-state HD-EEG microstate and functional-connectivity features were represented using a Transformer-based temporal encoder and a graph-based connectivity encoder. The source analysis used five-fold cross-validation with Bayesian hyperparameter optimisation. Because the archived fold-level and event-by-visit files were insufficient to verify fully nested preprocessing or the temporal eligibility of every longitudinal observation, the present report interprets performance as retrospective one-year recurrence-status classification rather than validated dynamic prediction.

Results

Participants with recurrence were older than recurrence-free participants (67.4 vs 63.2 years, $P = 0.008$) and had higher admission NIHSS scores (median 8 vs 5, $P < 0.001$). The full multimodal model yielded a reported ROC AUC of 0.91 and Brier score of 0.076; the source analysis also reported a 95% CI of 0.88–0.94 for the AUC, although the CI construction could not be independently reconstructed. A model without EEG yielded an AUC of 0.87 and a metabolomics-only model an AUC of 0.85, corresponding to nominal differences of 0.04 and 0.06 from the full model. At the observed recurrence prevalence (11.9%), a constant-risk reference has a Brier score of 0.105; thus the reported model Brier score represents a descriptive 27.6% reduction in mean squared probability error relative to that reference. This comparison does not substitute for calibration slope or intercept. Formal paired uncertainty estimates for between-model differences were unavailable.

Conclusions

Longitudinal metabolomic and HD-EEG features contained a strong internally evaluated signal for one-year recurrence-status classification, and the reduced-modality results suggest that metabolomics accounted for a substantial share of that signal. The nominal additional discrimination from HD-EEG warrants formal paired testing, but it should not yet be interpreted as confirmed incremental value. A participant-level, leakage-resistant nested reanalysis with auditable pre-event eligibility, full calibration and independent multicentre evaluation is required before this framework is described as prospective patient-level risk prediction.

Clinical trial registration

Chinese Clinical Trial Registry (ChiCTR2400079697) registered on 10 January 2024.

1. Introduction

Recurrent ischemic stroke remains one of the most consequential events after an initial cerebral infarction. A second event occurs in a patient who often already has residual neurological deficits, established vascular disease and substantial treatment burden; the resulting loss of function can therefore be disproportionate to the size of the recurrent lesion. Population studies consistently show that recurrence contributes materially to long-term mortality and disability despite improvements in antithrombotic therapy, lipid lowering, blood-pressure control and revascularisation strategies [1–3, 6, 7]. This residual risk creates a practical problem for secondary prevention: clinicians can identify many high-risk characteristics, but they still have limited tools for estimating how an individual patient's biological state evolves after discharge.

Existing clinical scores remain useful because they are inexpensive, transparent and immediately available. The Essen Stroke Risk Score has been evaluated for recurrent vascular risk after ischemic events, whereas ABCD² was designed for short-term stroke risk after transient ischemic attack [8, 9, 21, 25]. Their simplicity is an advantage, but it also limits the biological information they can represent. Age, vascular comorbidity and a small set of bedside variables are measured at a single time point; they do not directly measure the molecular consequences of tissue injury, treatment response, systemic inflammation or neural-network reorganisation. For a disease process that changes over weeks and months, a longitudinal biomarker strategy is therefore attractive, provided that the added complexity produces reproducible information beyond routine clinical data.

Metabolomics offers one route to that longitudinal information. The circulating metabolome reflects the combined effects of vascular disease, organ function, nutrition, medication exposure, inflammation and ongoing tissue repair. Metabolic signatures have been associated with cerebrovascular events, and prior work has shown that metabolomic or lipidomic profiles may contribute to recurrence-risk stratification after transient ischemic attack or minor cerebrovascular events [4, 5]. Repeated sampling has a particular conceptual advantage: it can distinguish a transient acute-phase perturbation from a persistent or progressively abnormal metabolic trajectory. That distinction may be important in recurrent

stroke, where the clinically relevant signal may reside not only in an abnormal value at baseline but also in failure to normalise during recovery.

HD-EEG provides a complementary view of the post-stroke patient. Resting-state EEG captures neuronal activity with millisecond temporal resolution, and microstate analysis summarises recurring global topographies of electrical activity. Phase-based connectivity measures such as PLV and wPLI describe coordinated activity between regions and can detect changes in distributed networks after stroke [11, 12]. These electrophysiological features are not substitutes for systemic vascular biomarkers. Their potential value lies in measuring a different biological scale: the organisation of large-scale neural networks during recovery. A multimodal model could therefore benefit if systemic biochemical trajectories and network-level electrophysiology contain partially non-redundant information.

Integrating these modalities is methodologically demanding. Longitudinal measurements are sequential, whereas functional-connectivity data are relational. Transformer architectures can represent temporal dependencies, and graph-based neural networks can represent structured relationships between nodes [13–17]. Similar strategies are increasingly used for multimodal clinical prediction, including multi-omics recurrence modelling [10]. At the same time, flexible models can produce overly optimistic estimates when the number of outcome events is small relative to the feature space or when preprocessing and hyperparameter selection are not fully isolated from held-out observations. Prediction-model reporting guidance therefore places equal emphasis on discrimination, calibration, transparent handling of missing data, internal evaluation and independent external evaluation [18, 26–30].

This study evaluated an exploratory model that combines longitudinal UPLC-MS/MS metabolomic information with 128-channel HD-EEG microstate and functional-connectivity features in 428 patients with acute ischemic stroke. The primary objective was to quantify internal discrimination for one-year recurrence status. Secondary objectives were to determine how much discrimination was retained after removing EEG information or restricting the model to metabolomics, and to place those estimates against simple clinical-score benchmarks. Because the accessible archive did not permit complete reconstruction of the temporal relationship between every longitudinal measurement and every recurrence event, we use the term classification for the present analysis and reserve prospective risk-prediction claims for a future pre-event, leakage-resistant validation. This distinction preserves the biological value of the longitudinal signal without overstating the temporal evidence available from the current dataset.

2. Methods

2.1 Study design and setting

This was a single-centre observational cohort study conducted at The Second Affiliated Hospital of Anhui University of Traditional Chinese Medicine. Consecutive adults with acute ischemic stroke were enrolled between January 2020 and December 2022 and followed for recurrent ischemic stroke within

one year. The study was subsequently registered with the Chinese Clinical Trial Registry (ChiCTR2400079697) on 10 January 2024. Because registration occurred after enrolment and outcome ascertainment, the modelling work is reported as retrospective and exploratory rather than as a prospectively prespecified prediction study.

The analysed cohort comprised 428 participants. During one-year follow-up, 51 participants were recorded as having recurrent ischemic stroke and 377 remained recurrence-free, corresponding to an observed recurrence proportion of 11.9%. The source modelling analysis treated recurrence as a fixed binary one-year endpoint. The present reconstruction therefore evaluates recurrence-status classification; it does not reinterpret the archived analysis as survival, landmark or continuously updated dynamic-risk modelling.

2.2 Clinical variables and outcome ascertainment

Baseline variables available for cohort description included age, sex, admission NIHSS score, TOAST etiological subtype, acute thrombolysis or neurointerventional treatment, and recanalisation status where applicable. These variables were used to characterise the study population and to provide clinical context for the biomarker analyses. Univariable comparisons were not used to make causal claims about recurrence.

The endpoint was recurrent ischemic stroke status recorded during one-year follow-up in the study database. The archived analysis identifies 51 recurrent events. The accessible manuscript archive did not contain a complete event-by-event adjudication package with recurrence dates, source-document review and an auditable link to every longitudinal sample. We therefore retain the recorded study endpoint without adding unverified adjudication details. Transient ischemic attack, haemorrhagic stroke and other vascular outcomes are not combined with the primary endpoint in this manuscript.

2.3 Longitudinal acquisition framework and temporal interpretation

The study programme scheduled molecular and electrophysiological assessments across four nominal study time points (T0–T3), spanning acute and follow-up phases (Fig. 1). Longitudinal UPLC-MS/MS metabolomics was the principal molecular modality and HD-EEG contributed microstate and functional-connectivity features. Repeated observations were intended to represent evolving patient state rather than a single cross-sectional snapshot.

Exact modality-completion counts at T0–T3 were not recoverable from the accessible analysis package, and the source-reported overall missingness range of approximately 5–11% did not include the denominators needed to derive visit-specific counts. We therefore do not impute or estimate those counts. The archive also did not preserve a complete event-by-visit map that could classify every longitudinal measurement as pre- or post-recurrence. Consequently, the present model is interpreted as retrospective recurrence-status classification. In a prospective risk-prediction analysis, only

measurements obtained before the prediction origin or recurrent event should be eligible, and all observations from a participant must remain within the same resampling unit.

2.4 Multimodal feature domains

The modelling framework combined three information domains (Fig. 2). Clinical data described the baseline patient context. Longitudinal metabolomic features represented systemic biochemical state over the scheduled follow-up. HD-EEG contributed microstate parameters and phase-based connectivity measures. The architecture was designed to preserve the different statistical structures of these inputs rather than concatenate all raw measurements into a single undifferentiated feature matrix.

2.5 UPLC-MS/MS metabolomic data

Plasma metabolomic profiling was performed using UPLC-MS/MS and generated repeated patient-level feature sets for model development. The source analysis used available longitudinal metabolomic information as a temporal input. The purpose was to retain within-patient change across the scheduled study period rather than reduce each participant to a single baseline vector.

High-dimensional metabolomic analyses are particularly sensitive to quality-control thresholds, missing-value rules, normalisation, transformation, scaling, batch handling and feature annotation. Complete vendor-level acquisition settings, peak-processing parameters, metabolite-identification confidence levels and the final post-filtering feature dimension were not preserved in the accessible analysis record. These details are therefore not replaced with generic parameters in the present manuscript. Their absence limits reproducibility and should be resolved from raw instrument and preprocessing records before a definitive methods supplement is submitted.

2.6 HD-EEG acquisition and feature extraction

The electrophysiological modality comprised 128-channel resting-state HD-EEG sampled at 2048 Hz. The recorded workflow included band-pass filtering and ICA-based artefact removal before feature derivation. Resting-state microstate parameters were extracted, with particular attention to microstate C, and functional-connectivity features included PLV and wPLI [11, 12]. These complementary measures were intended to summarise global electrical configurations and inter-regional phase relationships, respectively.

The EEG recordings used for recurrence modelling were offline resting-state acquisitions. They are therefore described as HD-EEG-derived biomarkers rather than brain-computer-interface outputs. This terminology matters because offline feature extraction does not demonstrate real-time decoding, adaptive stimulation or closed-loop control. Figure 3 summarises the EEG pathway used in the classification framework.

2.7 Missing data and feature handling

Missingness across the metabolomic and EEG modalities was reported in the source analysis to range from approximately 5% to 11%, and multiple imputation by chained equations was used under a missing-at-random assumption. The archived materials did not preserve the modality-by-time-point denominators or a fold-level imputation log; accordingly, the exact visit-specific missingness percentages and the number of imputed values are not reconstructed here.

For high-dimensional biomarker modelling, any operation that learns from the data—including imputation, normalisation, scaling, feature filtering, graph construction and dimensionality reduction—must be fitted on training participants only. The source documentation confirms five-fold cross-validation and Bayesian tuning but does not provide the fold assignments and preprocessing logs needed to prove that every learned operation was nested inside a participant-level outer loop. We therefore do not retrospectively label the original procedure as fully nested. The definitive reanalysis specification requires subject-level outer folds, inner-loop tuning and exactly one held-out prediction per eligible participant.

2.8 Multimodal model architecture

The exploratory architecture combined a Transformer-based temporal encoder with a graph-based connectivity encoder (Fig. 4). Longitudinal metabolomic and EEG-derived trajectories were represented by the temporal branch, allowing repeated measurements to contribute to a common patient-level representation. EEG functional-connectivity information was represented by the graph branch. The resulting representations were merged through a fusion component to estimate the probability of recurrence within one year.

The source workflow used Bayesian optimisation for hyperparameter tuning. Post hoc feature-attribution procedures were used during model exploration, but feature importance was interpreted as a description of model behaviour rather than biological causation. The manuscript does not claim that a feature receiving high model importance is a mechanistic driver or therapeutic target.

2.9 Internal evaluation and model comparisons

Five-fold cross-validation was used for internal evaluation in the source analysis, with Bayesian optimisation used for model tuning. ROC AUC was the principal discrimination measure and the Brier score quantified mean squared probabilistic error. Ablation analyses included the full multimodal architecture, a model with EEG information removed and a metabolomics-only model. The Essen Stroke Risk Score and ABCD² were retained as contextual clinical-score benchmarks rather than purpose-matched models.

The accessible source output did not include fold assignments, pooled out-of-fold probabilities, or a preprocessing log sufficient to verify a fully nested participant-level outer-validation procedure. It also did not provide calibration intercept, calibration slope, precision-recall AUC, or paired confidence intervals for the reduced-modality differences. The source analysis did report a 95% CI of 0.88–0.94 for the full-

model AUC; because the construction method cannot be independently reconstructed, that interval is identified as source-reported rather than newly estimated. Numerical AUC differences are treated as nominal ablation contrasts, not inferential evidence of incremental value.

2.10 Statistical analysis and reporting framework

Baseline cohort characteristics are presented as mean $\pm$ standard deviation, median with interquartile range, or number and percentage, as appropriate. The original cohort comparisons were retained to describe differences between recurrence groups. These P values are descriptive and are not interpreted as independent prognostic effects because the study objective is prediction rather than causal inference.

Prediction-model reporting was aligned where possible with TRIPOD + AI, while the observational cohort description was informed by STROBE [18, 19]. To extract the maximum defensible information from the verified aggregate results without fabricating patient-level outputs, two additional descriptive quantities were calculated. First, the Brier score of a constant predictor equal to the observed recurrence prevalence was calculated as $p(1 - p)$; with $p = 51/428$, this reference Brier score is 0.105. The descriptive Brier skill relative to that reference was calculated as $1 - (0.076/0.105) = 27.6\%$. Second, nominal AUC differences were calculated directly from the reported point estimates: +0.04 for the full model versus the model without EEG and +0.06 versus metabolomics alone. Neither calculation supplies missing calibration parameters or paired uncertainty. These derived quantities are used only to contextualise the archived performance estimates.

3. Results

3.1 Cohort characteristics

Among 428 analysed participants, 51 (11.9%) experienced recurrent ischemic stroke during one-year follow-up and 377 remained recurrence-free. Participants with recurrence were older (67.4 vs 63.2 years, $P = 0.008$) and had greater neurological deficit at admission, with a median NIHSS score of 8 compared with 5 in the recurrence-free group ($P < 0.001$). Sex distribution, TOAST subtype distribution and the reported use of recanalisation treatment did not differ significantly between groups (Table 1). These comparisons describe the observed cohort and are not interpreted as independent prognostic effects.

Table 1
Baseline characteristics of the study cohort.

Characteristic	Total (n = 428)	Non-recurrent (n = 377)	Recurrent (n = 51)	P
Age, years, mean $\pm$ SD	63.7 $\pm$ 10.4	63.2 $\pm$ 10.5	67.4 $\pm$ 9.2	0.008
Male, n (%)	264 (61.7%)	230 (61.0%)	34 (66.7%)	0.421
Admission NIHSS, median (IQR)	6 (3–9)	5 (3–8)	8 (5–12)	< 0.001
TOAST: large-artery atherosclerosis, n (%)	164 (38.3%)	142 (37.7%)	22 (43.1%)	0.354*
TOAST: cardioembolism, n (%)	116 (27.1%)	104 (27.6%)	12 (23.5%)	
TOAST: small-vessel occlusion, n (%)	94 (21.9%)	80 (21.2%)	14 (27.5%)	
TOAST: other/undetermined, n (%)	54 (12.7%)	51 (13.5%)	3 (5.9%)	
Neurointerventional/thrombolysis, n (%)	163 (38.1%)	145 (38.5%)	18 (35.3%)	0.652
Successful recanalisation (mTICI 2b–3), %†	84.6%	86.0%	75.0%	0.215

*P value refers to the overall TOAST subtype distribution. †Calculated within the mechanical-thrombectomy subgroup. SD, standard deviation; IQR, interquartile range; mTICI, modified Thrombolysis in Cerebral Infarction.

3.2 Data structure and modality contribution

The recurrence model combined longitudinal plasma metabolomic information with HD-EEG microstate and connectivity features. The modalities were intentionally non-redundant: metabolomics described systemic biochemical state, whereas HD-EEG represented large-scale neuronal organisation. The temporal encoder represented repeated observations, while the graph encoder represented connectivity structure. Table 2 summarises the information supplied by each domain. Because complete participant-by-visit modality counts were not available, no assumption of complete paired sampling at T0–T3 is made.

Table 2
Data domains, feature families and analytic roles.

Domain	Measurement	Feature family	Analytic role
Clinical context	Baseline assessment	Age, sex, NIHSS, TOAST subtype, recanalisation variables	Cohort description / comparator context
Metabolomics	Longitudinal plasma UPLC-MS/MS; T0–T3 scheduled	Repeated metabolomic feature trajectories	Principal molecular input
HD-EEG	128-channel resting-state EEG; 2048 Hz	Microstate parameters, PLV, wPLI	Electrophysiological / connectivity input
Outcome	One-year follow-up	Binary recurrent ischemic stroke status	Primary modelling target

3.3 Internal classification performance

The full multimodal model yielded a reported ROC AUC of 0.91 and Brier score of 0.076. The source analysis additionally reported a 95% CI of 0.88–0.94 for the AUC, but the archived materials were insufficient to reconstruct the exact confidence-interval procedure. We therefore retain the interval as a source-reported uncertainty estimate rather than presenting it as a newly reproduced result.

The cohort prevalence was 11.9%, giving a Brier score of 0.105 for a constant-risk predictor that assigns every participant the observed cohort risk. Relative to this reference, the reported Brier score of 0.076 corresponds to a 27.6% descriptive reduction in mean squared probability error. This provides useful context for the Brier score but does not demonstrate calibration-in-the-large or calibration slope. Those quantities require patient-level predicted probabilities and cannot be inferred from a single Brier score. With only 51 recurrent events supporting a high-dimensional model, the observed discrimination remains vulnerable to model flexibility and optimisation choices.

3.4 Reduced-modality analyses and clinical-score benchmarks

Removing EEG information reduced the reported ROC AUC from 0.91 to 0.87, a nominal difference of 0.04. A metabolomics-only model using the same temporal modelling concept retained an AUC of 0.85, a nominal difference of 0.06 from the full model and 0.02 from the non-EEG model. These contrasts show the direction and magnitude of the archived ablation results, but they are not paired Δ AUC estimates in the inferential sense because the necessary matched out-of-fold predictions and resampling uncertainty were not available.

The source analysis also reported an AUC of 0.68 for the Essen Stroke Risk Score and 0.65 for ABCD². The Essen score is a reasonable contextual benchmark for recurrent vascular risk. ABCD² has a different clinical purpose—it was designed for early stroke risk after transient ischemic attack—and is therefore shown only for context. A purpose-built clinical baseline constructed from the same cohort and

evaluated on exactly the same held-out participants would provide a more informative comparator for quantifying added biomarker value. The reported internal classification performance of the full model, reduced-modality models, and contextual clinical-score benchmarks is summarised in Fig. 5 and Table 3.

Table 3

Reported internal classification performance and contextual benchmarks. *The 95% CI for the full-model AUC was reported in the source analysis; the confidence-interval construction could not be independently reconstructed from the archived output. Nominal Δ AUC values are arithmetic differences between reported point estimates, not paired inferential comparisons.

Model / benchmark	ROC AUC	Brier score	Evidence-constrained interpretation
Full multimodal model	0.91 (reported 95% CI 0.88–0.94*)	0.076	Primary internal estimate; descriptive Brier skill 27.6% vs prevalence-only reference
Model without EEG	0.87	Not reported	Nominal Δ AUC – 0.04 vs full; paired uncertainty unavailable
Metabolomics-only model	0.85	Not reported	Nominal Δ AUC – 0.06 vs full; retains substantial discrimination
Essen Stroke Risk Score	0.68	Not reported	Contextual recurrent-risk benchmark
ABCD ² score	0.65	Not reported	Context only; different intended clinical use

3.5 Evidence-constrained interpretation of calibration and incremental value

The available aggregate results permit a stronger interpretation than AUC alone, but not the full set of analyses that would be obtained from locked out-of-fold probabilities. The Brier comparison indicates lower mean squared probability error than a prevalence-only reference, while the ablation point estimates indicate that the full model carries 0.04 more AUC than the non-EEG model and 0.06 more than metabolomics alone. Taken together, these findings support a potentially useful multimodal signal and a substantial molecular contribution. They do not establish calibration slope, calibration intercept, or statistically significant incremental value.

This distinction is clinically relevant. Repeated 128-channel HD-EEG requires specialised equipment, preparation time and signal-processing expertise. A reproducible paired gain of 0.04 could be clinically interesting if it also improved calibration or classification of patients near actionable risk thresholds; an unstable gain would favour a metabolomics-centred strategy. The next inferential step is therefore a matched comparison on identical held-out patients, not a stronger verbal claim based on the point estimates already available. The translational implications and unresolved validation requirements for the candidate modelling strategies are summarised in Table 4.

Table 4

Translational interpretation of candidate modelling strategies after evidence-constrained review.

Candidate strategy	Implementation burden	Reported AUC	Potential advantage	Unresolved requirement
Routine clinical model	Low	Not available as a matched model	Immediate availability and interpretability	Build and evaluate on the same outer folds
Metabolomics only	Moderate	0.85	Captures most reported discrimination with simpler workflow	Calibration, paired comparison, external evaluation
Model without EEG	Moderate–high	0.87	Retains non-EEG multimodal information	Define exact retained input set and paired uncertainty
Full metabolomics + HD-EEG	High	0.91	Highest reported internal discrimination	Demonstrate reproducible incremental value of EEG

4. Discussion

4.1 Principal findings

This study addresses a difficult but clinically important question: do repeated molecular measurements and brain-network features contain complementary information about one-year recurrent ischemic stroke status? In 428 patients, 51 of whom experienced recurrence, the full model yielded a reported ROC AUC of 0.91 and Brier score of 0.076. The source-reported AUC interval was 0.88–0.94. The Brier score was 27.6% lower than the error of a constant predictor fixed at the cohort recurrence prevalence, providing descriptive evidence that the model carried probability information beyond prevalence alone. Metabolomics alone retained an AUC of 0.85, while removing EEG yielded 0.87.

This pattern is more informative than simply highlighting the largest AUC. The nominal improvement from 0.87 to 0.91 after adding EEG is large enough to merit formal testing, but it is not itself a paired statistical result. The data therefore support two simultaneous conclusions: a strong internally evaluated multimodal classification signal is present, and a simpler molecular model retains much of that signal. The relative clinical value of those alternatives remains an empirical question for leakage-resistant, participant-level evaluation.

4.2 Biological rationale for longitudinal metabolomics

The metabolomic component is biologically plausible because the circulating metabolome sits close to the expressed phenotype of vascular disease. It reflects lipid handling, energy metabolism, inflammatory

activity, organ function, diet and pharmacological exposure. Stroke itself can perturb many of these processes, but their trajectories after the acute event are unlikely to be uniform. A patient whose metabolic state normalises during recovery may differ fundamentally from a patient with persistent dysregulation despite similar baseline clinical characteristics.

Previous work has linked metabolic and lipidomic patterns to recurrence or subsequent vascular events after transient ischemic attack and stroke [4, 5]. The contribution of the present framework is the emphasis on repeated measurements rather than a single snapshot. Longitudinal profiling can, in principle, separate acute injury markers from sustained susceptibility. That is a biologically appealing approach to recurrence because the outcome is not determined only at the moment of the index infarction; it emerges from the interaction of underlying vascular disease, recovery, treatment exposure and ongoing systemic stress over time.

The present dataset does not yet allow us to claim that a specific metabolite pathway drives recurrence. Complete feature-selection stability and post-filtering metabolite annotation were not available for secondary audit. The stronger conclusion is therefore methodological: repeated molecular information retained substantial predictive signal even when EEG was removed. A locked reanalysis should determine which metabolite features recur across resampling iterations and whether their selection is robust to reasonable preprocessing choices before mechanistic interpretation is emphasised.

4.3 What HD-EEG may add

HD-EEG addresses a different level of post-stroke biology. Cerebral ischemia disturbs distributed networks as well as focal tissue, and recovery is accompanied by large-scale reorganisation. Microstate analysis provides compact descriptions of recurring whole-brain electrical configurations, while PLV and wPLI quantify aspects of phase synchronisation and reduce the complex connectivity matrix to features that can be modelled [11, 12]. These measurements do not directly quantify vascular risk, but that is also their potential advantage: they may contain information that is not already present in blood-based molecular profiles.

The current ablation result is compatible with such complementarity. The nominal AUC increase from 0.87 without EEG to 0.91 in the full model is 0.04. That magnitude justifies a paired test on identical held-out participants, particularly if an EEG contribution improves calibration or discrimination among patients with intermediate clinical risk. However, a point-estimate difference observed during development can narrow or disappear when feature construction and tuning are fully isolated within an outer evaluation loop. We therefore describe the EEG contribution as a candidate increment rather than a confirmed one.

Clinical feasibility must be part of that comparison. Repeated 128-channel recordings are substantially more demanding than a blood sample, and the analytic pipeline adds signal-processing expertise that may not be available in every stroke unit. If HD-EEG produces a stable gain in discrimination or calibration, it may be valuable in specialised centres or in patients whose risk remains uncertain after

clinical and molecular assessment. If the gain is small or unstable, metabolomics may offer a more scalable route to translation.

4.4 Model complexity, calibration and comparison with clinical predictors

The structure of the data gives a legitimate reason to consider a Transformer and a graph-based encoder: repeated biomarker observations are sequential, whereas functional connectivity is relational. Nevertheless, model flexibility must be matched to the amount of outcome information. Only 51 recurrences were observed. Contemporary prediction-model guidance emphasises that sample adequacy depends on the number of candidate parameters, anticipated model fit and the precision required for predicted risks rather than on a simple events-per-variable rule alone [26]. In a high-dimensional neural architecture, this consideration is especially important because effective model complexity can be far larger than the number of clinically recognisable predictors.

Calibration is the second major issue. ROC AUC measures ranking, whereas clinical risk communication requires predicted probabilities to correspond reasonably to observed risk [27]. The reported Brier score of 0.076 is lower than the prevalence-only reference value of 0.105, equivalent to a descriptive Brier skill of 27.6%. That is encouraging, but Brier score combines several aspects of probabilistic performance and cannot identify systematic overprediction or underprediction. Calibration-in-the-large, calibration slope and a calibration curve must therefore be calculated from held-out patient-level probabilities before a prospective risk-prediction claim is made.

The clinical-score comparisons also need context. The Essen Stroke Risk Score is relevant to recurrent vascular risk, but ABCD² was developed for a different population and time horizon [8, 9, 21, 25]. More importantly, neither score is a substitute for a matched clinical model developed from the present cohort. The most persuasive assessment of added value would compare four prespecified models on identical outer folds: routine clinical predictors; clinical predictors plus metabolomics; clinical predictors plus HD-EEG; and the full multimodal combination. This design would answer the practical question that matters to clinicians—what does each biomarker layer add beyond information already available at the bedside?

4.5 Strengths

The study has several strengths that justify continued development. First, it combines systemic molecular information and neurophysiological data within the same patient-level prediction framework rather than treating them as unrelated exploratory analyses. Second, it uses repeated biomarker measurements, which is closer to the evolving biology of secondary prevention than a single cross-sectional assessment. Third, the reduced-modality analyses provide an initial view of parsimony: they show that much of the signal may be recoverable without the complete acquisition burden of the full model. Finally, the cohort is larger than many early-stage multimodal biomarker studies and contains a

clinically meaningful number of recurrent events, even though that event count remains modest for the complexity of the architecture.

4.6 Limitations

Several limitations define the evidentiary boundary of the present analysis. The study was single-centre and retrospectively registered. Although four longitudinal time points were scheduled, the accessible archive did not contain modality-specific T0–T3 completion denominators or a complete event-by-visit chronology, so post-event measurements cannot be ruled out for every recurrent participant and the analysis is not labelled validated dynamic prediction. The archived workflow also did not include the fold-assignment and preprocessing logs required to independently verify participant-level nested validation. Finally, calibration slope/intercept and matched out-of-fold predictions for formal Δ AUC testing were unavailable.

The event count was modest for a high-dimensional architecture, increasing the risk of unstable feature selection and optimistic performance. Source-reported missingness of approximately 5–11% was handled with MICE, but the visit-specific denominators and fold-level imputation provenance were not recoverable. Complete vendor-level metabolomic acquisition and processing parameters, final feature dimensions and a fully auditable recurrence-adjudication package should also accompany a definitive reproducibility supplement.

These gaps cannot be repaired responsibly by estimating plausible values. They can, however, be isolated so that the current manuscript does not depend on unverified claims: performance is presented as retrospective internal classification evidence; the Brier score is contextualised against a prevalence-only reference; ablation differences are reported only as nominal contrasts; and T0–T3 availability is described as scheduled rather than complete. This evidence-constrained framing preserves the reproducible core of the study while making the requirements for confirmatory analysis explicit.

4.7 Next-stage validation and clinical translation

A confirmatory analysis should be locked before examining new performance results. The cohort must first be reconstructed at participant level with exact recurrence dates, T0–T3 modality availability and an explicit flag indicating whether each longitudinal observation occurred before the event or prediction origin. All measurements occurring after recurrence must be excluded from any future-risk model. Participants—not visits—must define the resampling unit. Within each outer training fold, imputation, scaling, feature filtering, dimensionality reduction, graph construction and hyperparameter optimisation should be fitted without access to the outer test fold, after which exactly one out-of-fold probability should be generated for every eligible participant.

Performance should be reported from those pooled out-of-fold predictions using ROC AUC and precision-recall AUC with confidence intervals, Brier score, calibration intercept, calibration slope and a calibration plot. Model comparisons should be paired because the same participants contribute predictions from each model. A low-capacity comparator such as penalised logistic regression should

be included to determine whether the neural architecture provides a gain that justifies its complexity. Finally, the primary model should be locked and evaluated in an independent multicentre cohort without retuning. The sequence from development evidence to clinical translation is summarised in Fig. 6.

Importantly, a simpler result would not represent failure. If leakage-resistant reanalysis brings the full model close to the metabolomics-only model, the study would identify a more scalable biomarker strategy. If HD-EEG retains a stable and clinically meaningful increment, the case for selective multimodal assessment becomes stronger. The translational value of this programme therefore lies not in preserving a particular architecture, but in determining the minimum information required to produce reliable patient-level risk estimates.

5. Conclusions

In this single-centre cohort of 428 patients with acute ischemic stroke, longitudinal UPLC-MS/MS metabolomic information and HD-EEG microstate and connectivity features formed a coherent multimodal representation of one-year recurrence status. The full model showed strong reported internal discrimination (ROC AUC 0.91; source-reported 95% CI 0.88–0.94) and a Brier score of 0.076, which was 27.6% lower than a prevalence-only reference error. Reduced-modality estimates suggested that metabolomics retained a substantial proportion of the signal, while the nominal EEG-associated AUC increment of 0.04 remains to be tested with paired held-out predictions. The current evidence supports exploratory multimodal classification, not yet validated dynamic or prospective risk prediction. Participant-level nested evaluation with pre-event temporal eligibility, complete calibration and independent multicentre validation is required before clinical implementation.

Abbreviations

Abbreviation	Full term
ABCD ²	Age, Blood Pressure, Clinical Features, Duration, and Diabetes Score
HD-EEG	High-Density Electroencephalography
ICA	Independent Component Analysis
MICE	Multiple Imputation by Chained Equations
NIHSS	National Institutes of Health Stroke Scale
PLV	Phase-Locking Value
ROC AUC	Area Under the Receiver Operating Characteristic Curve
TOAST	Trial of Org 10172 in Acute Stroke Treatment
UPLC-MS/MS	Ultra-Performance Liquid Chromatography–Tandem Mass Spectrometry
wPLI	Weighted Phase-Lag Index

Declarations

Clinical trial registration: Chinese Clinical Trial Registry (ChiCTR2400079697); registered on 10 January 2024.

Human ethics approval: The study was approved by the Institutional Review Board of The Second Affiliated Hospital of Anhui University of Traditional Chinese Medicine. All procedures were conducted in accordance with the Declaration of Helsinki.

Consent to participate: Written informed consent was obtained from all participants within 72 hours of admission.

Consent for publication: Not applicable.

Availability of data and materials: The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests: The authors declare that they have no competing interests.

Funding: The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Author Contribution

S.L. contributed to the conception and design of the study, data acquisition, data analysis and interpretation, and drafted the manuscript. M.S., H.L., Y.W., F.Z., N.L., and T.Q. contributed to clinical data collection, patient follow-up, data curation, and interpretation of the results. L.Z. contributed to methodological support, data interpretation, and critical revision of the manuscript. P.S. supervised the study, contributed to study conception and design, interpretation of the results, and critically revised the manuscript for important intellectual content. All authors reviewed and approved the final version of the manuscript and agreed to be accountable for all aspects of the work. S.L. = Shiyang Liu, M.S. = Meixiang Sun, H.L. = Hui Liu, Y.W. = Yu Wang, F.Z. = Fang Zhang, N.L. = Nan Li, T.Q. = Tingting Qian, L.Z. = Ling Zou, P.S. = Peiyang Sun

References

1. Skajaa N, Adelborg K, Horváth-Puhó E, et al. Risks of stroke recurrence and mortality after first and recurrent strokes in Denmark: a nationwide registry study. *Neurology*. 2022;98(4):e329-e342. doi:10.1212/WNL.00000000000013118.
2. Chen Y, Wright N, Guo Y, et al. Mortality and recurrent vascular events after first incident stroke: a 9-year community-based study of 0.5 million Chinese adults. *Lancet Glob Health*. 2020;8(4):e580-e590. doi:10.1016/S2214-109X(20)30069-3.

3. Ke C, Bhatt DL, Austin PC, et al. Residual risk in patients with recurrent ischemic stroke: a nationwide cohort study. *Stroke*. 2023;54(9):2280-2289. doi:10.1161/STROKEAHA.123.042765.
4. Jové M, Mauri-Capdevila G, Suárez I, Cambray S, Sanahuja J, Quílez A, et al. Metabolomics predicts stroke recurrence after transient ischemic attack. *Neurology*. 2015;84(1):36-45. doi:10.1212/WNL.0000000000001093.
5. Purroy F, et al. Lipidomic signature of stroke recurrence after transient ischemic attack. *Sci Rep*. 2023. doi:10.1038/s41598-023-40838-7.
6. Lin B, Zhang Z, Mei Y, et al. Cumulative risk of stroke recurrence over the last 10 years: a systematic review and meta-analysis. *Neurol Sci*. 2021;42(3):889-898. doi:10.1007/s10072-020-04797-5.
7. Xia X, Li Y, Bai Y, et al. Recent advances in the prevention of secondary ischemic stroke. *Brain Circ*. 2024;10(4):283-295. doi:10.4103/bc.bc_159_24.
8. Weimar C, Diener HC, Alberts MJ, Steg PG, Bhatt DL, Wilson PWF, Mas JL, Röther J. The Essen Stroke Risk Score predicts recurrent cardiovascular events: a validation within the REduction of Atherothrombosis for Continued Health (REACH) registry. *Stroke*. 2009;40(2):350-354.
9. Johnston SC, Rothwell PM, Nguyen-Huynh MN, Giles MF, Elkins JS, Bernstein AL, Sidney S. Validation and refinement of scores to predict very early stroke risk after transient ischaemic attack. *Lancet*. 2007;369(9558):283-292. doi:10.1016/S0140-6736(07)60150-0.
10. Miao R, Li S, Fan D, et al. An integrated multi-omics prediction model for stroke recurrence based on Lnet transformer layer and dynamic weighting mechanism. *Comput Biol Med*. 2024;179:108823. doi:10.1016/j.compbimed.2024.108823.
11. Hao W, Zhai S, Cheng M, Pan C, Dou M. EEG microstate-specific functional connectivity and stroke-related alterations in brain dynamics. *Front Neurosci*. 2022;16:848737. doi:10.3389/fnins.2022.848737.
12. Michel CM, Koenig T. EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: a review. *Neuroimage*. 2018;180:577-593. doi:10.1016/j.neuroimage.2017.11.062.
13. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. *Adv Neural Inf Process Syst*. 2017;30:5998-6008.
14. Veličković P, Cucurull G, Casanova A, et al. Graph attention networks. *International Conference on Learning Representations*. 2018. doi:10.48550/arXiv.1710.10903.
15. Li MM, Huang K, Zitnik M. Graph representation learning in biomedicine and healthcare. *Nat Biomed Eng*. 2022;6(12):1353-1369. doi:10.1038/s41551-022-00942-x.
16. D'Souza NS, Wang H, Giovannini A, et al. Fusing modalities by multiplexed graph neural networks for outcome prediction from medical data and beyond. *Med Image Anal*. 2024;93:103064. doi:10.1016/j.media.2023.103064.
17. Zhou HY, Yu Y, Wang C, et al. A transformer-based representation-learning model with unified processing of multimodal input for clinical diagnostics. *Nat Biomed Eng*. 2023;7(6):743-755. doi:10.1038/s41551-022-00984-1.

18. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*. 2024;385:e078378. doi:10.1136/bmj-2023-078378.
19. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet*. 2007;370:1453-1457. doi:10.1016/S0140-6736(07)61602-X.
20. Adams HP Jr, Bendixen BH, Kappelle LJ, et al. Classification of subtype of acute ischemic stroke: definitions for use in a multicenter clinical trial. *Stroke*. 1993;24:35-41.
21. Liu Y, Wang Y, Li WA, Yan A, Wang Y. Validation of the Essen Stroke Risk Score in different subtypes of ischemic stroke. *Neurol Res*. 2017;39(6):504-508. doi:10.1080/01616412.2017.1313364.
22. Abedi V, Avula V, Chaudhary D, Shahjouei S, Khan A, Griessenauer CJ, Li J, Zand R. Prediction of long-term stroke recurrence using machine learning models. *J Clin Med*. 2021;10(6):1286. doi:10.3390/jcm10061286.
23. Purroy F, et al. How predictors and patterns of stroke recurrence after a TIA differ during the first year of follow-up. *J Neurol*. 2014;261:1614-1621. doi:10.1007/s00415-014-7390-z.
24. Ois A, Gomis M, Rodríguez-Campello A, et al. Factors associated with a high risk of recurrence in patients with transient ischemic attack or minor stroke. *Stroke*. 2008;39:1717-1721. doi:10.1161/STROKEAHA.107.505438.
25. Sumi S, Origasa H, Houkin K, et al. A modified Essen stroke risk score for predicting recurrent cardiovascular events: development and validation. *Int J Stroke*. 2013;8(4):251-257. doi:10.1111/j.1747-4949.2012.00841.x.
26. Riley RD, Snell KIE, Ensor J, Burke DL, Harrell FE Jr, Moons KGM, Collins GS. Minimum sample size for developing a multivariable prediction model: Part II - binary and time-to-event outcomes. *Stat Med*. 2019;38(7):1276-1296. doi:10.1002/sim.7992.
27. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW. Calibration: the Achilles heel of predictive analytics. *BMC Med*. 2019;17:230. doi:10.1186/s12916-019-1466-7.
28. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med*. 2019;170(1):51-58. doi:10.7326/M18-1376.
29. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J*. 2014;35(29):1925-1931. doi:10.1093/eurheartj/ehu207.
30. Moons KGM, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*. 2015;162(1):W1-W73. doi:10.7326/M14-0698.
31. Moons KGM, Damen JAA, Kaul T, et al. PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. *BMJ*. 2025;388:e082505. doi:10.1136/bmj-2024-082505.
32. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction

Figures

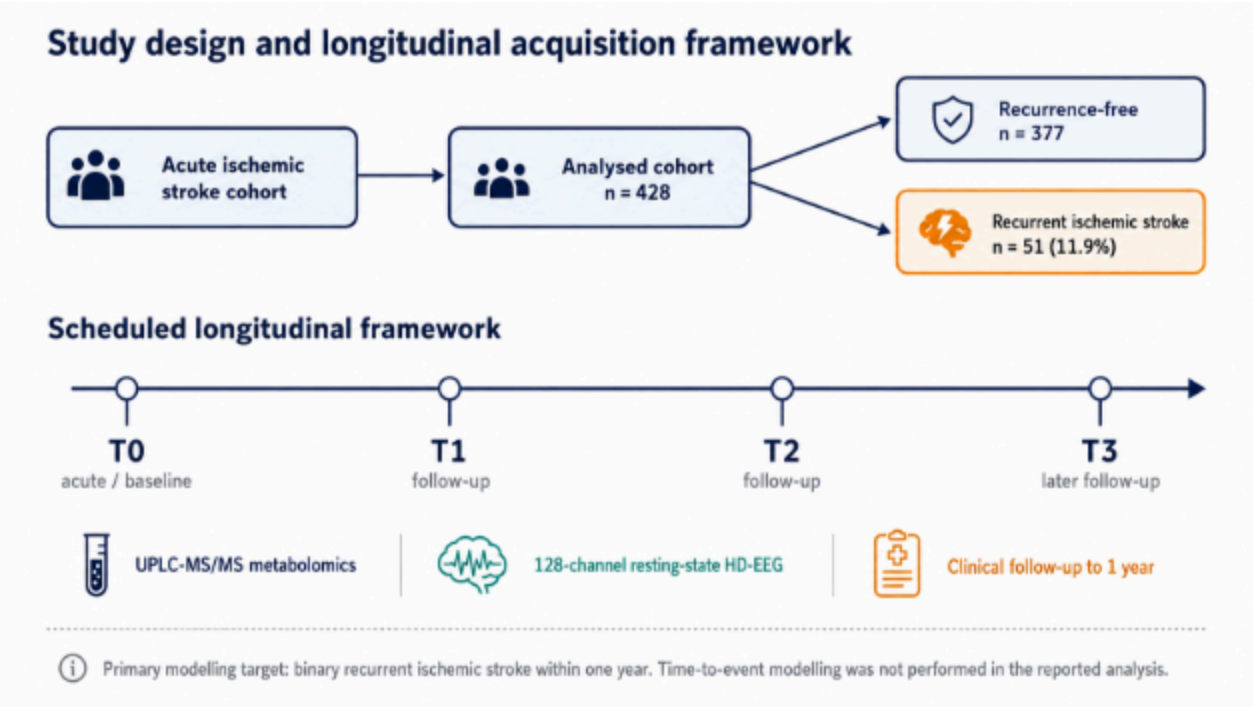
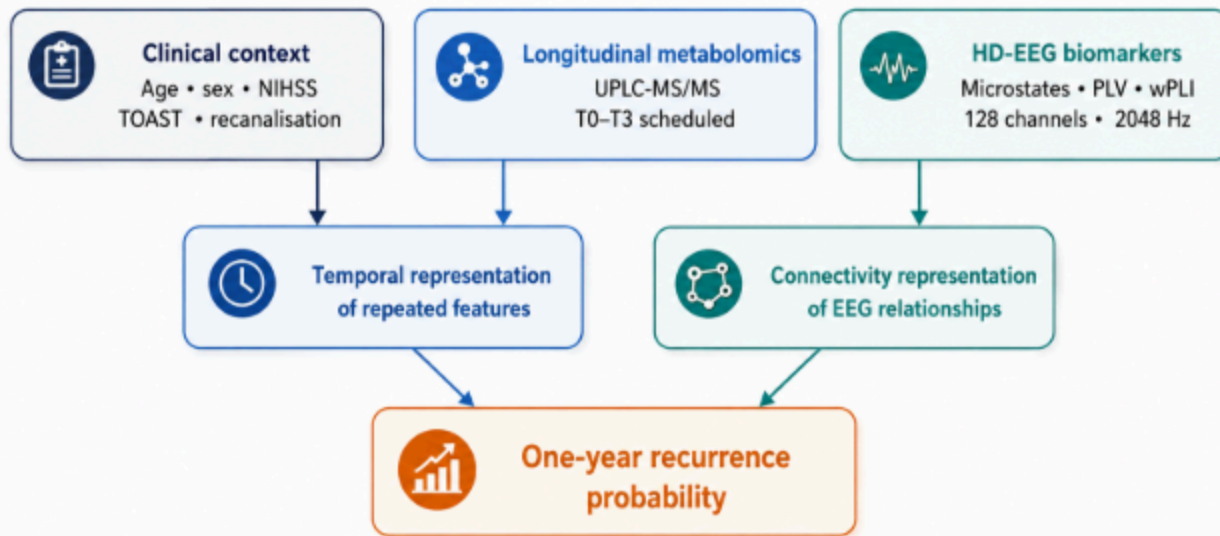


Figure 1

Study design, scheduled longitudinal acquisition and one-year recurrence-status structure. T0–T3 indicate nominal study time points. The figure shows the intended acquisition framework and known cohort totals; it does not imply complete paired metabolomic and HD-EEG sampling at every visit or that every observation preceded recurrence.

Multimodal feature domains and analytic roles

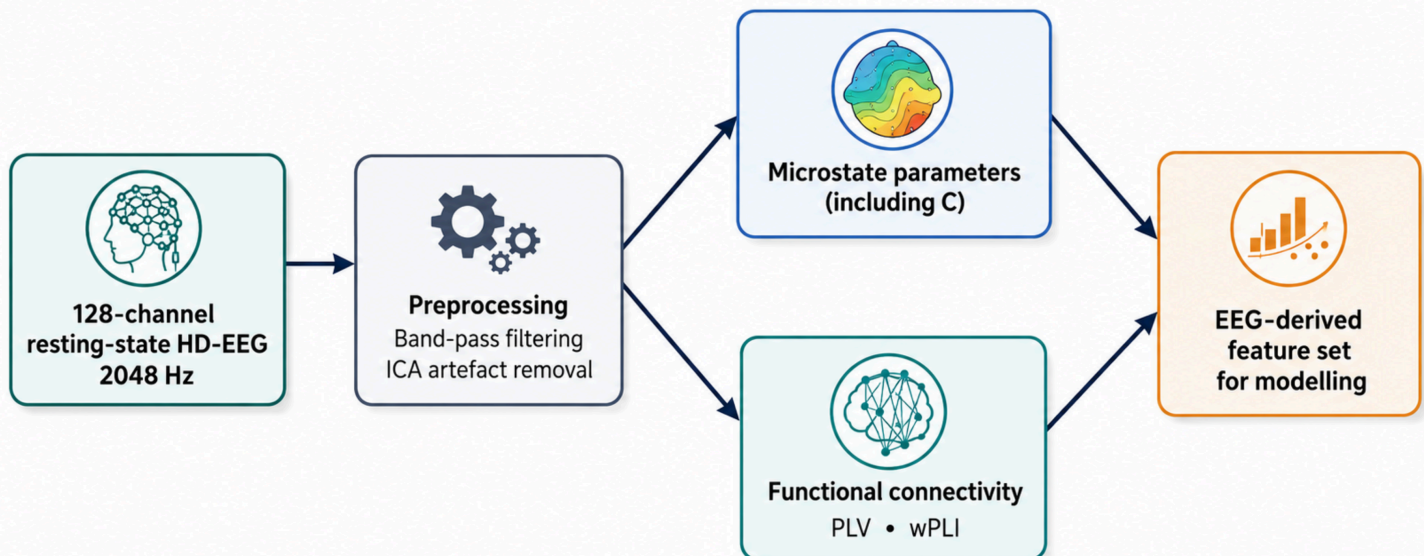


The figure describes the feature domains used for model development; it does not imply that every participant contributed every modality at every scheduled visit.

Figure 2

Multimodal feature domains and their analytic roles in the recurrence-classification framework. Clinical variables contextualised the cohort; longitudinal metabolomics supplied the principal molecular input; HD-EEG supplied microstate and functional-connectivity features.

HD-EEG acquisition and feature-extraction pathway



i Offline resing-state recordings were used for recurrence-status classification; the analysis does not imply real-time BCI decoding or prospective dynamic prediction.

Figure 3

HD-EEG acquisition and feature-extraction pathway. PLV, phase-locking value; wPLI, weighted phase-lag index; ICA, independent component analysis.

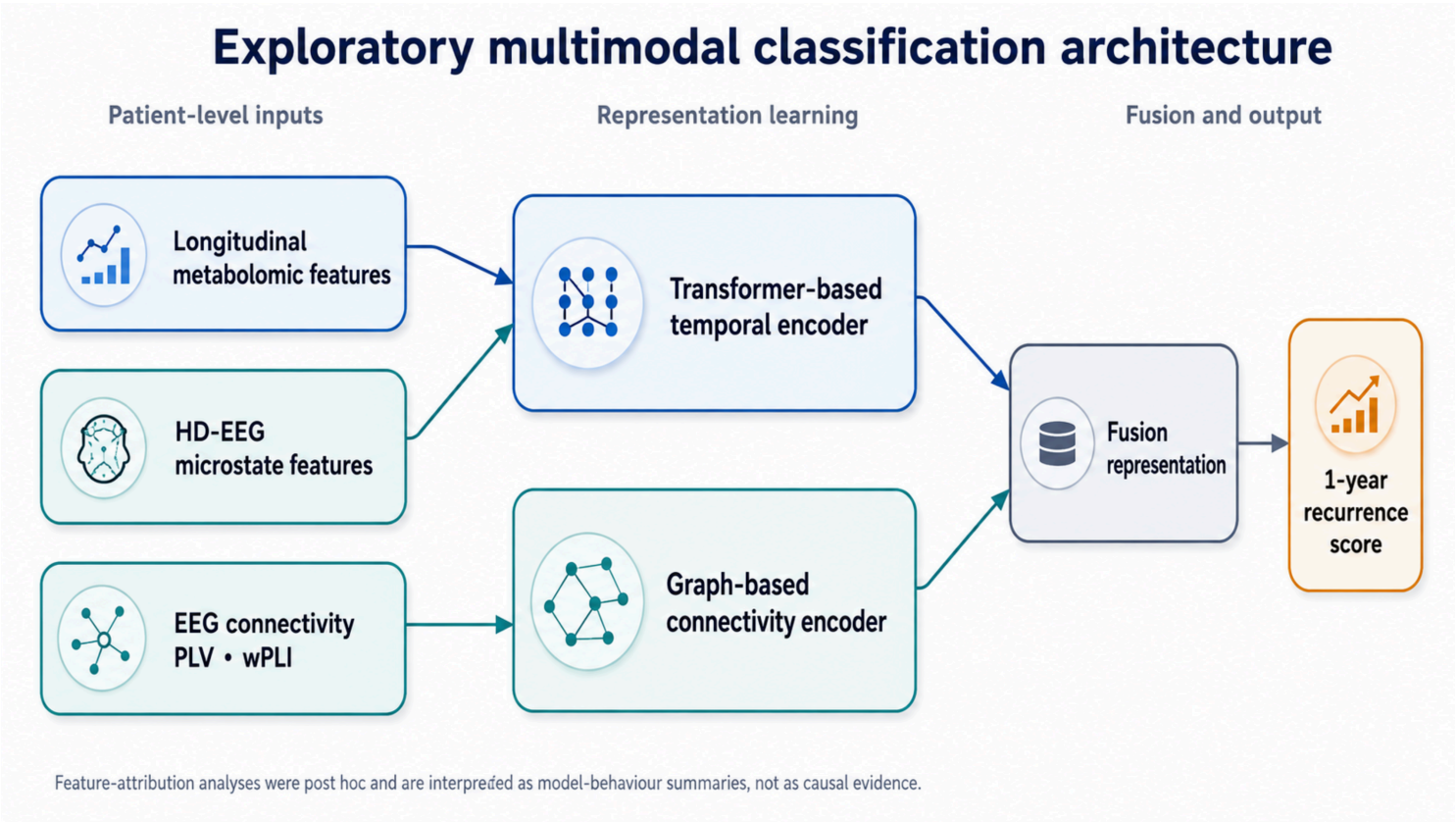


Figure 4

Simplified architecture of the exploratory multimodal model. Longitudinal features are encoded temporally, EEG connectivity is encoded graphically, and the learned representations are fused to estimate one-year recurrence status. The current analysis is interpreted as retrospective classification rather than validated dynamic prediction.

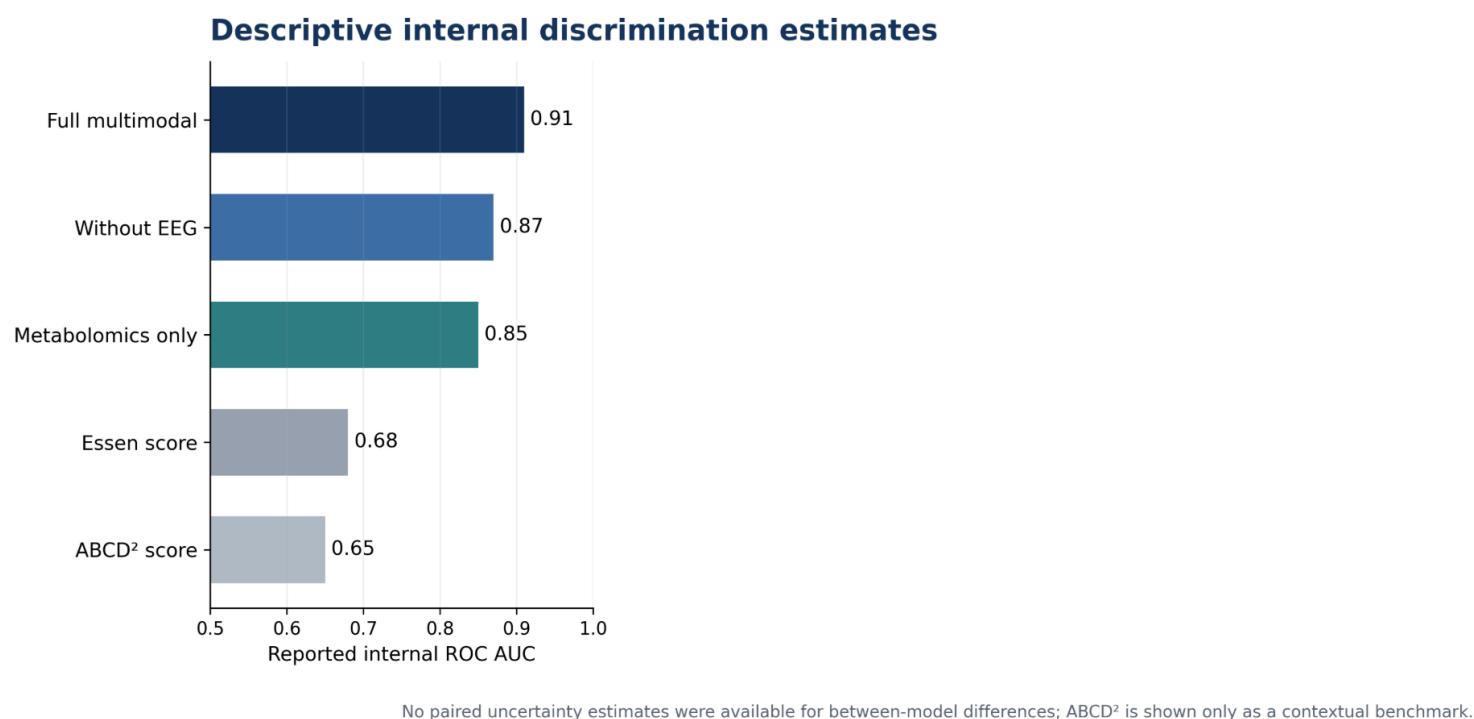


Figure 5

Reported internal ROC AUC values for the full and reduced-modality models and contextual clinical-score benchmarks. Between-model differences were not accompanied by paired uncertainty estimates and should not be interpreted as confirmed incremental classification value.

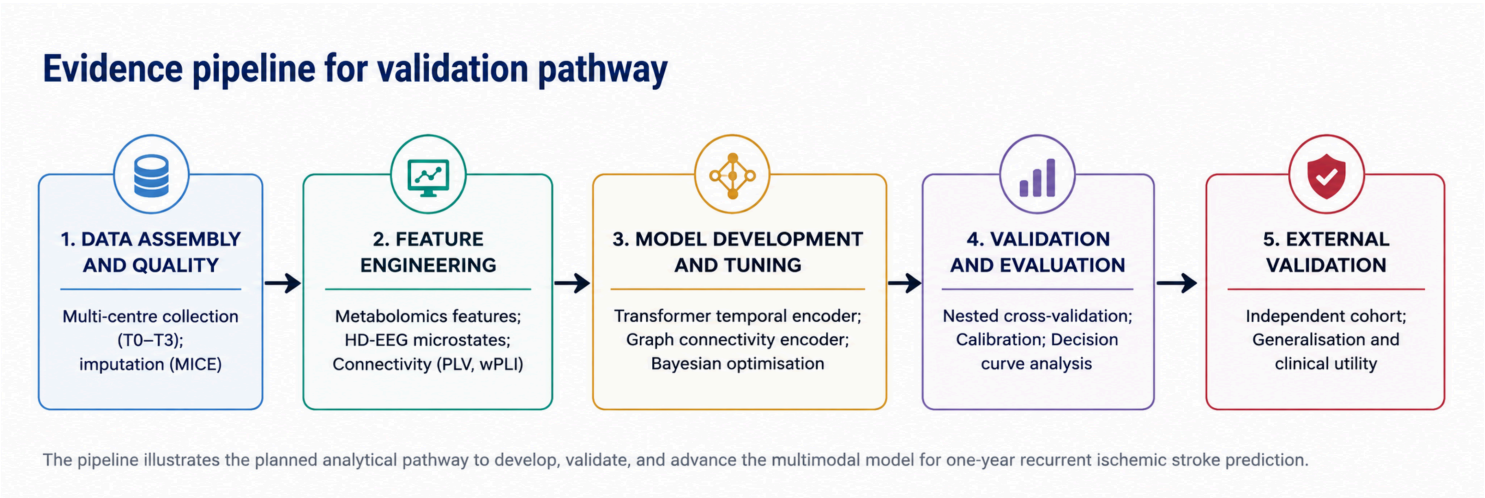


Figure 6

Evidence maturity and validation pathway. The current study provides retrospective model-development and internal classification evidence. Participant-level nested evaluation with verified pre-event inputs, complete calibration and independent multicentre evaluation are required before prospective clinical risk prediction.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMethodologicalEvidenceMatrix.docx](#)