

Explainable models for predicting long-term outcomes in patients with spontaneous intracerebral haemorrhage: a retrospective cohort study

Kai-Cheng Yang , Yu-Jia Jin, Li-Li Tang, Feng Gao, Lusha Tong 

To cite: Yang K-C, Jin Y-J, Tang L-L, *et al.* Explainable models for predicting long-term outcomes in patients with spontaneous intracerebral haemorrhage: a retrospective cohort study. *Stroke & Vascular Neurology* 2025;**0**. doi:10.1136/svn-2024-003864

► Additional supplemental material is published online only. To view, please visit the journal online (<https://doi.org/10.1136/svn-2024-003864>).

K-CY and Y-JJ contributed equally.

Received 8 November 2024
Accepted 14 February 2025



© Author(s) (or their employer(s)) 2025. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ Group.

Department of Neurology, The Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

Correspondence to

Dr Lusha Tong;
2310040@zju.edu.cn

Dr Feng Gao;
2202012@zju.edu.cn

ABSTRACT

Background and aim Recently, long-term outcomes in patients with spontaneous intracerebral haemorrhage (sICH) have gained increasing attention besides acute-phase characteristics. Predictive models for long-term outcomes are valuable for risk stratification and treatment strategies. This study aimed to develop and validate an explainable model for predicting long-term recurrence and all-cause death in patients with ICH, using clinical and imaging markers of cerebral small vascular diseases from MRI.

Method We retrospectively analysed data from a prospectively collected large-scale cohort of patients with acute ICH admitted to the Neurology Department of The Second Affiliated Hospital of Zhejiang University between November 2016 and April 2023. After comprehensive variable selection using least absolute shrinkage and selection operator and stepwise Cox regression, we constructed Cox proportional hazards models to predict recurrence and all-cause death. Model performance was evaluated using the concordance index, integrated Brier score and time-dependent area under the curve. Global and local interpretability were assessed using variable importance calculated as SurvSHAP(t) and SurvLIME methods for the entire training set and individual patients, respectively.

Results A total of 842 eligible patients were included. Over a median follow-up of 36 months (IQR: 12–51), 86 patients (9.1%) died, and 62 patients (6.6%) experienced recurrence of ICH. The concordance indexes for the all-cause death and recurrence models were 0.841 (95% CI 0.767 to 0.913) and 0.759 (95% CI 0.651 to 0.867), respectively, with integrated Brier scores of 0.079 and 0.063. The interpretability maps highlighted age, aetiology of ICH and low haemoglobin as key predictors of long-term death, while cortical superficial siderosis and previous haemorrhage were crucial for predicting recurrence.

Conclusions This model demonstrates high predictive accuracy and emphasises the crucial factors in predicting long-term outcomes of patients with sICH.

INTRODUCTION

Spontaneous intracerebral haemorrhage (sICH) is one of the most devastating forms of stroke, accounting for approximately 10–15% of all strokes.¹ Despite advances in acute management, the poor prognosis

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Spontaneous intracerebral haemorrhage (sICH) is associated with high mortality, disability and recurrence rates, contributing significantly to the global burden of stroke. Existing prognostic models mainly focus on the prediction of short-term mortality but have limitations in predicting long-term outcomes.

WHAT THIS STUDY ADDS

⇒ This study proposes an explainable predictive model for long-term death and recurrence in patients with sICH, using comprehensive clinical and MRI variables. It incorporates interpretability methods, such as SurvSHAP and SurvLIME, to enhance clinical usability.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The model offers a practical tool for clinicians to stratify long-term risk in patients with ICH and guide personalised management strategies.

remains. In addition to the mortality rate up to 50% within 30 days, the long-term recurrence rate, disability rate and mortality bring 79.5 million disability-adjusted life years globally, accounting for nearly half of all types of stroke.^{2–3} Given these challenges, there is an urgent need for accurate and clinically applicable models that can predict long-term outcomes in patients with ICH, to guide treatment decisions and to identify high-risk individuals requiring closer surveillance and stricter secondary preventive strategies.

The ICH score was proposed since 2001 as a prognosis model for patients with ICH. Based on several simple clinical variables, it has provided efficient short-term mortality risk stratification.⁴ Subsequent studies have modified the ICH score and improved its predictive accuracy in these decades.^{5–7} Recently, with the surge of artificial intelligence (AI) algorithms, prognostic models have been developed to further improve predictive performance. But these AI-based prognostic

models are mostly aimed at acute phase. For instance, a machine learning algorithm that incorporates clinical variables such as age, National Institutes of Health Stroke Scale (NIHSS) score and haematoma volume achieved an area under the curve (AUC) of over 0.8 in predicting 90-day outcomes for patients with ICH in the validation set.⁸ The prognostic performances of the two AI models combined with CT were developed and showed prediction accuracy exceeding ICH score.^{9–10} However, models predicting long-term events are scarce. Additionally, small sample sizes, short follow-up or failure to include various variables such as the aetiology of ICH are becoming the major limitation to form a long-term perspective for patients with ICH.

One of the primary challenges in the clinical application of prediction models is the ‘black box’ nature of algorithms.¹¹ While these models often demonstrate high accuracy, their lack of interpretability makes it difficult for clinicians to understand the prediction process for individual patients and to trust the model’s reliability. Given the complexity of ICH management, the development of explainable model is also crucial.

In this context, we proposed a clinically explainable predictive model for long-term death and recurrence in patients with sICH, using a prospective cohort from a tertiary stroke centre. Our model incorporates a comprehensive resource of clinical, imaging and aetiological variables and provides interpretability through Survival SHapley Additive exPlanations (SurvSHAP) and Survival Local Interpretable Model-agnostic Explanations (SurvLIME).¹²

METHOD

Patient selection

The cohort included patients with ICH consecutively admitted to the Neurology Department of The Second Affiliated Hospital of Zhejiang University between November 2016 and April 2023. Eligible patients underwent both CT and MRI during their hospitalisation. Patients were excluded if they had secondary ICH, modified Rankin Scale score ≥ 3 prior to the ICH, isolated intraventricular haemorrhage (IVH), or severe liver or renal disease. Additionally, patients using anticoagulants within 1 week prior to the ICH event were also excluded.

Data collection

Comprehensive clinical and laboratory data were extracted from the hospital’s electronic medical record system. Key variables included baseline assessments at admission, such as age, sex, Glasgow Coma Scale (GCS) and NIHSS score. Laboratory data encompassed blood routine examination, creatinine, fasting blood glucose, cholesterol, uric acid and coagulation function. Due to missing values in some laboratory factors (creatinine 7.24%, fasting blood glucose 7.60%, low-density lipoprotein-cholesterol 7.95% and uric acid 23.87%), multiple imputations using predictive mean matching

method were performed using MICE package in R. Missing values were imputed using five datasets generated through iterative regression models, which considered correlations among variables. Visualisations of the distribution of original and imputed values across five imputation iterations were provided in online supplemental figure S1. Relevant medical history, including smoking and alcohol use, hypertension, diabetes, coronary artery disease, atrial fibrillation, ischaemic stroke or transient ischaemic attack, previous ICH and medication history, was collected. We also collected data on mechanical ventilation, severe pneumonia, heart failure and renal failure during hospitalisation, which were commonly associated with the poor prognosis.^{13–14} Imaging assessments involved the location and volume of haemorrhage, presence of ventricular or subarachnoid extension, cerebral microbleed (CMB) and cortical superficial siderosis (CSS). Two neurologists assessed the presence of CSS and the number of CMBs. The CSS was defined as a curvilinear signal loss on the susceptibility-weighted image (SWI) in compliance with the gyral cortical surface within the subarachnoid space, away from at least two sulci of the haemorrhage with no corresponding signal hyperintensity on the baseline CT scan.¹⁵ The CMB was defined as round or ovoid signal voids between 2 and 10 mm in diameter, with associated blooming on the SWI sequences.¹⁶ CMB locations were categorised as lobar, deep structures (basal ganglia, thalamus and internal capsule), the brainstem or the cerebellum.¹⁷ The assessment of white matter hyperintensity (WMH) included both periventricular and deep regions. The severity of WMH was visually evaluated on axial fluid-attenuated inversion recovery sequence images using the Fazekas scale.¹⁸ The enlarged perivascular space was graded based on the number of dilated spaces in the contralateral centrum semiovale and deep regions, categorised into four grades: 0, 1–10, 11–20 and >20 . Detailed protocols of variable collection have been described in our previous work.^{17–19} Aetiological classification was based on the SMASH-U system, which categorizes sICH into structural vascular abnormalities, medication-related, amyloid angiopathy (CAA), systemic disease, hypertension, and undetermined causes.²⁰ As patients with secondary ICH were excluded, the classification was simplified into three categories: hypertensive vasculopathy, CAA and undetermined origin. Follow-up data were obtained through in-person or telephone interviews at regular intervals to determine patient survival and ICH recurrence. Termination of follow-up was defined as the last follow-up date before censoring, or the end of the longest 60-month follow-up period, or all-cause death.

Model development and testing

All patients were randomly divided into the training cohort (70%) and testing cohort (30%). The variable selection process was predefined. Initially, univariate Cox regression was performed to identify predictors with a p value <0.2 . The least absolute shrinkage and selection operator (LASSO) regression and backward

stepwise regression were implemented for further selection. LASSO was based on the minimum lambda, while stepwise regression employed the Akaike information criterion (AIC) for model refinement. The details of variable selection were presented in the online supplemental materials tables S1-6. The final set of variables was determined by integrating those selected from the LASSO and stepwise regression with key variables supported by published work. Using the training cohort, multivariable Cox proportional hazards models were trained to predict all-cause death and recurrence. Multicollinearity among variables was assessed using the variance inflation factor (VIF), with a VIF <5 indicating an absence of multicollinearity.

Model performance was evaluated in the testing cohort. Harrell's concordance index (C-index) was used to present overall discriminative ability, while the integrated Brier score presented overall calibration. Additionally, time-dependent receiver operating characteristic (ROC) curves and calibration curves were generated to visualise the model's predictive performance at 1, 3 and 5 years. The nomogram was constructed to represent the model's predictive capacity graphically. We also provided a nomogram and time-dependent ROC curves for 1, 3 and 6-month deaths in the online supplemental figures S2 and S3 to comprehensively evaluate the model's performance for prediction of both short-term and long-term deaths in patients with sICH.

To evaluate the robustness of our model, we conducted a temporal validation using patients admitted between April 2023 and May 2024 as an independent validation cohort. Due to the relatively short follow-up duration in the temporal validation cohort (maximum of 20 months), we evaluated the model's performance specifically for predicting 1-year outcomes. Baseline characteristics and model performance on the temporal validation cohort are provided in the online supplemental table S7 and figure S4.

Model interpretability

To enhance the clinical interpretability of the model, we employed both global and local interpretability methods. For global interpretability across the training cohort, we applied time-dependent variable importance using the SurvSHAP(t) method.²¹ SurvSHAP(t) enabled the assessment of the significance of variables over time and offered insights into the overall feature influence on survival predictions. For local interpretability at the individual level, we used the SurvLIME method, which facilitates patient-specific prediction analysis by identifying the influential features contributing to individual outcomes.²² This approach allows for a deeper understanding of personalised risk factors, thereby enhancing clinical decision-making.

Statistical analysis

Continuous variables that adhered to a normal distribution were reported as mean±SD, while non-normally

distributed data were presented as median and IQR. Comparisons between groups for continuous variables were performed using the t-test or Mann-Whitney U test. Categorical variables were summarised as percentages, with group differences analysed using Pearson's χ^2 test or Fisher's exact test as appropriate. Survival analysis was carried out with a Cox proportional hazards model, and the results were presented as HR with 95% CI. A two-tailed $p < 0.05$ was considered statistically significant.

Statistical analyses and visualisations were conducted using R software V.4.3.1 (R Foundation for Statistical Computing, Vienna, Austria). The R packages used included 'survival', 'MASS', 'rms', 'glmnet', 'survminer', 'riskRegression', 'timeROC', 'survex', 'ggplot2' and 'VIM'.

RESULTS

Baseline characteristics

The baseline characteristics of the cohort are summarised in table 1. A total of 842 patients with primary ICH were included in the analysis, with 592 allocated to the training set and 250 to the testing set. The mean age was comparable (61.4 years in the training set vs 60.3 years in the testing set, $p = 0.260$) between groups, and the majority of patients were male (64.7% in the training set vs 63.2% in the testing set, $p = 0.738$). Mortality rates (10.30% vs 10.00%, $p = 0.993$) and recurrence rates (6.93% vs 8.40%, $p = 0.546$) were comparable between the two datasets. However, renal disease (6.00% vs 2.87%, $p = 0.049$) and mean fasting blood glucose levels (6.56 mmol/L vs 6.14 mmol/L, $p = 0.024$) differed significantly between groups. Other variables, including GCS, coronary heart disease, coagulation, haematoma volume, location, CSS, CMBs and aetiology of ICH, showed no significant differences between training and testing datasets.

Variables predicting long-term all-cause death and ICH recurrence

Table 2 summarises the multivariate Cox models used to predict long-term all-cause death and ICH recurrence. In the all-cause death model, factors associated with an increased risk of death included advanced age (HR 1.07; 95% CI 1.04 to 1.10; $p < 0.001$), coronary heart disease (HR 2.97; 95% CI 1.23 to 7.16; $p = 0.015$), previous haemorrhage (HR 2.65; 95% CI 1.35 to 5.19; $p = 0.004$), a higher burden of total CMBs (HR 1.02; 95% CI 1.01 to 1.03; $p < 0.001$), a lower haemoglobin level (HR 0.97; 95% CI 0.95 to 0.99; $p = 0.003$) and intraventricular extension (HR 2.03; 95% CI 1.23 to 7.16; $p = 0.011$). Patients with CAA (HR 2.42; 95% CI 1.32 to 4.43; $p = 0.004$) or undetermined aetiology (HR 3.55; 95% CI 1.76 to 7.17; $p < 0.001$) had a significantly higher risk of death compared with hypertensive ICH.

In the ICH recurrence model, predictors of recurrent haemorrhage included a history of previous haemorrhage (HR 4.79; 95% CI 2.38 to 9.62; $p < 0.001$), the presence of CSS (HR 3.63; 95% CI 1.88 to 6.99; $p < 0.001$), a higher

Table 1 Baseline characteristics and outcomes in training and testing sets

	Training data (n=592)	Testing data (n=250)	P value
Age (years), mean (SD)	61.4 (13.0)	60.3 (12.7)	0.260
Male, n (%)	383 (64.7)	158 (63.2)	0.738
Admission NIHSS, median (IQR)	4 (2–10)	4 (2–10)	0.992
Admission GCS, median (IQR)	15 (14–15)	15 (14–15)	0.828
ODT (hours), mean (SD)	8.68 (10.3)	8.68 (10.3)	1.000
Hypertension, n (%)	450 (76.0)	192 (76.8)	0.876
Diabetes, n (%)	43 (17.2)	91 (15.4)	0.576
Atrial fibrillation, n (%)	8 (1.35)	9 (3.60)	0.064
Coronary heart disease, n (%)	25 (4.22)	8 (3.20)	0.614
Previous intracerebral haemorrhage, n (%)	47 (7.94)	14 (5.60)	0.293
AIS/TIA, n (%)	56 (9.46)	29 (11.6)	0.414
Renal disease, n (%)	17 (2.87)	15 (6.00)	0.049
Antiplatelet use, n (%)	61 (10.3)	25 (10.0)	0.993
Statin use, n (%)	44 (7.43)	17 (6.80)	0.859
Smoke, n (%)	191 (32.3)	85 (34.0)	0.682
Drink, n (%)	196 (33.1)	71 (28.4)	0.208
INR, mean (SD)	1.02 (0.39)	1.00 (0.09)	0.413
Platelet ($\times 10^9/L$), mean (SD)	199 (64.7)	206 (68.6)	0.180
Haemoglobin (g/L), mean (SD)	138 (16.6)	139 (16.4)	0.413
White blood cell ($\times 10^9/L$), mean (SD)	8.50 (2.90)	8.36 (2.67)	0.487
Total cholesterol ($\mu\text{mol/L}$), mean (SD)	4.71 (1.04)	4.65 (1.02)	0.389
LDL-cholesterol ($\mu\text{mol/L}$), mean (SD)	2.48 (0.78)	2.46 (0.69)	0.661
FBG (mmol/L), mean (SD)	6.14 (1.93)	6.56 (2.62)	0.024
Uric acid ($\mu\text{mol/L}$), mean (SD)	283 (104)	284 (106)	0.958
Creatinine ($\mu\text{mol/L}$), mean (SD)	74.8 (98.4)	83.2 (121)	0.334
Mechanical ventilation, n (%)	6 (0.8)	2 (1.01)	1.000
Severe pneumonia, n (%)	8 (1.35)	4 (1.60)	0.757
Heart failure, n (%)	9 (1.52)	4 (1.60)	1.000
Renal failure, n (%)	21 (3.55)	14 (5.60)	0.240
OMT (day), mean (SD)	5.39 (3.01)	5.34 (2.88)	0.794
Haematoma volume (mL), mean (SD)	12.8 (13.3)	12.6 (12.7)	0.870
Haematoma location, n (%)			0.664
Lobar	116 (19.6)	47 (18.8)	
Deep	396 (66.9)	157 (62.8)	
Cerebellum	38 (6.42)	21 (8.40)	
Brainstem	49 (8.28)	29 (11.6)	
Intraventricular extension, n (%)	160 (27.0)	74 (29.6)	0.498
Subarachnoid extension, n (%)	51 (8.61)	23 (9.20)	0.888
Cerebral microbleeds, mean (SD)			
Lobar	3.56 (9.39)	3.30 (10.1)	0.729
Deep	2.73 (5.24)	2.62 (4.28)	0.742
Cerebellum	0.55 (1.66)	0.74 (2.70)	0.285
Brainstem	0.75 (1.71)	0.62 (1.47)	0.259
Total	7.59 (14.3)	7.28 (14.6)	0.780
Cortical superficial siderosis, n (%)	114 (19.3)	41 (16.4)	0.379

Continued

Table 1 Continued

	Training data (n=592)	Testing data (n=250)	P value
PWMH, n (%)			0.360
None	223 (37.7)	90 (36.0)	
Fazekas I	144 (24.3)	63 (25.2)	
Fazekas II	111 (18.8)	58 (23.2)	
Fazekas III	114 (19.3)	39 (15.6)	
DWMH, n (%)			0.792
None	275 (46.5)	115 (46.0)	
Fazekas I	174 (29.4)	81 (32.4)	
Fazekas II	86 (14.5)	33 (13.2)	
Fazekas III	57 (9.63)	21 (8.40)	
PEPVS, n (%)			0.275
None	368 (62.2)	162 (64.8)	
1–10	158 (26.7)	63 (25.2)	
11–20	39 (6.59)	20 (8.00)	
Over 20	27 (4.56)	5 (2.00)	
DEPVS, n (%)			0.432
None	203 (34.3)	94 (37.6)	
1–10	252 (42.6)	92 (36.8)	
11–20	87 (14.7)	38 (15.2)	
Over 20	50 (8.45)	26 (10.4)	
Aetiology, n (%)			0.943
Amyloid angiopathy	94 (15.9)	38 (15.2)	
Hypertension	380 (64.2)	160 (64.0)	
Undetermined	118 (19.9)	52 (20.8)	
Death, n (%)	61 (10.3)	25 (10.0)	0.993
Recurrent ICH, n (%)	41 (6.93)	21 (8.40)	0.546
Follow-up time (days), median (IQR)	36 (12–51)	36 (12–49)	0.945
AIS, acute ischaemic stroke; DEPVS, deep enlarged perivascular spaces; DWMH, deep white matter hyperintensities; FBG, fasting blood glucose; GCS, Glasgow Coma Scale; ICH, intracerebral haemorrhage; INR, international normalised ratio; LDL, low-density lipoprotein; NIHSS, National Institutes of Health Stroke Scale; ODT, onset to door time; OMT, onset to MRI time; PEPVS, periventricular enlarged perivascular spaces; PWMH, periventricular white matter hyperintensities; TIA, transient ischaemic attack.			

burden of lobar CMBs (HR 1.02; 95% CI 1.00 to 1.04; $p=0.041$) and a lower haemoglobin level (HR 0.98; 95% CI 0.95 to 0.99; $p=0.019$).

Model performance

The performance metrics of the predictive models were summarised in [table 3](#). For the prediction of all-cause death within 5 years, the model achieved a C-index of 0.830 (95% CI 0.883 to 0.776) in the training set and 0.841 (95% CI 0.767 to 0.913) in the testing set, indicating strong discriminative ability. The integrated Brier scores were 0.069 and 0.079 for the training and testing sets, respectively. Time-dependent AUCs at 1, 3 and 5 years in the testing set were 0.821 (95% CI 0.729 to 0.913), 0.845 (95% CI 0.748 to 0.942) and 0.892 (95% CI 0.810 to 0.974), respectively ([figure 1](#)). The recurrence model demonstrated a C-index of 0.740 (95% CI 0.832 to 0.648)

in the training set and 0.759 (95% CI 0.651 to 0.867) in the testing set, with integrated Brier scores of 0.058 and 0.063, respectively. Time-dependent AUCs for recurrence at 1, 3 and 5 years in the testing set were 0.808 (95% CI 0.642 to 0.973), 0.780 (95% CI 0.643 to 0.918) and 0.789 (95% CI 0.657 to 0.921), respectively. Calibration curves ([figure 2](#)) indicated good concordance between predicted and observed outcomes for both all-cause death and recurrence models. Nomograms were developed ([figure 3](#)) to facilitate clinical decision-making, enabling individualised predictions of survival and ICH recurrence at 1, 3 and 5 years based on key prognostic variables.

Feature importance and model interpretability

To enhance model interpretability, both global and local feature importance were assessed. The results were visualised in [figures 4 and 5](#). Global feature importance

Table 2 Multivariate models for predicting all-cause death and recurrent ICH in patients with spontaneous haemorrhage

Variable	Coefficient	HR (95% CI)	P value	VIF
Cox model for predicting all-cause death within 5 years				
Age	0.069	1.072 (1.045, 1.099)	<0.001	1.197
Aetiology				1.339
Hypertension	Ref	Ref	Ref	
Amyloid angiopathy	0.884	2.421 (1.323, 4.431)	0.004	
Undetermined	1.268	3.554 (1.762, 7.170)	<0.001	
Coronary heart disease	1.088	2.967 (1.229, 7.160)	0.015	1.087
Previous haemorrhage	0.973	2.645 (1.348, 5.190)	0.004	1.063
Haemoglobin	−0.028	0.972 (0.954, 0.990)	0.003	1.228
Total CMB	0.021	1.021 (1.013, 1.029)	<0.001	1.191
IVH extension	0.710	2.034 (1.229, 7.157)	0.011	1.110
Cox model for predicting recurrence of ICH within 5 years				
Prior haemorrhage	1.566	4.786 (2.382, 9.618)	<0.001	1.116
Haemoglobin	−0.023	0.977 (0.958, 0.996)	0.019	1.073
Lobar CMB	0.019	1.020 (1.001, 1.039)	0.041	1.105
Cortical superficial siderosis	1.289	3.629 (1.884, 6.989)	<0.001	1.128

CMB, cerebral microbleed; ICH, intracerebral haemorrhage; IVH, intraventricular haemorrhage; VIF, variance inflation factor.

was assessed via SurvSHAP values, showing that in the all-cause death model, age, aetiology of ICH and admission haemoglobin were consistently the most influential factors, significantly impacting mean SurvSHAP values throughout the follow-up period. In the recurrence model, previous haemorrhage, haemoglobin and CSS were identified as key predictors. Local interpretability methods enable the model to predict individual outcomes and assess how specific factors impact each patient's survival, sometimes uncovering unexpected differences.

As shown in [figure 5](#), if haemoglobin is identified as a key predictor, clinicians may prioritise anaemia correction. Similarly, if prior haemorrhage is recognised as a risk factor, stricter blood pressure control or closer follow-up may be necessary.

DISCUSSION

In the present study, we developed robust and explainable models using a large dataset to predict long-term

Table 3 Models' performance for prediction of sICH all-cause death and recurrence

Performance metrics	Training data	Test data
Cox model for predicting all-cause death within 5 years		
Concordance index	0.830 (0.883, 0.776)	0.841 (0.767, 0.913)
Integrated Brier score	0.069	0.079
Time-dependent AUC		
1 year	0.880 (0.805, 0.955)	0.821 (0.729, 0.913)
3 years	0.869 (0.815, 0.923)	0.845 (0.748, 0.942)
5 years	0.814 (0.737, 0.890)	0.892 (0.810, 0.974)
Cox model for predicting recurrence of ICH within 5 years		
Concordance index	0.740 (0.832, 0.648)	0.759 (0.651, 0.867)
Integrated Brier score	0.058	0.063
Time-dependent AUC		
1 year	0.782 (0.658, 0.907)	0.808 (0.642, 0.973)
3 years	0.751 (0.649, 0.853)	0.780 (0.643, 0.918)
5 years	0.749 (0.647, 0.850)	0.789 (0.657, 0.921)

AUC, area under the curve; sICH, spontaneous intracerebral haemorrhage.

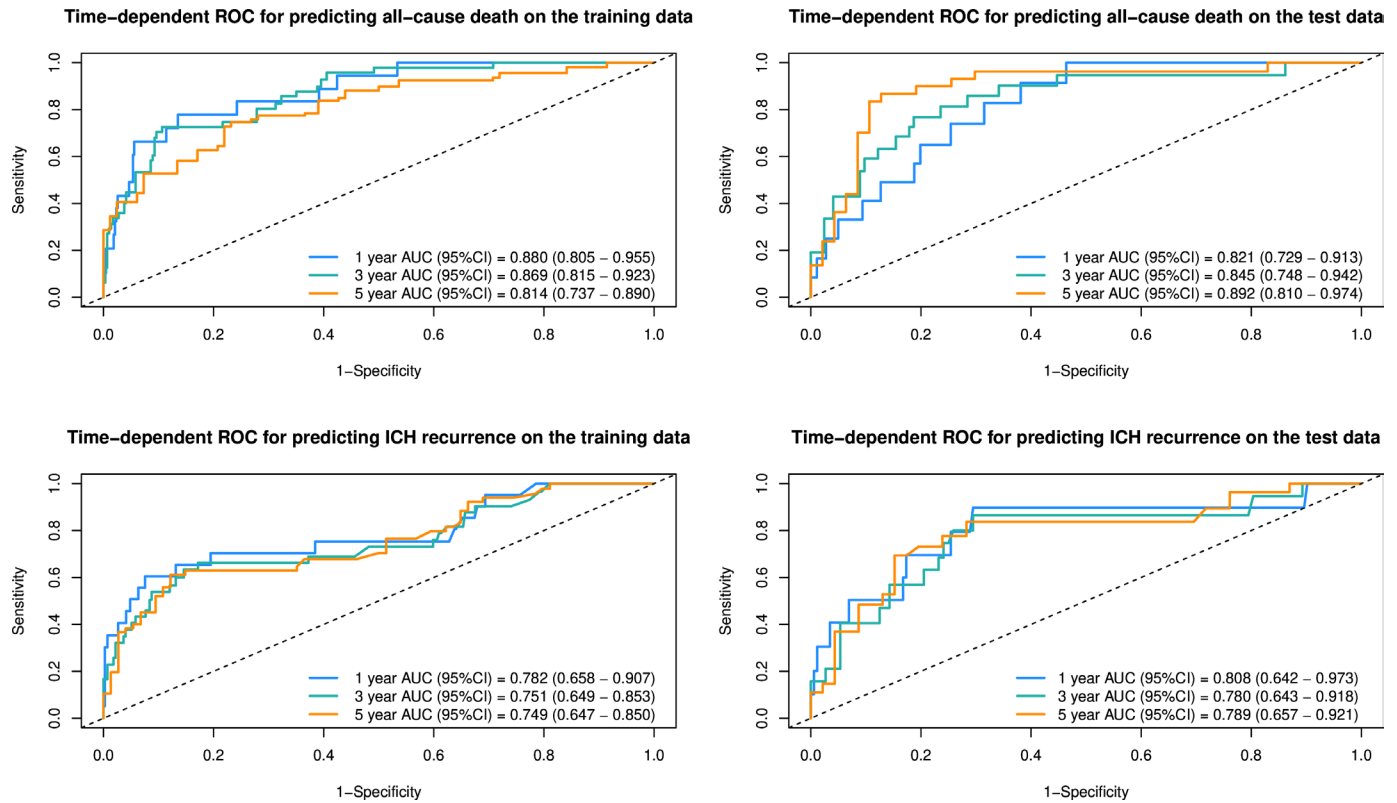


Figure 1 Time-dependent receiver operating characteristic (ROC) curves for predicting spontaneous intracerebral haemorrhage (sICH) all-cause death and recurrence. AUC, area under the curve.

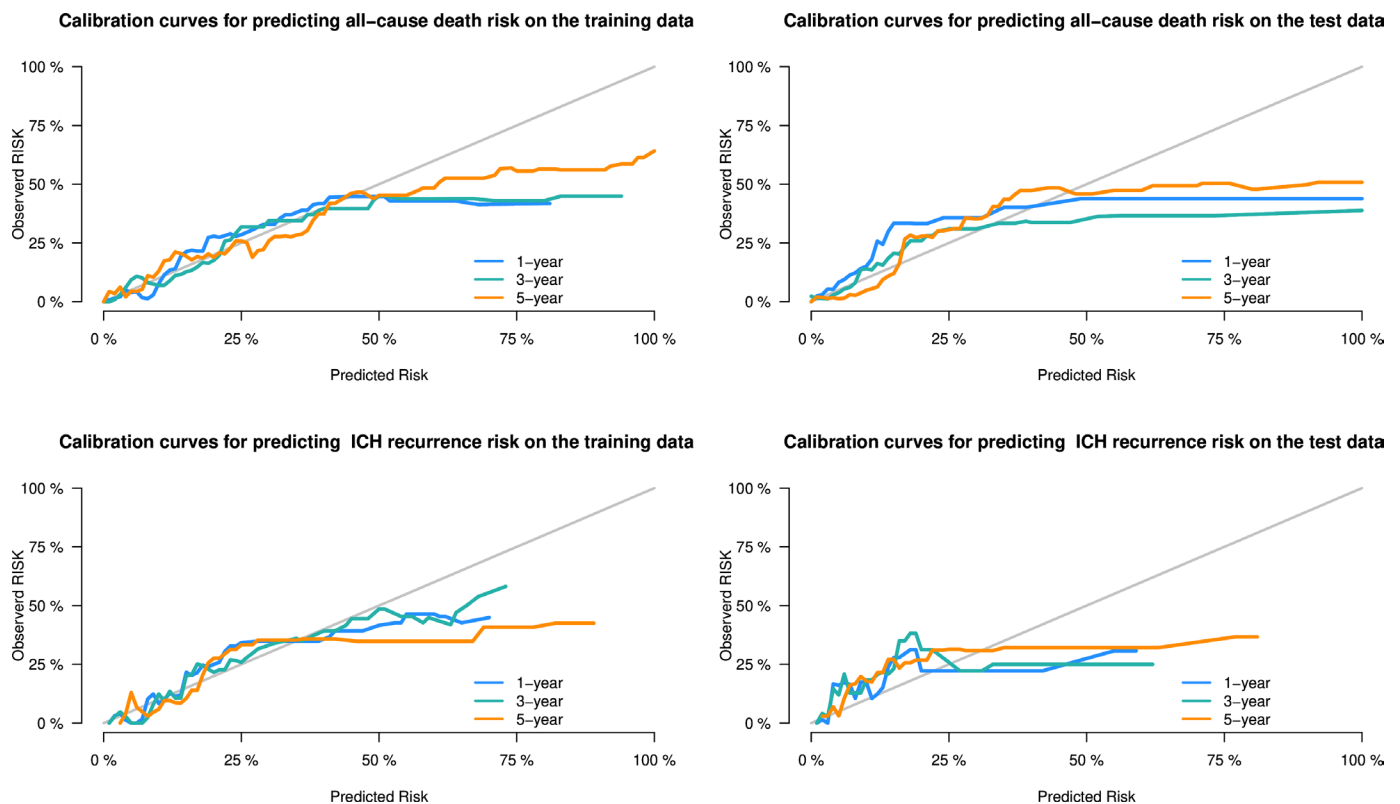


Figure 2 Calibration curves for predicting spontaneous intracerebral haemorrhage (sICH) all-cause death and recurrence.

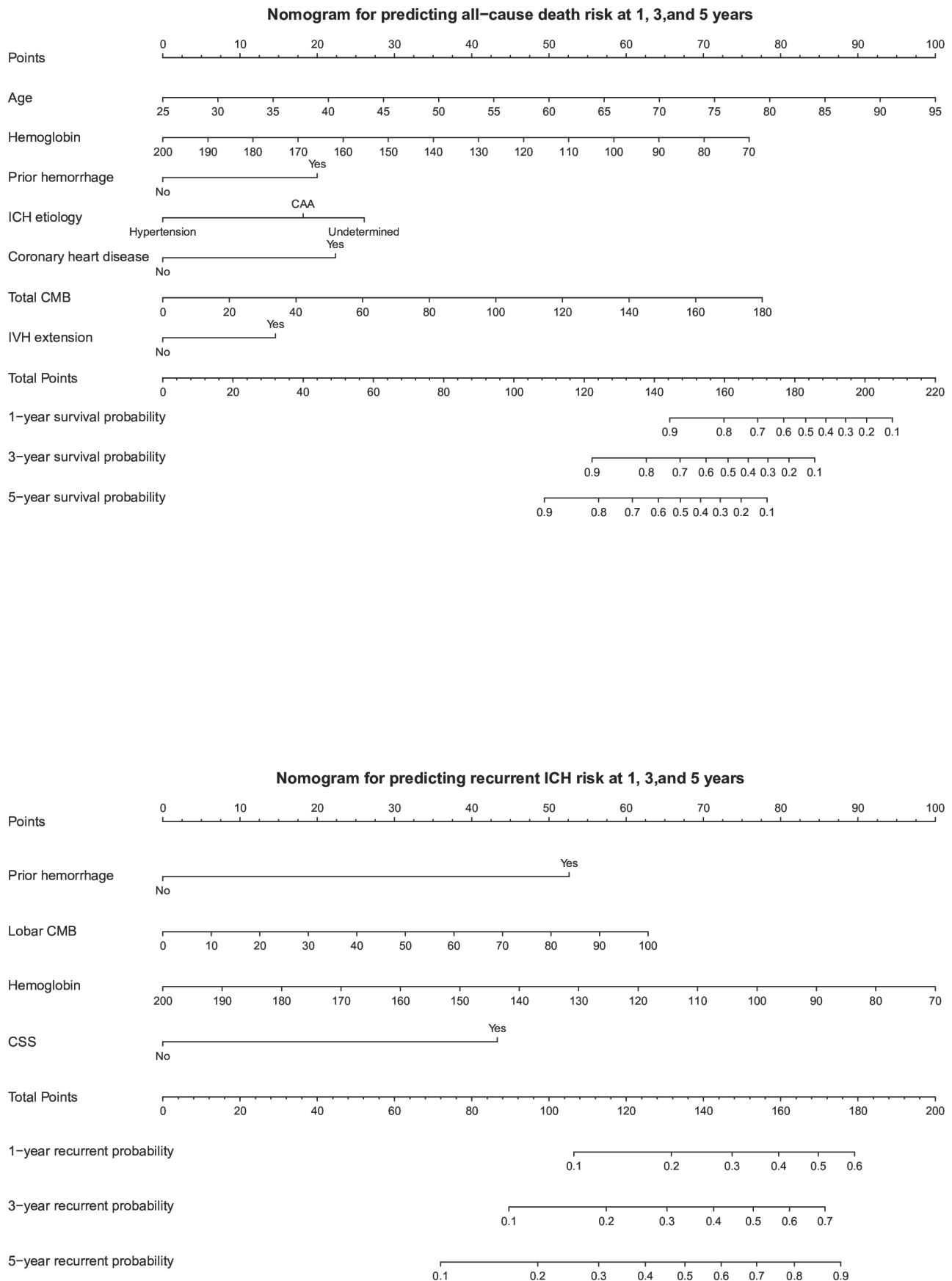
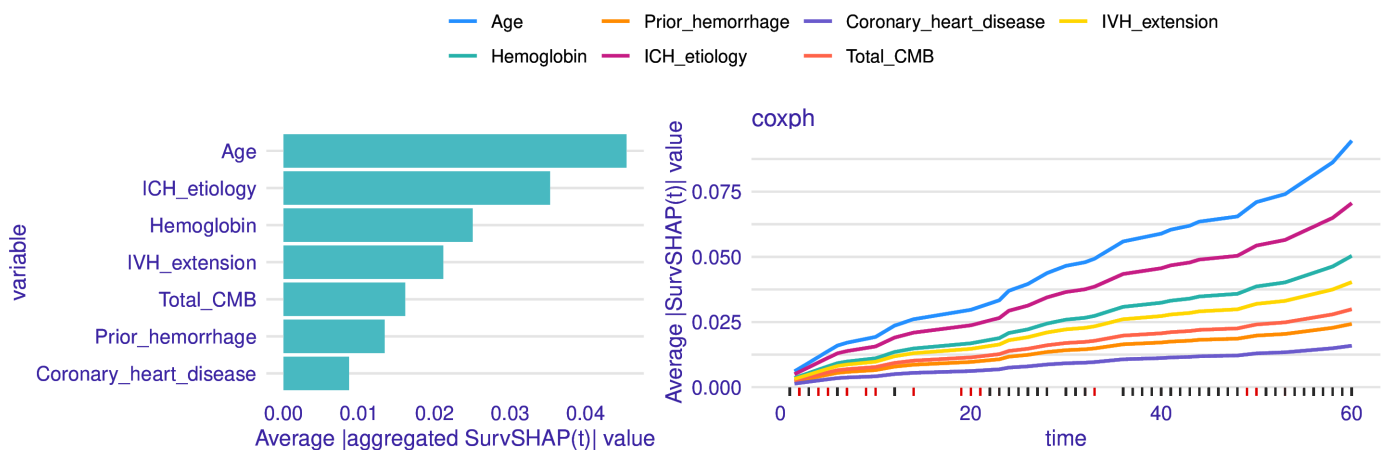


Figure 3 Nomogram for predicting long-term all-cause death and recurrence in patients with spontaneous intracerebral haemorrhage (sICH). CAA, cerebral amyloid angiopathy; CMB, cerebral microbleed; CSS, cortical superficial siderosis; IVH, intraventricular haemorrhage.

Feature importance according to aggregated |SurvSHAP(t)|

Based on a model predicting all-cause death



Feature importance according to aggregated |SurvSHAP(t)|

Based on a model predicting ICH recurrence

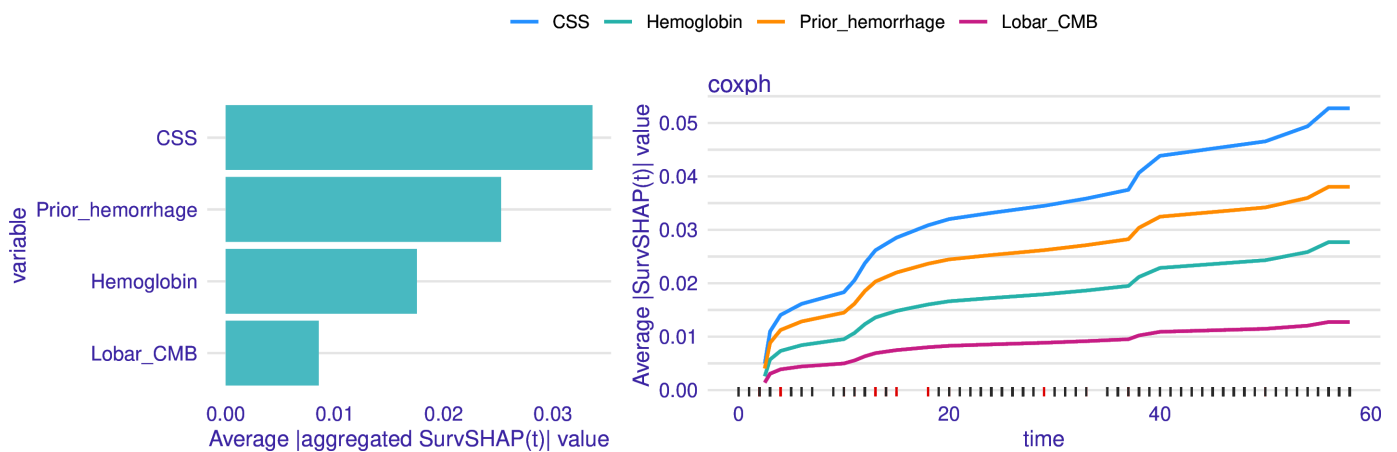


Figure 4 Global explanation for predicting long-term all-cause death and recurrence on training data. The one on the left presented overall importance of variables based on mean absolute SurvSHAP(t) value; longer bars indicate greater contribution to the prediction, with age being the most important factor for all-cause death and cortical superficial siderosis (CSS) for spontaneous intracerebral haemorrhage (sICH) recurrence. The right panels demonstrated the time-dependent feature importance of each variable based on the value at each time point across all observations, showing the trend of the importance of each variable over time. CMB, cerebral microbleed; IVH, intraventricular haemorrhage.

outcomes in patients with sICH with high accuracy and reliability. Our findings highlight that the aetiology of ICH, small vessel disease markers and haemoglobin levels were critical variables for models' performance. These models hold promise for stratifying long-term mortality and recurrence risks in patients with ICH, offering significant potential for optimising management and treatment strategies.

The variable selection process combined stepwise regression and LASSO, balancing the strengths and limitations of regularisation techniques and AIC. This approach minimised biases with using a single selection method within a single-centre dataset.²³ To prevent overfitting, we controlled the number of variables by excluding those with collinearity based on the VIF.

This strategy enhanced predictive efficiency, ensuring a concise yet powerful set of variables, thus maximising clinical utility. The present model demonstrated strong concordance and calibration, maintaining robustness on testing. This can be attributed to two key factors: the relatively large sample size, facilitating accurate parameter estimation, and the rigorous variable selection. However, the incidence of outcomes in the dataset was lower than reported in previous studies, likely due to the relatively mild conditions of our ICH cohort.²⁴ The lower incidence of outcomes, particularly for recurrence, led to an overestimation of individual risk, reflected by an elevated false-positive rate in the calibration curve.

In recent years, research on the ICH prognosis has primarily focused on outcomes within the first 3 months.

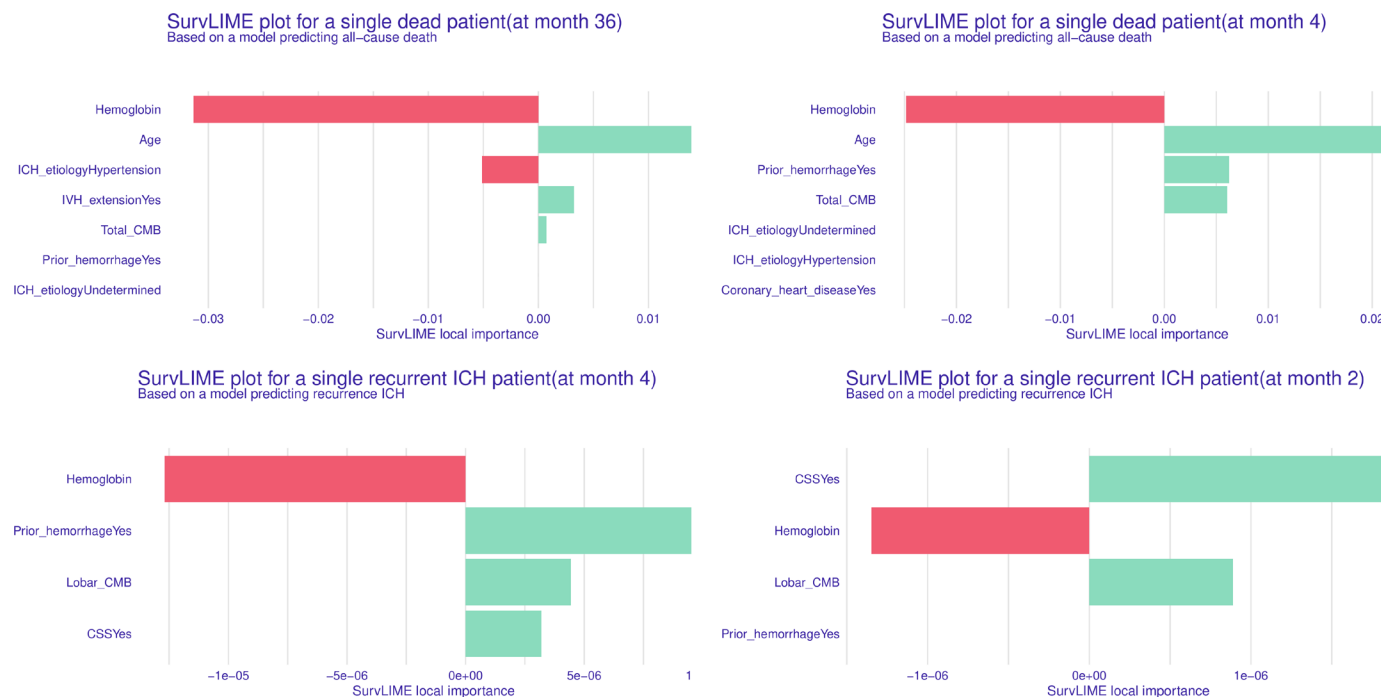


Figure 5 Local explanation: feature importance map based on SurvLIME for a single patient. Longer bars indicate greater contribution to the prediction. Green bars indicate factors associated with higher risk while red bars indicate protective effects. The top-left and top-right panels represent patients who died at 36 and 4 months of follow-up, respectively. The bottom-left and bottom-right panels depict patients who experienced intracerebral haemorrhage (ICH) recurrence at 4 and 2 months, with haemoglobin and cortical superficial siderosis (CSS) identified as the key predictive factors for recurrence in these two patients. CMB, cerebral microbleed; IVH, intraventricular haemorrhage.

Published models often incorporated variables such as age, IVH and haematoma volume, achieving C-statistics ranging from 0.7 to 1.0, with an average of approximately 0.88.^{8 25 26} However, few studies have explored long-term outcomes in the sICH population. Short-term mortality is often associated with acute-phase injuries and complications, while long-term mortality reflects a broader interplay of factors, including baseline health and chronic conditions.²⁴ Despite these differences, predicting long-term outcomes remains crucial for stratifying risk early and guiding clinical decision-making, which can help identify patients who may benefit from intensive rehabilitation, closer monitoring or preventative interventions targeting complications like recurrent haemorrhage. Baseline variables at admission, such as age, GCS and haematoma volume, remain critical predictors of long-term outcomes as they reflect the overall severity of ICH and patient health status. A recent study involving a training cohort of 480 patients used age, GCS and hydrocephalus secondary to IVH to predict 5-year mortality, achieving a concordance index of 0.76, which is lower than the performance of our proposed model.²⁷ Compared with previous studies, our research analysed a larger cohort and expanded the scope of model variables, incorporating both clinical characteristics and cerebral small vessel disease markers on MRI.^{28 29} These markers better reflect the severity of vascular damage. Our model achieved time-dependent AUCs exceeding 0.8 for 1, 3 and 5-year mortality predictions in both the training and

testing sets, with low integrated Brier scores, demonstrating satisfactory discrimination and calibration. More importantly, the use of visualisation tools enhances the model's interpretability, facilitating its potential application in clinical practice.

To enhance interpretability, we employed SurvSHAP and SurvLIME for global and local feature importance assessment, visualised in figures 4 and 5. Global interpretation using SurvSHAP identified age, aetiology and admission haemoglobin levels as the most influential predictors in the all-cause mortality model, consistent with findings from previous studies.^{2 5 24} The aetiology of ICH emerged as a crucial factor in prognosis, with CAA-related or undetermined aetiologies exhibiting a higher risk of long-term mortality compared with hypertensive ICH. This aligns with previous studies showing that haemorrhages caused by medication, CAA or systemic disease, as classified by the SMASH-U system, are associated with poorer outcomes due to factors such as older age, higher stroke recurrence and poorer overall health conditions.^{20 30 31} The system of ICH aetiology has been regarded as significantly enhancing the predictive ability of the max-ICH score.³¹ In our cohort, similar findings were observed. According to the mean SurvSHAP values, the aetiology of ICH emerged as a key predictor of all-cause death. The total CMB burden, an indicator of deteriorated small cerebrovascular condition, was also an independent predictor of all-cause death. Notably, our model introduced admission haemoglobin levels

as a predictor, a novel addition supported by growing evidence linking lower haemoglobin concentrations with poorer outcomes after stroke. Low haemoglobin may contribute to secondary brain injury through reduced oxygen delivery, metabolic disruption and impaired cellular energy processes.^{32 33} Other studies have associated low haemoglobin with haematoma expansion and increased small vessel damage, further elevating mortality and recurrent risks.^{34 35} Haemoglobin was shown to be a predictor of comparable importance to intraventricular extension, according to interpretability maps. For ICH recurrence, CSS and lobar CMBs, both markers of CAA, were strong predictors, likely due to increased vascular fragility caused by β -amyloid deposition.^{36 37} The local SurvLIME method provides interpretability by quantifying individual risk factors and can be seamlessly integrated into clinical workflows. It offers visualisations of key prognostic factors, helping clinicians effectively stratify patient mortality risk and develop personalised management plans. For instance, identifying haemoglobin as a dominant factor may prompt timely anaemia intervention, while prior haemorrhage may require enhanced secondary prevention.

This study has several limitations. It is a single-centre study, introducing potential selection bias. Patients included in this cohort were those hospitalised in the neurology department and capable of undergoing MRI, which indicates relatively mild ICH cases. This may partly explain the higher survival rates observed in our cohort compared with previous studies.^{24 28} Moreover, acute-phase complications, such as severe pneumonia and organ failure, which are known predictors of poor short-term outcomes, were infrequent in our cohort.^{13 14} As a result, these variables could not be incorporated as robust predictors in our model. These biases reflect the limitation in the generalisability of our findings to severe cases. Second, while the proposed model demonstrated robust predictive performance on the testing set, the lack of multicentre external validation remains a key limitation. Future work will focus on external validation using multicentre cohorts, incorporating a broader range of ICH severity. Additionally, recalibration with external datasets featuring higher outcome incidence could address the overestimation of risks observed in the calibration curves. Prospective validation in real-world clinical settings will also be essential to further confirm the model's reliability and applicability. Our predictions rely solely on admission data, which may not fully capture the complexity of long-term outcomes. For instance, dynamic changes in blood pressure and follow-up MRI findings of small vessel disease markers may also be associated with the prognosis of sICH.³⁸ However, collecting dynamic data poses significant challenges in practice, requiring collaboration across teams and sufficient resources. Future large-scale, multicentre cohort studies could explore the use of time-dependent predictors to refine prediction models. Finally, this study primarily focused on baseline clinical and imaging features, without explicitly accounting for

clinical management strategies, such as intensive blood pressure control, mannitol use and surgical interventions. These interventions can potentially influence ICH outcomes by preventing haematoma expansion and alleviating intracranial hypertension.^{39 40} However, the effectiveness and timing of these treatments can vary widely depending on individual patient characteristics and clinical settings. Quantifying their impact on prognosis within observational studies remains a significant challenge.

CONCLUSION

In conclusion, our explainable models demonstrate substantial clinical applicability for predicting long-term death and recurrence in patients with spontaneous intracerebral haemorrhages ICH. The integration of interpretable methods strengthens the model's utility by providing visualisations of individual predictions, enabling clinicians to make more appropriate and personalised treatment decisions. For improved accuracy, future studies should aim to include a larger sample size and a more diverse range of haemorrhage. Additionally, prospective validation in real-world clinical settings will be essential to confirm the model's robustness and generalisability.

Contributors K-cY: conceptualisation, methodology, software, validation, formal analysis, investigation, visualisation, data curation, writing—original draft preparation. Y-JJ: conceptualisation, methodology, formal analysis, investigation, data curation, writing—original draft preparation. LTa: conceptualisation, methodology, investigation, writing—review and editing. FG: project administration, supervision, methodology, writing—review and editing. LTo: project administration, supervision, methodology, writing—review and editing, funding acquisition. LTo is the guarantor.

Funding This research project was supported by the Zhejiang Provincial Medical and Health Science and Technology Project (2022KY174) and the 'Pioneer' R&D Program of Zhejiang (2024C03006).

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval The studies involving human participants were reviewed and approved by the Human Investigation Committee (IRB) of The Second Affiliated Hospital of Zhejiang University (ID: I20200011153). Trial participants signed written informed consent for data collection and follow-up before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. The data that support the findings of the study are available from the corresponding author upon reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Kai-Cheng Yang <http://orcid.org/0009-0006-8711-2882>

REFERENCES

- 1 Puy L, Parry-Jones AR, Sandset EC, *et al.* Intracerebral haemorrhage. *Nat Rev Dis Primers* 2023;9:14.
- 2 van Asch CJ, Luitse MJ, Rinkel GJ, *et al.* Incidence, case fatality, and functional outcome of intracerebral haemorrhage over time, according to age, sex, and ethnic origin: a systematic review and meta-analysis. *Lancet Neurol* 2010;9:167–76.
- 3 Pinho J, Costa AS, Araújo JM, *et al.* Intracerebral hemorrhage outcome: A comprehensive update. *J Neurol Sci* 2019;398:54–66.
- 4 Hemphill JC 3rd, Bonovich DC, Besmertis L, *et al.* The ICH score: a simple, reliable grading scale for intracerebral hemorrhage. *Stroke* 2001;32:891–7.
- 5 Ruiz-Sandoval JL, Chiquete E, Romero-Vargas S, *et al.* Grading scale for prediction of outcome in primary intracerebral hemorrhages. *Stroke* 2007;38:1641–4.
- 6 Cho DY, Chen CC, Lee WY, *et al.* A new Modified Intracerebral Hemorrhage score for treatment decisions in basal ganglia hemorrhage—a randomized trial. *Crit Care Med* 2008;36:2151–6.
- 7 Parry-Jones AR, Abid KA, Di Napoli M, *et al.* Accuracy and clinical usefulness of intracerebral hemorrhage grading scores: a direct comparison in a UK population. *Stroke* 2013;44:1840–5.
- 8 Geng Z, Yang C, Zhao Z, *et al.* Development and validation of a machine learning-based predictive model for assessing the 90-day prognostic outcome of patients with spontaneous intracerebral hemorrhage. *J Transl Med* 2024;22:236.
- 9 Zhao X, Zhou B, Luo Y, *et al.* CT-based deep learning model for predicting hospital discharge outcome in spontaneous intracerebral hemorrhage. *Eur Radiol* 2023;34:4417–26.
- 10 Zhang H, Yang YF, Song XL, *et al.* An interpretable artificial intelligence model based on CT for prognosis of intracerebral hemorrhage: a multicenter study. *BMC Med Imaging* 2024;24:170.
- 11 Reyes M, Meier R, Pereira S, *et al.* On the Interpretability of Artificial Intelligence in Radiology: Challenges and Opportunities. *Radiol Artif Intell* 2020;2:e190043.
- 12 Spytek M, Krzyżiński M, Langbein SH, *et al.* survex: an R package for explaining machine learning survival models. *Bioinformatics* 2023;39.
- 13 Hannawi Y, Hannawi B, Rao CPV, *et al.* Stroke-associated pneumonia: major advances and obstacles. *Cerebrovasc Dis* 2013;35:430–43.
- 14 Saeed F, Adil MM, Piracha BH, *et al.* Acute renal failure worsens in-hospital outcomes in patients with intracerebral hemorrhage. *J Stroke Cerebrovasc Dis* 2015;24:789–94.
- 15 Moulin S, Casolla B, Kuchcinski G, *et al.* Cortical superficial siderosis: A prospective observational cohort study. *Neurology (Ecricon)* 2018;91:e132–8.
- 16 Magid-Bernstein JR, Li Y, Cho S-M, *et al.* Cerebral Microbleeds and Acute Hematoma Characteristics in the ATACH-2 and MISTIE III Trials. *Neurology (Ecricon)* 2022;98:e1013–20.
- 17 Jin YJ, Li JW, Wu J, *et al.* Cortical superficial siderosis, hematoma volume, and outcomes after intracerebral hemorrhage: a mediation analysis. *Front Neurol* 2023;14.
- 18 Fazekas F, Chawluk JB, Alavi A, *et al.* MR signal abnormalities at 1.5 T in Alzheimer's dementia and normal aging. *AJR Am J Roentgenol* 1987;149:351–6.
- 19 Jin Y, Huang YH, Chen YP, *et al.* Combined effect of cortical superficial siderosis and cerebral microbleed on short-term and long-term outcomes after intracerebral haemorrhage. *Stroke Vasc Neurol* 2024;9:429–38.
- 20 Meretoja A, Strbian D, Putaala J, *et al.* SMASH-U: a proposal for etiologic classification of intracerebral hemorrhage. *Stroke* 2012;43:2592–7.
- 21 Krzyżiński M, Spytek M, Baniecki H, *et al.* SurvSHAP(t): Time-dependent explanations of machine learning survival models. *Knowl Based Syst* 2023;262:110234.
- 22 Kovalev MS, Utikin LV, Kasimov EM. SurvLIME: A method for explaining machine learning survival models. *Knowl Based Syst* 2020;203:106164.
- 23 García RI, Ibrahim JG, Zhu H. Variable selection in the cox regression model with covariates missing at random. *Biometrics* 2010;66:97–104.
- 24 Poon MTC, Fonville AF, Al-Shahi Salman R. Long-term prognosis after intracerebral haemorrhage: systematic review and meta-analysis. *J Neurol Neurosurg Psychiatry* 2014;85:660–7.
- 25 Gregório T, Pipa S, Cavaleiro P, *et al.* Prognostic models for intracerebral hemorrhage: systematic review and meta-analysis. *BMC Med Res Methodol* 2018;18:145.
- 26 Peng SY, Chuang YC, Kang TW, *et al.* Random forest can predict 30-day mortality of spontaneous intracerebral hemorrhage with remarkable discrimination. *Eur J Neurol* 2010;17:945–50.
- 27 Lin F, He Q, Zhuo L, *et al.* A nomogram predictive model for long-term survival in spontaneous intracerebral hemorrhage patients without cerebral herniation at admission. *Sci Rep* 2023;13:3126.
- 28 Pasi M, Casolla B, Kyheng M, *et al.* Long-term mortality in survivors of spontaneous intracerebral hemorrhage. *Int J Stroke* 2021;16:448–55.
- 29 Loan JJ, Gane AB, Middleton L, *et al.* Association of baseline hematoma and edema volumes with one-year outcome and long-term survival after spontaneous intracerebral hemorrhage: A community-based inception cohort study. *Int J Stroke* 2021;16:828–39.
- 30 Mosconi MG, Paciaroni M, Agnelli G, *et al.* SMASH-U classification: a tool for aetiology-oriented management of patients with acute haemorrhagic stroke. *Intern Emerg Med* 2021;16:109–14.
- 31 Jia Y, Li G, Song G, *et al.* SMASH-U aetiological classification: A predictor of long-term functional outcome after intracerebral haemorrhage. *Eur J Neurol* 2022;29:178–87.
- 32 Kuramatsu JB, Gerner ST, Lücking H, *et al.* Anemia is an independent prognostic factor in intracerebral hemorrhage: an observational cohort study. *Crit Care* 2013;17:R148.
- 33 Acosta JN, Leasure AC, Kuohn LR, *et al.* Admission Hemoglobin Levels Are Associated With Functional Outcome in Spontaneous Intracerebral Hemorrhage. *Crit Care Med* 2021;49:828–37.
- 34 Roh DJ, Boehme A, Mamoon R, *et al.* Relationships of Hemoglobin Concentration, Ischemic Lesions, and Clinical Outcomes in Patients With Intracerebral Hemorrhage. *Stroke* 2023;54:1021–9.
- 35 Wolters FJ, Zonneveld HI, Licher S, *et al.* Hemoglobin and anemia in relation to dementia risk and accompanying changes on brain MRI. *Neurology (Ecricon)* 2019;93:e917–26.
- 36 Charidimou A, Gang Q, Werring DJ. Sporadic cerebral amyloid angiopathy revisited: recent insights into pathophysiology and clinical spectrum. *J Neurol Neurosurg Psychiatry* 2012;83:124–37.
- 37 Schrag M, Kirshner H. Management of Intracerebral Hemorrhage: JACC Focus Seminar. *J Am Coll Cardiol* 2020;75:1819–31.
- 38 Mengel A, Siokas V, Buesink R, *et al.* Continuous Blood Pressure Indices During the First 72 Hours and Functional Outcome in Patients with Spontaneous Intracerebral Hemorrhage. *Neurocrit Care* 2024.
- 39 Cordonnier C, Demchuk A, Ziai W, *et al.* Intracerebral haemorrhage: current approaches to acute management. *Lancet* 2018;392:1257–68.
- 40 Ma L, Hu X, Song L, *et al.* The third Intensive Care Bundle with Blood Pressure Reduction in Acute Cerebral Haemorrhage Trial (INTERACT3): an international, stepped wedge cluster randomised controlled trial. *Lancet* 2023;402:27–40.