Check for updates

OPEN ACCESS

EDITED BY Hsiao-Lung Chan, Chang Gung University, Taiwan

REVIEWED BY Amirmasoud Ahmadi, Max Planck Institute for Biological Intelligence, Germany Sushma Jaiswal, Guru Ghasidas Vishwavidyalaya, India Muhana Fawwazy Ilyas, Sebelas Maret University, Indonesia

*CORRESPONDENCE Dong-Wha Kang I dwkang@amc.seoul.kr Sun U. Kwon I sukwon@amc.seoul.kr

RECEIVED 31 January 2025 ACCEPTED 26 June 2025 PUBLISHED 25 July 2025

CITATION

Namgung E, Kim YS, Kwon SU and Kang D-W (2025) Predicting cognitive decline in cognitively impaired patients with ischemic stroke with high risk of cerebral hemorrhage: a machine learning approach. *Front. Neurol.* 16:1569073. doi: 10.3389/fneur.2025.1569073

COPYRIGHT

© 2025 Namgung, Kim, Kwon and Kang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic

practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Predicting cognitive decline in cognitively impaired patients with ischemic stroke with high risk of cerebral hemorrhage: a machine learning approach

Eun Namgung¹, Young Sun Kim², Sun U. Kwon³* and Dong-Wha Kang^{2,3}*

¹Asan Institute for Life Sciences, Asan Medical Center, Seoul, Republic of Korea, ²Nunaps Inc., Seoul, Republic of Korea, ³Department of Neurology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea

Background and objective: Cognitive decline progresses rapidly in stroke patients, increasing risks of stroke recurrence. Predicting deterioration within a year in patients with poststroke cognitive impairment (PSCI) could guide targeted interventions for dementia prevention and better prognosis. In this Preventlon of CArdiovascular events in iSchemic Stroke patients with high risk of cerebral hemOrrhage for reducing cognitive decline substudy, machine learning on clinical and imaging data was used to predict cognitive decline over 9 months in PSCI patients.

Methods: This retrospective study included 109 patients with acute ischemic stroke and high-risk cerebral hemorrhage with PSCI (baseline Korean-Mini Mental Status Examination [K-MMSE] < 24), along with baseline clinical imaging and K-MMSE assessments at baseline and after 9 months. Four machine learning algorithms were trained, Categorical Boosting (CatBoost), Adaptive Boosting (AdaBoost), eXtreme Gradient Boosting (XGBoost), and logistic regression, to predict cognitive decliners, defined as a decline of \geq 3 K-MMSE points over 9 months, and ranked variable importance using the SHapley Additive exPlanations methodology.

Results: CatBoost outperformed the other models in classifying cognitive decliners within 9 months. In the test set, CatBoost achieved a mean area under the curve (AUC) of 0.897, with an accuracy of 0.873; other models performed as follows: logistic regression (AUC 0.775), AdaBoost (AUC 0.767), and XGBoost (AUC 0.722). Higher baseline K-MMSE scores (total, language, orientation to place, and recall), longer interval between stroke and baseline K-MMSE, initial National Institutes of Health Stroke Scale scores, and lesion volume ratio were identified as key predictors of cognitive decline in CatBoost. Cognitive decliners showed longer interval between stroke onset and pharmacotherapy initiation than non-decliners.

Conclusion: CatBoost effectively recognized patients with ischemic stroke at high risk of cognitive decline over 9 months. Recognizing these high-risk individuals and their risk and protective factors allows for timely and targeted interventions to improve prognosis in PSCI patients.

KEYWORDS

machine learning, cognitive decline, ischemic stroke, cerebral hemorrhage, poststroke cognitive impairment

1 Introduction

Stroke, which causes long-term disability, is a global health challenge (1). Cognitive impairment frequently follows a stroke, remarkably exacerbating disability and quality of life (2, 3). Individuals who have experienced a stroke typically show rapider decline in cognitive functions, increased risk of recurrent strokes, and higher mortality rates (4, 5). Recurrent strokes are considered strong predictors of cognitive decline, with affected patients exhibiting higher rates of dementia than those with a first-ever stroke (6, 7).

Higher risk of poststroke cognitive impairment (PSCI) is related to ischemic stroke with intracerebral hemorrhage or small vessel occlusive disease, requiring optimal secondary prevention (8, 9). Although conventional antiplatelet or statin therapy may increase hemorrhage risks (10, 11), cilostazol (a phosphodiesterase-3 inhibitor) or probucol (a non-statin lipid-lowering agent) may reduce the incidence of cardiovascular events without increasing hemorrhage risk (12–14). Preserving cognitive function and preventing dementia are vital for patients with ischemic stroke and high risk of cerebral hemorrhage, particularly in Asian populations due to regional stroke differences (15, 16).

Recent meta-analyses highlight the multifactorial nature of PSCI, involving vascular, demographic, and neuroanatomical factors (17). PSCI typically manifests within 3–6 months poststroke, with recovery often plateauing beyond this period (18, 19). Given the heterogeneity in cognitive trajectories, accurate prediction of cognitive deterioration beyond this window could support the development of personalized interventions aimed at preventing dementia and enhancing quality of life (17, 20). In this context, machine learning algorithms offer a promising approach by capturing non-linear interactions among baseline variables—such as stroke severity, white matter hyperintensity, and initial cognitive status—and identifying key predictors through feature importance analysis (20, 21).

In this substudy of PreventIon of CArdiovascular events in iSchemic Stroke patients with high risk of cerebral hemOrrhage for reducing COGnitive decline (PICASSO-COG) (12, 22), we propose machine learning algorithms leveraging clinical and imaging data to predict cognitive decline over a 9-month period in patients with acute ischemic stroke with cognitive impairment and high risk of cerebral hemorrhage – a population that requires timely prevention and intervention strategies.

2 Materials and methods

2.1 Participants and study design

This retrospective analysis focused on a subset of PICASSO-COG substudy, which evaluated the effects of cilostazol and/or probucol on cognitive functions in patients with ischemic stroke and high risk of cerebral hemorrhage from the PICASSO cohort (22). PICASSO (PreventIon of CArdiovascular events in iSchemic Stroke patients with high risk of cerebral hemOrrhage) trial is a multicenter, randomized, double-blind, placebo-controlled 2 × 2 factorial trial that compared the efficacy and safety of cilostazol versus aspirin, with and without probucol, for preventing hemorrhagic stroke and major vascular events ischemic in these patients (ClinicalTrials.gov, no. NCT01013532) (12). PICASSO study was approved by the site ethics committees and conducted according to Good Clinical Practice and the Declaration of Helsinki, with written consent obtained from all participants.

Key inclusion criteria for the PICASSO cohort were (1) age >20 years; (2) non-cardioembolic ischemic stroke or transient ischemic attack within the 180 days prior to screening; (3) previous intracerebral hemorrhage or multiple cerebral microbleeds based on clinical or radiological findings; and (4) asymptomatic intracerebral hemorrhage identified as a slit-like curvilinear lesion on magnetic resonance imaging, with no obvious history of intracerebral hemorrhage within the past 6 months, contraindications to long-term antiplatelet therapy, severe cardiomyopathy or heart failure, and recent myocardial infarction or coronary procedures within the previous 4 weeks (12).

Figure 1 shows the subject flow diagram. From the PICASSO cohort, 892 patients were included in the PICASSO-COG cohort after excluding those unable to undergo cognitive testing due to severe dysphasia or neurological deficits. Of these, 376 patients both the baseline evaluation (3–7 months post-stroke) and the first follow-up evaluation (\geq 9 months post-stroke) using the Korean Mini-Mental Status Examination (K-MMSE). Baseline Fluid-Attenuated Inversion Recovery (FLAIR) MRI scans for quantifying stroke lesion volume ratios were available for 376 of these patients. The final analysis included 109 patients with PSCI, defined by baseline K-MMSE scores <24 (23). No demographic, clinical, or imaging data were missing in the final analytic sample. Patients with incomplete cognitive assessments or missing imaging data were excluded during cohort selection; thus, imputation was not required.

2.2 Demographic and clinical assessment

We evaluated baseline characteristics of the participants, including demographics and clinical data (Table 1). Demographics included age, gender, and years of education, and clinical data included vital signs, lipid levels, blood glucose levels, and smoking history. We also included the concomitant pharmacotherapy assigned in the PICASSO trial (cilostazol vs. aspirin with probucol vs. no probucol) and the time between stroke onset and randomization in the PICASSO trial. Stroke-related factors included time since stroke onset, classification of ischemic events, and a high-risk index for intracerebral hemorrhage, which encompassed a history or radiological evidence of intracerebral hemorrhage and multiple microbleeds. Stroke severity was assessed using the National Institutes of Health Stroke Scale (NIHSS) score at admission (24). Fazekas scores, which indicate the extent of white matter hyperintensities (25), and lesion volume ratio, assessed by a neuroradiologist based on FLAIR images, were included as strokerelated imaging features. Volume of ischemic stroke lesions on



baseline FLAIR images was automatically quantified using the lesion prediction algorithm, implemented in the Lesion Segmentation Tool (LST) within SPM12 (26). The algorithm generated a lesion probability map in which each voxel was assigned a probability of being a lesion. Voxels with a probability \geq 0.5 were considered lesional. The total lesion volume (cm³) was calculated by multiplying the number of suprathreshold voxels by the spatial resolution of the scan. To normalize for interindividual differences in brain size, the lesion volume ratio was computed by dividing the segmented lesion volume by the total intracranial volume.

2.3 Cognitive impairment assessment

Our study focused on patients with acute ischemic stroke, cognitive impairment, and a high risk of cerebral hemorrhage— a population in urgent need of timely prevention and intervention strategies. The K-MMSE, administered by a certified neurologist, was used at baseline and follow-up to assess cognitive impairment and its progression, with total scores ranging from 0 to 30, where lower scores indicate greater impairment. Scores <24 indicated cognitive impairment (23). We focused on patients with acute ischemic stroke, cognitive impairment, and a high risk of cerebral hemorrhage— a population that requires timely prevention and intervention strategies.

The K-MMSE evaluated the following seven domains: orientation to time, orientation to place, registration, attention and calculation, memory recall, language, and visuospatial ability (23). The total scores and seven domain subscores of the K-MMSE at baseline are presented in Table 1.

Baseline K-MMSE was conducted 3–7 months after stroke onset (mean \pm standard deviation, 4.7 \pm 0.6 months; minimum–maximum, 3.5–6.6). The follow-up K-MMSE, conducted after 9 months (9.3 \pm 1.0 months; minimum–maximum, 8.3–13.3), was between 12 and 20 months after stroke onset (14.0 \pm 1.2 months; minimum– maximum 12.2–19.3). A \geq 3-point decline in K-MMSE total scores over 9 months indicated cognitive decline, according to studies suggesting significant MMSE changes of four points over 5 years and reliable annual changes of 1.3–2.7 points (27, 28).

2.4 Training and testing

Participants were randomly divided into a training (n = 87) set and a test (n = 22) set in an 8:2 ratio, with no significant differences observed in baseline characteristics (Table 2). Patients were categorized into two groups for labeling: those with a decrease of \geq 3 K-MMSE points over 9 months received a positive label (cognitive decliners, n = 27, 24.7%), whereas those with a < 3-point decrease in TABLE 1 Baseline characteristics between cognitive decliners and cognitive non-decliners.

(n = 27)(n = 82)Age (sens)68.2.1.797.2.3.1.8.76.0.7.0Fancia65.5.5.624.0.6.40.000Elacatiny years64.4.1.5.24.3.1.4.20.0.80Nothis bettern stoke out and buedne K-MOSE9.9.1.84.4.1.8.20.2.1.6.9K MSR near shadine23.0.1.4.17.7.9.1.3.84.001Orientation tonis out and buedne K-MOSE4.9.1.84.4.1.8.10.001C MSR near shadine4.5.1.13.0.1.1.60.004*Orientation tonis outs out and buedne K-MOSE4.7.1.6.64.4.1.1.00.004*Orientation tonis outs out and buedne K-MOSE2.2.9.40.9.7.1.6.60.001*Robal outs outs out and buedne K-MOSE2.2.9.40.001*0.001*Naterian and outsiton tonis outs6.5.1.1.60.0.1.40.001*Naterian and outsiton outsiton0.5.9.1.10.0.1.6.10.0.0.1.1Naterian and outsiton tonis outs0.5.9.1.10.0.1.10.0.1.1Naterian and outsiton outsiton0.5.9.1.10.0.1.10.0.1.1Internation outsiton outsiton0.5.9.1.10.0.1.10.0.1.1Naterian and outsiton outsiton0.5.9.1.10.0.1.10.0.1.1Internation outsiton outsiton0.5.9.1.10.0.1.10.0.1.1	Baseline characteristics	Decliners	Non-decliners	p	
Ag (yana)(92257)723 5 87(0.07)Fanale13555022(8.4)8.39Fanale view allow145124.34.200.000Follow grinario (rearch)9.54.119.24.090.331Menha brives attok const and bashine KAMSR432.0845.61.600.001*Constation to ine subscore142.1330.61.640.001*Orientation base subscore127.26.641.11.00.001*Orientation base subscore27.24.6442.77.060.001*Recall subscore12.12.160.001*0.001*Recall subscore12.12.160.001*0.001*Recall subscore13.1113.11.113.01.160.001*Recall subscore13.1113.11.10.001*0.001*Langage subscore13.11.113.11.10.001*0.001*Vinsopsital Maly subscore13.11.113.11.10.001*0.001*Namespital Maly subscore13.11.113.11.10.001*0.001*113.11.113.11.10.001*0.001*0.001*113.1213.12.10.001*0.001*0.001*113.1213.12.10.001*0.001*0.001*113.11.113.12.10.001*0.001*113.1213.12.10.001*0.001*113.1213.12.10.001*0.001*113.1213.12.10.001*0.001*113.1213.12.10.001*0.001*1 <th></th> <th>(n = 27)</th> <th>(<i>n</i> = 82)</th> <th></th>		(n = 27)	(<i>n</i> = 82)		
Fend151550920340.000Baccio yan64.12520.000Baccio yan64.12364.1240.000Bacter abundan64.02464.0240.000Construction to abundan C MMNE42.02417.023.830.000*Construction to abundan22.024.413.0.140.000*Orientation to instore64.0133.0.140.000*Registration subscore22.02427.040.000*Registration subscore22.02427.040.000*Registration subscore60.0145.0140.000*Registration subscore60.0145.0140.000*Registration subscore60.0145.0120.000*Registration subscore60.0145.012.0140.000*Registration subscore60.0145.012.0140.000*Registration subscore60.0145.012.0140.000*Registration subscore60.0143.012.0140.000*Registration subscore60.0143.012.0140.000*Registration subscore10.0700.02.0140.000*Registration subscore10.0700.010.0140.000*Registration subscore10.0100.02.0140.000*Registration subscore10.0100.02.010.000*Registration subscore10.0100.02.010.010*Registration subscore10.0100.010*0.010*Registration subscore10.0100.010*0.010*Registration subscore10.010*	Age (years)	69.2 ± 7.9	72.3 ± 8.7	0.070	
Induction (nameline)6.4.1.5.24.3.1.4.20.000Relow-quantion (nameline)9.9.1.1.39.2.1.0.90.2.3.3Northan between stock over and alkadine K-MMSE1.9.2.9.4.11.7.9.2.1.8K-MMSE cover structures1.3.0.1.1.60.0.0.0.1.1Orientation to time subscore4.4.0.1.1.30.1.1.60.0.0.1.1Orientation to place subscore2.9.2.0.42.7.9.6.40.0.0.1.1Regarding subscore2.9.2.0.40.0.0.1.10.0.0.1.1Regarding subscore2.9.2.0.40.0.0.1.10.0.0.1.1Attention and calculation subscore0.9.2.1.2.10.00.1.10.0.0.1.1Regard subscore0.9.2.2.1.40.0.0.1.10.0.0.1.1Attention and calculation subscore0.9.2.2.1.10.0.0.1.10.0.0.1.1Regard subscore0.9.2.2.1.10.0.1.10.0.1.10.0.1.1Viscogatal ability subscore0.9.2.2.1.10.0.1.10.0.1.1Viscogatal ability subscore0.9.2.2.1.10.0.1.10.0.1.1Viscogatal ability subscore0.9.2.2.1.10.0.1.10.0.1.1Viscogatal ability subscore1.1.1.92.1.1.1.90.0.1.1Viscogatal ability subscore1.1.1.90.1.1.1.10.0.1.1Viscogatal ability subscore1.1.1.90.0.1.1.10.0.1.1Viscogatal ability subscore1.1.1.90.0.1.1.10.0.1.1Viscogatal ability subscore1.1.1.90.0.1.1.10.0.1.1Viscogatal ability subscore1.1.1.90.1.1.1.10.0.1.1Viscogatal abil	Female	15 (55.6)	52 (63.4)	0.500	
Pellow up duration (month)0.9.9.130.9.2.1.090.0.3.1Month sevens stuck onat and baseline K-MMSE4.9.9.1084.6.9.6.000.000.10Construction to me subscore0.4.0.1.1.00.000.100.000.10Orientation to me subscore0.4.7.1.0.60.0.1.1.00.000.10Registration and accluitor stuckere0.2.1.1.1.80.0.9.1.1.00.0.0.1.1.0America and scalation stuckere0.1.2.1.1.1.00.0.0.1.1.00.0.0.1.1.0America and scalation stuckere0.1.2.1.1.1.00.0.0.1.1.00.0.0.1.1.0America and scalation stuckere0.2.1.1.1.00.0.0.1.1.00.0.0.0.1.0America and scalation stuckere0.2.1.1.0.00.0.0.0.1.00.0.0.0.1.0NIRSS once and scalation stuckere0.2.1.1.1.00.0.0.0.1.00.0.0.0.0NIRSS once and scalation stuckere0.0.0.1.0.00.0.0.0.00.0.0.0.011.0.1.0.00.0.0.0.00.0.0.0.00.0.0.0.011.0.1.0.00.0.0.0.00.0.0.0.00.0.0.0.011.0.1.0.00.0.0.0.00.0.0.0.00.0.0.0.011.0.1.0.00.0.0.0.0.0.00.0.0.0.0.00.0.0.0.0.0.011.0.1.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0	Education years	6.4 ± 5.2	4.3 ± 4.2	0.080	
Months between stroke onset and baseline K-MMSE4.9 ± 0.84.4 ± ± 0.80.6 ± 0.00 °K-MMSE score at baseline2.5 ± 2.4 11.7 ± 2.8 d0.000 °Orientation tine nalwone4.4 ± ± 1.00.000 °0.000 °Regizration indocore2.9 ± 0.4 12.7 ± 0.6 d0.011 °Artention and calculation subscore1.9 ± 1.1 10.012 °0.000 °Regizration indocore1.9 ± 1.1 11.1 ± 1.10.012 °Artention and calculation subscore0.6 ± 2.1 ± 1.80.6 ± 1.2 00.000 °Regizration indocore0.6 ± 2.1 ± 1.92.1 ± 2.0 0.00 °0.000 °Vanospizita Malking valscore0.6 ± 2.6 50.2 ± 0.4 0.000 °0.000 °Nutsongizita Malking valscore0.001 °0.2 ± 0.4 0.000 °0.001 °Nutsongizita Malking valscore0.012 °0.011 °0.011 °Lesion ratio in the whole brain (N)2.1 ± 1.9 °0.011 °0.010 °11.1 ∴ 1.1 °0.010 °1.1 ∴ 1.1 °0.000 °11.1 ∴ 1.1 °0.010 °0.010 °0.010 °11.1 ∴ 1.1 °0.010 °1.1 ∴ 1.1 °0.000 °11.1 ∴ 1.1 °0.010 °0.010 °0.010 °11.1 ∴ 1.1 °0.010 °0.010 °0.010 °11.1 ∴ 1.1 °0.010 °0.010 °0.010 °11.1 ∴ 1.1 °0.010 °0.010 °0.010 °11.1 ∴ 1.1 °0.010 °0.010 °0.010 °11.1 ∴ 1.1 °0.010 °0.010 °0.010 °1 <td>Follow-up duration (months)</td> <td>9.5 ± 1.3</td> <td>9.2 ± 0.9</td> <td>0.293</td>	Follow-up duration (months)	9.5 ± 1.3	9.2 ± 0.9	0.293	
K-MMSE acore at baseline230 ± 4.11.79 ± 3.84.0001*Orientation bylase valuescor1.40 ± 1.33.0 ± 1.60.000*Orientation bylase valuescor1.72 ± 0.64.12 ± 1.00.000*Registration subscore2.92 ± 0.42.72 ± 0.60.011*Attention and calculation audescor1.11 ± 1.80.92 ± 1.20.000*Becall subscore1.52 ± 1.40.019*0.019*Language subscore0.52 ± 5.50.22 ± 0.40.000*NURSS accon calculation subscore0.52 ± 5.50.22 ± 0.40.000*NURSS accon calculation subscore1.52 ± 1.60.000*0.014Language subscore0.52 ± 5.50.22 ± 0.40.000*NURSS accon calculation subscore1.52 ± 1.60.000*0.014Language subscore1.52 ± 1.60.0100.014Language subscore1.63 ± 0.02 ± 0.40.000*NURSS accon calculation in the whole brain (%)2.5 ± 1.80.017Jacon ration in the whole brain (%)1.63 ± 0.100.000*11.61 ± 0.100.000*1.61 ± 0.000*11.61 ± 0.000*1.61 ± 0.000*0.01411.61 ± 0.000*1.61 ± 0.000*0.000*11.62 ± 0.000*1.61 ± 0.000*0.000*11.61 ± 0.000*1.61 ± 0.000*0.000*11.62 ± 0.000*1.61 ± 0.000*0.000*11.62 ± 0.000*1.61 ± 0.000*0.000*11.61 ± 0.000*1.61 ± 0.000*0.000*11.61 ± 0.00	Months between stroke onset and baseline K-MMSE	4.9 ± 0.8	4.6 ± 0.6	0.089	
Orientation to time solucore4.0 ± 1.33.0 ± 1.60.064*Orientation to place subscore4.7 ± 0.64.1 ± 1.00.008*Regitariano subscore2.9 ± 0.42.7 ± 0.60.011*Attention and adculation subscore2.1 ± 1.80.9 ± 1.20.001*Recil subscore1.9 ± 1.11.5 ± 1.10.008*Languag subscore0.6 ± 1.65.6 ± 1.44.000 *Viscospatial ability subscore0.6 ± 0.50.2 ± 0.40.008*Viscospatial ability subscore0.5 ± 0.50.2 ± 0.40.008*Viscospatial ability subscore0.5 ± 0.50.2 ± 0.40.008*Viscospatial ability subscore0.5 ± 0.50.2 ± 0.40.008*Viscospatial ability subscore0.1 1.70.010.914Lastin artio in be whole hard (%)2.1 ± 1.90.118.0326Functas score10 (07.0)34 (41.5)0.463210 (07.0)34 (41.5)0.463210 (07.0)34 (41.5)0.999210 (07.0)34 (41.5)0.999210 (07.0)9.0 (76.7)0.999Trinstend reverts10 (17.0)8.0 (77.3)0.999Hobers crystore ord plarmacotherapg5.16.5)14 (17.1)8.0 (9.7)Radiological finding of intracerbral benorrhage5.16.5)14 (27.1)0.821Radiological finding of intracerbral benorrhage5.16.5)14 (47.3)0.03*Clistazel (10 migday)16 (59.3)14 (52.1)0.821Abiting of probacel (25 migday)<	K-MMSE score at baseline	23.0 ± 4.1	17.9 ± 3.8	<0.001*	
Orientation to place abscore4.4 ± 0.64.1 ± 1.00.008*Registration subscore2.9 ± 0.42.7 ± 0.60.011*Atternito and calculation subscore2.1 ± 1.80.0 ± 1.20.001*Recall subscore0.6 ± 1.65.6 ± 1.40.6 0.01*Langage subscore0.6 ± 1.65.6 ± 1.40.6 0.01*Visuospital billy subscore0.5 ± 1.65.6 ± 1.40.008*Visuospital billy subscore0.2 ± 1.83.2 ± 0.40.008*INTSS vore and simision1.1 ± 1.90.12 ± 0.00.32 ± 0.2Paretars score0.1 (3.7)0.00)0.2 ± 0.40.32 ± 0.221.1 (3.7)0.00)3.4 ± 0.20.63 ± 0.221.1 (3.7)0.00)3.4 ± 0.20.63 ± 0.221.1 (3.7)0.00)3.4 ± 0.20.63 ± 0.221.1 (3.7)0.00 ± 0.4 ± 0.20.63 ± 0.221.1 (3.7)0.60 (9.6)3.4 ± 0.20.63 ± 0.231.3 ± 0.2	Orientation to time subscore	4.0 ± 1.3	3.0 ± 1.6	0.004*	
Registration subscore2.9 ± 0.42.7 ± 0.60.011*Artention and calculation subscore2.1 ± 1.80.9 ± 1.10.018*Becall subscore1.9 ± 1.11.3 ± 1.10.018*Languag subscore0.5 ± 5.40.400.40.000*Visuogatial ability subscore0.5 ± 5.40.200.40.000*Visuogatial ability subscore0.5 ± 5.40.21 ± 1.92.1 ± 2.1 ± 0.30.33 ± 1.80.35 ± 0.	Orientation to place subscore	4.7 ± 0.6	4.1 ± 1.0	0.008*	
Attention and calculation subscore2.1 ± 1.80.9 ± 1.10.01*Reall subscor1.9 ± 1.11.1 ± 1.11.0 ± 1.8Language subscore6.9 ± 1.65.6 ± 1.440.01*Visuopatial shifty subscore0.5 ± 0.50.2 ± 0.40.006*NIRS score on admission2.1 ± 1.92.1 ± 2.00.914Lesion ratio in the whole brain (%)0.1 ± 1.10.010.026Tackas score0.010.000.0262 cackas score0.010.000.02613.1 (1.1)0.90 (0.1)0.000.02620.01 (3.01)0.90 (0.1)0.000.0020.01 (3.01)0.90 (0.1)0.000.0030.01 (3.01)0.00 (0.1)0.000.0030.00 (0.1)0.00 (0.1)0.000.0011.6 text0.00 (0.1)0.00 (0.1)0.0011.6 text0.00 (0.1)0.00 (0.1)0.0011.6 text0.00 (0.1)0.00 (0.1)0.00 (0.1)11.6 text0.00 (0.1)0.00 (0.1)0.00 (0.1) <t< td=""><td>Registration subscore</td><td>2.9 ± 0.4</td><td>2.7 ± 0.6</td><td>0.011*</td></t<>	Registration subscore	2.9 ± 0.4	2.7 ± 0.6	0.011*	
Recall subscore19 ± 1.11.1 ± 1.10.018*Langag subscore6.6 ± 1.65.5 ± 1.4<0001*	Attention and calculation subscore	2.1 ± 1.8	0.9 ± 1.2	0.001*	
Languag subscore6.6 ± 1.65.6 ± 1.4.0001*Vacuoquial ability subscore6.0 ± 1.6 56.0 ± 1.6 40.006*NIHSS score on admission2.0 ± 1.6 40.0 ± 1.6 ± 0.0	Recall subscore	1.9 ± 1.1	1.3 ± 1.1	0.018*	
Vancopatial ability subscore0.05 ± 0.50.2 ± 0.40.006°NHESS score on admission2.1 ± 1.92.1 ± 2.00.914Lesion ratio in the whole brain (%)2.9 ± 1.83.3 ± 1.80.326Packas score11 (3.7)0 (0)113 (11.1)8 (98.0)3 (44.15)210 (07.0)34 (41.5)3313 (48.2)40 (48.8)-46.89113 (48.2)40 (48.8)-46.99113 (48.2)40 (48.8)-46.99114 (17.1)80 (97.6)-40.99114 (17.1)80 (97.6)-40.99114 ± 1.510 4 ± 1.50.821114 ± 1.510 4 ± 1.00.821111 (40.7)36 (45.9)-40.99111 (40.7)36 (45.9)-40.99111 (40.7)36 (45.9)-40.92110 (10.0)15 (5.5)41 (2.4)-0.827111 (40.7)36 (45.9)-40.92110 (10.0)15 (5.5)41 (2.4)-40.92110 (10.2)10 (10.2)-40.92110 (10.2)10 (10.2)-40.92115 (5.5)13 (5.6)13 (5.2)-40.92115 (5.5)13 (5.6)13 (5.2)-40.92114 ± 1.510 (10.2)-40.92114 ± 1.510 (10.2)-40.92114 ± 1.510 (10.2)-40.92115 (5.5)14 (5.2)-40.921 <td>Language subscore</td> <td>6.9 ± 1.6</td> <td>5.6 ± 1.4</td> <td><0.001*</td>	Language subscore	6.9 ± 1.6	5.6 ± 1.4	<0.001*	
NHSS score on edmission2.1 ± 1.92.1 ± 2.00.914Lesion ratio in the whole brain (%)2.9 ± 1.83.3 ± 1.80.326Parkats score1(3.7)0 (0)13 (11.1)8 (9.80)	Visuospatial ability subscore	0.5 ± 0.5	0.2 ± 0.4	0.006*	
Leson ratio in the whole brain (%)2.9 ± 1.83.3 ± 1.80.326Fackas score011(3.7)0(0)113(11.1)8 (9.80)0.463210(0'5.0)34 (4.15.0)0.46331.13 (48.2)40 (48.8)0.46331.3 (48.2)40 (48.8)0.402131.3 (48.2)40 (48.8)0.40211.6 (19.2)1.6 (19.2)0.6020.60211.6 (19.2)1.6 (19.2)0.6020.60211.6 (19.2)1.6 (19.2)0.6020.60211.6 (19.2)1.6 (19.2)0.6120.60311.6 (19.2)1.6 (19.2)0.6230.62311.6 (19.3)4.6 (19.1)0.6230.62311.6 (19.2)1.6 (19.2)0.6230.62311.6 (19.3)4.6 (19.1)0.6230.62311.6 (19.3)4.6 (19.1)0.6230.62311.6 (19.3)4.6 (19.1)0.6230.62311.6 (19.3)1.6 (19.1)0.6 (19.1)0.62311.6 (19.3)1.6 (19.1)0.6 (19.1)0.61411.6 (19.1)1.6 (19.1)1.6 (19.1)0.6 (19.1)11.6 (19.1)1.6 (19.1)1.6 (19.1)0.6 (19.1)11.6 (19.1)1.6 (19.1)1.6 (19.1)0.6 (19.1)11.6 (19.1)1.6 (19.1)1.6 (19.1)0.6 (19.1)11.6 (19.1)1.6 (19.1)1.6 (19.1)0.6 (NIHSS score on admission	2.1 ± 1.9	2.1 ± 2.0	0.914	
Interclase score Interclase score 0 1.0.37 0.00 1 3.01.10 8.08.00 2 0.0370 34.(41.5) 3 0.13 (48.2) 0.00 (40.00) Ischemic events 0.00 2.(24.00) Index intscender 0.00 2.(24.00) High-risk index intracerebral hemorrhage 0.00 2.(24.00) High-risk index intracerebral hemorrhage 5.(18.5) 14.(17.1) Radiogical findings of intracerebral hemorrhage 5.(18.5) 1.4.(17.1) Radiogical findings of intracerebral hemorrhage 5.(18.5) 1.4.(17.1) Multiple microbleeds 1.07.63.00 4.02.12.56.00 Multiple microbleeds 1.07.63.00 4.03.00 Cilotazol or ageirin therapy 1.4.1.5 0.8.21.00 Cilotazol or angeirin therapy 1.1.6.07.00 3.65.63.00 Addition of probucio (250 mg/day) 1.5.(55.6) 4.3.(52.4) 0.0.327 Current smoking stuts 1.5.(56.00 3.13.52.16.70 0.43.12.10 Quit smoking in the pat 3 years 2.(7.4) 6.66.3.	Lesion ratio in the whole brain (%)	2.9 ± 1.8	3.3 ± 1.8	0.326	
01100013(1.1)8(9.80)0.463210(57.0)34(41.5)0.463313(42.2)40(48.8)0113(42.2)40(48.8)0113(42.2)40(48.8)0113(42.2)80(97.6)2(.24)2(.9997710(0)80(97.6)2(.99971314(17.1)3(.15)14(17.1)3(.999114/15.314/17.1)14/15.30.0217131314/15.30.1213(.16)713/16.3014/17.1)14/15.30.131817(63.0)47(57.3)0.3213(.16)111/16.3014/15.30.1313(.12)111/16.3014/15.30.1323(.12)111/16.3014/15.30.1323(.12)111/16.3014/15.30.1323(.12)111/16.3014/15.314/15.33(.12)111/16.3011/16.314/15.33(.12)111/16.3015/15.6)14/15.33(.12)3(.12)115/15.6)15/15.6)14/15.33(.12)3(.12)115/15.6)15/15.6)15/15.6)3(.12)3(.12)115/15.6)15/15.6)15/15.6)3(.12)3(.12)115/15.6)15/15.6)15/15.6)3(.12)3(.12)115/15.6)15/15.6)15/15.6)3(.12)<	Fazekas score				
13 (11.1)8 (980)0.43210 (37.0)34 (41.5)0.43313 (43.2)40 (48.0)0.681schemic events13 (43.2)40 (68.0)-0.999Ischemic stroke27 (100)80 (97.6)2.0.999Transient ischemic event0 (0)2 (2.4)-0.999High-risk index intracerebral hemorrhage5 (18.5)14 (17.1)ARadiologial findings of intracerebral hemorrhage5 (18.5)14 (17.1)AMultipe microbleeds17 (63.0)47 (57.3)0.033*Clostarol or aprim therapy14 ± 1.50.8 ± 1.00.033*Clostarol or aprim therapy11 (40.7)36 (43.9)-0.826Addition of probucol or none12 (24.4)30 (45.6)-0.827Addition of probucol C50 mg/day)15 (55.6)43 (52.4).0.827No probucol12 (24.4)30 (47.6)-0.827Current smoking strats2 (7.4)56 (68.3)	0	1 (3.7)	0 (0)		
211100313 (42.)40 (48.8)00113 (48.2)40 (48.8)00113 (48.2)40 (48.8)00113 (48.2)40 (48.8)00110 (0)2 (2.4) 20.999 110 (0)2 (2.4)0110 (0)2 (2.4)011 <t< td=""><td>1</td><td>3 (11.1)</td><td>8 (9.80)</td><td></td></t<>	1	3 (11.1)	8 (9.80)		
311 (48.2)04 0 (48.8)313 (48.2)00 (48.8)Ischemic events27 (100)80 (97.6)Tansient ischemic event0 (0)2 (2.4)High-risk index intracerbal hemorrhage0 (0)2 (2.4)History of intracerbal hemorrhage5 (18.5)14 (17.1)Radiological findings of intracerban hemorrhage5 (18.5)14 (17.1)Multiple microbleeds17 (63.0)47 (57.3)Months between stroke onset and pharmacotherapy1.4 ± 1.50.8 ± 1.0Cilostazol or aspirin therapy16 (59.3)46 (56.1)Cilostazol or angerin (100 mg/day)11 (40.7)36 (43.9)Addition of probucol or none15 (55.6)43 (52.4)Addition of probucol or none15 (56.6)30.827Never smoked18 (66.7)56 (68.3)Current smoking status2 (7.4)6 (7.3)Never smoking for >3 years2 (7.4)6 (7.3)Quit smoking for >3 years4 (14.8)10 (12.2)Quit smoking for >3 years4 (14.8)10 (12.2)Quit smoking for >3 years4 (14.8)10 (12.2)Systoli blood pressure (nm Hg)136 ± 2.1377.0 ± 11.0Diastolic blood pressure (nm Hg)16 (24.30.6)17.6 ± 14.20Diastolic blood pressure (nm Hg)16 (24.30.6)0.711Diastolic blood pressure (nm Hg)16 (24.30.6)17.6 ± 4.80Diastolic blood pressure (nm Hg)16 (24.30.6)17.6 ± 4.80Diastolic blood pressure (nm Hg)16 (24.30.6)17.6 ± 4.80Diastolic bloo	2	10 (37.0)	34 (41.5)	0.463	
Index Index Index Ischemic events 27(100) 80(97.6) <0.999 Transient ischemic event 0(0) 2(2.4) <0.999 High-risk index intracerebral hemorrhage 5(18.5) 14(17.1) <0.999 Radiological findings of intracerebral hemorrhage 5(18.5) 21 (25.6) <0.821 Multiple microbleeds 17(63.0) 47 (57.3) $<0.033*$ Caloszal (100 mg/day) 14 ± 1.5 <0.81 $<0.033*$ Caloszal (100 mg/day) 16 (59.3) 46 (56.1) <0.821 Addition of probucol or none 11 (40.7) 36 (43.9) <0.827 Addition of probucol (250 mg/day) 15 (55.6) 43 (52.4) <0.827 No probucol 15 (55.6) 43 (52.4) <0.827 Quit snoking status 15 (55.6) 43 (52.4) <0.827 Never snoked 18 (66.7) 56 (68.3) <0.827 Quit snoking in the past 3 years 2 (7.4) 6 (7.3) <0.978 Quit snoking for >3 years 4 (14.8) 10 (12.2) <0.491	3	13 (48.2)	40 (48.8)		
Ischemic stroke27 (100)80 (97.6) <0.999 Transient ischemic event0 (0)2 (2.4) <0.999 High-risk index intracerebral hemorrhage5 (18.5)14 (17.1) 0.821 Radiological findings of intracerebral hemorrhage5 (18.5)21 (25.6) 0.821 Multiple microbleeds17 (63.0)47 (57.3) $0.033*$ Months between stroke onset and pharmacotherapy14 ± 1.5 0.8 ± 1.0 $0.033*$ Cilostazol or aspirin therapy14 (16.7) 36 (43.9) 0.826 Cilostazol or aspirin therapy11 (65.3)46 (56.1) 0.826 Aspirin (100 mg/day)16 (59.3)46 (56.1) 0.826 Addition of probucol or none12 (44.4)39 (47.6) 0.827 Addition of probucol 205 mg/day)15 (55.6)43 (52.4) 0.827 No probucol12 (44.4)39 (47.6) 0.827 Quit smoking in the past 3 years2 (7.4)6 (7.3) 0.978 Quit smoking in the past 3 years2 (7.4)6 (7.3) 0.978 Quit smoking in the past 3 years2 (7.4)6 (7.3) 0.978 Yastolic blood pressure (mm Hg)13 (62 20.313 (2.5 ± 16.7)0.491Distolic blood pressure (mm Hg)13 (62 4± 30.6)17.5 ± 12.00.228Heart rate (betas per min J16 (2.4 ± 30.6)17.5 ± 14.2.00.226LDL cholesterol (mg/dL)16 (2.4 ± 30.6)110.7 ± 36.10.055LDL cholesterol (mg/dL)16 (2.4 ± 30.6)110.7 ± 36.10.055LDL cholesterol (mg/dL)16	Ischemic events				
Instant NameInterfact of the sector of the sec	Ischemic stroke	27 (100)	80 (97.6)		
International memory International memory High-risk index intracerebral hemory 5 (18.5) 14 (17.1) Radiological findings of intracerebral hemory 5 (18.5) 21 (25.6) 0.821 Multiple microbleeds 17 (63.0) 47 (57.3) 0.033* Months between stroke onset and pharmacotherapy 1.4 ± 1.5 0.8 ± 1.0 0.033* Cilostazol or aspirin therapy 16 (59.3) 46 (56.1) 0.826 Aspirin (100 mg/day) 11 (40.7) 36 (43.9) 0.826 Aspirin (100 mg/day) 11 (659.3) 43 (52.4) 0.827 Addition of probucol (250 mg/day) 15 (55.6) 43 (52.4) 0.827 No probucol 12 (44.4) 39 (47.6) 0.827 Current smoking status 18 (66.7) 56 (68.3) 0.978 Quit smoking in the past 3 years 2 (7.4) 6 (7.3) 0.978 Quit smoking in the past 3 years 11 (10.2) 0.238 Hare quit smoking for >3 years 11 (10.2) 0.238 Ystolic blood pressure (mm Hg) 150 c.2 0.3 13 (25.1 c.6) 0.491 <td< td=""><td>Transient ischemic event</td><td>0 (0)</td><td>2 (2.4)</td><td><0.999</td></td<>	Transient ischemic event	0 (0)	2 (2.4)	<0.999	
History of intracerbral hemorrhage $5(18.5)$ $14(17.1)$ 0.821 History of intracerbral hemorrhage $5(18.5)$ $11(25.6)$ 0.821 Multiple microbleeds $17(63.0)$ $47(57.3)$ 0.033^* Months between stroke onset and pharmacotherapy 1.4 ± 1.5 0.8 ± 1.0 0.033^* Cilostazol (100 mg/day) $16(59.3)$ $46(56.1)$ 0.826 Aspirin (100 mg/day) $11(40.7)$ $36(43.9)$ 0.826 Addition of probucol or none $12(44.4)$ $39(47.6)$ 0.827 Addition of probucol (250 mg/day) $15(55.6)$ $43(52.4)$ 0.827 No probucol $18(66.7)$ $56(68.3)$ 0.827 Current smoking status $11(40.7)$ $10(12.2)$ 0.978 Qut smoking for >3 years $2(7.4)$ $6(7.3)$ 0.978 Yastolic blood pressure (mm Hg) 136.0 ± 20.3 132.5 ± 16.7 0.491 Diastolic blood pressure (mm Hg) 80.2 ± 12.3 77.0 ± 11.0 0.238 Heart rate (beats per min) 81.8 ± 15.4 82.9 ± 13.6 0.741 Total cholesterol (mg/d1) 162.4 ± 30.6 175.1 ± 42.0 0.226 LDL cholesterol (mg/d1) 45.3 ± 12.7 48.0 ± 10.0 0.076 Huncholesterol (mg/d1) $61.61.2$ $61.61.2$ 0.623	High-risk index intracerebral hemorrhage	0(0)	2 (2.1)		
Initial function interarge $1000000000000000000000000000000000000$	History of intracerebral hemorrhage	5 (18 5)	14 (17 1)		
Native get intercept of intercept17 (63.0)17 (67.3)Multiple microbleeds17 (63.0)47 (57.3)Months between stroke onset and pharmacotherapy1.4 ± 1.50.8 ± 1.00.033*Cilostazol (100 mg/day)16 (59.3)46 (56.1) $$	Radiological findings of intracerebral hemorrhage	5 (18.5)	21 (25.6)	0.821	
InterpretationInterpretationInterpretationInterpretationMonths between stroke onest and pharmacotherapy 1.4 ± 1.5 0.8 ± 1.0 0.033^* Cilostazol or aspirin therapyInterpretation 0.033^* Cilostazol (100 mg/day)Interpretation 0.626 0.826 Aspirin (100 mg/day)Interpretation 0.826 0.826 Addition of probucol or noneInterpretation 0.826 0.827 Addition of probucol (250 mg/day)Interpretation 0.826 0.827 No probucolInterpretation 0.826 0.827 Current smoking statusInterpretation 0.826 0.827 Currently smokingInterpretation $0.666.3$ 0.978 Quit smoking in the past 3 years $2(7.4)$ $6(7.3)$ 0.978 Have quit smoking for >3 years $4(14.8)$ $10(12.2)$ 0.978 Systolic blood pressure (mm Hg)Interpretation 136.0 ± 20.3 132.5 ± 16.7 0.491 Diastolic blood pressure (mm Hg) 80.2 ± 12.3 77.0 ± 11.0 0.238 Heart rate (beats per min) 81.8 ± 15.4 82.9 ± 13.6 0.741 Total cholesterol (mg/d1)Interpretation 162.4 ± 30.6 175.1 ± 42.0 0.226 IDL cholesterol (mg/d1) 95.5 ± 29.7 110.7 ± 36.1 0.055 HDL cholesterol (mg/d1) $61.61.40.0$ $61.61.40.0$ 60.876 Hencoglobin Alc (%) $61.41.2$ $61.41.0$ 0.873	Multiple microbleeds	17 (63.0)	47 (57 3)	01021	
Initial contraint parametering (1)Initial (1)Initial (1)Initial (1)Initial (1)Cilostazol or aspirin therapy16 (59.3)46 (56.1) 0.826 Aspirin (100 mg/day)11 (40.7)36 (43.9) 0.826 Addition of probucol or none11 (40.7)36 (43.9) 0.827 Addition of probucol (250 mg/day)15 (55.6)43 (52.4) 0.827 No probucol12 (44.4)39 (47.6) 0.827 Current smoking status18 (66.7)56 (68.3) 0.978 Currently smoking3 (11.1)10 (12.2) 0.978 Quit smoking for >3 years2 (7.4)6 (7.3) 0.978 Ystolic blood pressure (mm Hg)136.0 ± 20.3132.5 ± 16.70.491Diastolic blood pressure (mm Hg)80.2 ± 12.377.0 ± 11.00.238Heart rate (beats per min)81.8 ± 15.482.9 ± 13.60.741Total cholesterol (mg/d1)95.5 ± 29.7110.7 ± 36.10.055HDL cholesterol (mg/d1)45.3 ± 12.748.0 ± 10.00.076Hemoglobin Alc (%)6.1 ± 126.1 ± 0.90.623	Months between stroke onset and pharmacotherapy	14+15	0.8 + 1.0	0.033*	
Cilostazol (100 mg/day)16 (59.3)46 (56.1)0.826Cilostazol (100 mg/day)11 (40.7)36 (63.9)0.826Addition of probucol or none1Addition of probucol (250 mg/day)15 (55.6)43 (52.4)0.827No probucol12 (44.4)39 (47.6)0.827Current smoking status1Current smoking status1Quit smoking in the past 3 years2 (7.4)6 (67.3)0.978Quit smoking for >3 years4 (14.8)10 (12.2)0.491Diastolic blood pressure (mm Hg)136.6 ± 20.3132.5 ± 16.70.491Diastolic blood pressure (mm Hg)80.2 ± 12.377.0 ± 11.00.238Heart rate (beats per min)81.8 ± 15.482.9 ± 13.60.741Total cholesterol (mg/dL)95.5 ± 29.7110.7 ± 36.10.055HDL cholesterol (mg/dL)45.3 ± 12.748.0 ± 10.00.076Huroglobin Alc (%)61.± 1261 ± 0.90.673	Cilostazol or aspirin therapy			0.000	
Initiation (vol mg/day) Init (vol mg/day) <thinit (vol="" day)<="" mg="" th=""> Init (vol mg/day) Init (vol mg/day)</thinit>	Cilostazol (100 mg/day)	16 (59 3)	46 (56 1)		
Addition of probucol or none 11 (10.17) 10 (10.17) Addition of probucol or none 15 (55.6) 43 (52.4) 0.827 No probucol (250 mg/day) 12 (44.4) 39 (47.6) 0.827 Current smoking status 12 (44.4) 39 (47.6) 0.827 Current smoking status 18 (66.7) 56 (68.3)	Aspirin (100 mg/day)	11 (40 7)	36 (43 9)	0.826	
Addition of probucol (250 mg/day) 15 (55.6) 43 (52.4) 0.827 No probucol 12 (44.4) 39 (47.6) 0.827 Current smoking status 18 (66.7) 56 (68.3)	Addition of probucol or none	11 (100)			
No produced (25 mg/s) 16 (25 mg/s) 16 (25 mg/s) 16 (25 mg/s) 0.827 No probucol 12 (44.4) 39 (47.6) 0.827 Current smoking status 18 (66.7) 56 (68.3)	Addition of producol (250 mg/day)	15 (55.6)	43 (52.4)		
Interpretation Interpr	No probucol	12 (44.4)	39 (47.6)	0.827	
Never smoked 18 (66.7) 56 (68.3) Currently smoking 3 (11.1) 10 (12.2) Quit smoking in the past 3 years 2 (7.4) 6 (7.3) Have quit smoking for >3 years 4 (14.8) 10 (12.2) Systolic blood pressure (mm Hg) 136.0 ± 20.3 132.5 ± 16.7 0.491 Diastolic blood pressure (mm Hg) 80.2 ± 12.3 77.0 ± 11.0 0.238 Heart rate (beats per min) 81.8 ± 15.4 82.9 ± 13.6 0.741 Total cholesterol (mg/dL) 162.4 ± 30.6 175.1 ± 42.0 0.226 LDL cholesterol (mg/dL) 95.5 ± 29.7 110.7 ± 36.1 0.055 HDL cholesterol (mg/dL) 45.3 ± 12.7 48.0 ± 10.0 0.076 Glucose (mg/dL) 116.0 ± 49.0 116.8 ± 42.8 0.888	Current smoking status	()			
Currently smoking B (ctrr)	Never smoked	18 (66.7)	56 (68.3)		
Quit smoking in the past 3 years 2 (7.4) 6 (7.3) Have quit smoking for >3 years 4 (14.8) 10 (12.2) Systolic blood pressure (mm Hg) 136.0 ± 20.3 132.5 ± 16.7 0.491 Diastolic blood pressure (mm Hg) 80.2 ± 12.3 77.0 ± 11.0 0.238 Heart rate (beats per min) 81.8 ± 15.4 82.9 ± 13.6 0.741 Total cholesterol (mg/dL) 162.4 ± 30.6 175.1 ± 42.0 0.226 LDL cholesterol (mg/dL) 95.5 ± 29.7 110.7 ± 36.1 0.055 HDL cholesterol (mg/dL) 45.3 ± 12.7 48.0 ± 10.0 0.076 Glucose (mg/dL) 116.0 ± 49.0 116.8 ± 42.8 0.888 Hemoglobin A1c (%) 6.1 ± 1.2 6.1 ± 0.9 0.673	Currently smoking	3(11.1)	10 (12.2)		
Have quit smoking for >3 years 4 (14.8) 10 (12.2) Systolic blood pressure (mm Hg) 136.0 ± 20.3 132.5 ± 16.7 0.491 Diastolic blood pressure (mm Hg) 80.2 ± 12.3 77.0 ± 11.0 0.238 Heart rate (beats per min) 81.8 ± 15.4 82.9 ± 13.6 0.741 Total cholesterol (mg/dL) 162.4 ± 30.6 175.1 ± 42.0 0.226 LDL cholesterol (mg/dL) 95.5 ± 29.7 110.7 ± 36.1 0.055 HDL cholesterol (mg/dL) 45.3 ± 12.7 48.0 ± 10.0 0.076 Glucose (mg/dL) 116.0 ± 49.0 116.8 ± 42.8 0.888 Hemoglobin ALc (%) 6.1 ± 1.2 6.1 ± 0.9 0.673	Quit smoking in the past 3 years	2 (7 4)	6 (7 3)	0.978	
Systolic blood pressure (mm Hg) 136.0 ± 20.3 132.5 ± 16.7 0.491 Diastolic blood pressure (mm Hg) 80.2 ± 12.3 77.0 ± 11.0 0.238 Heart rate (beats per min) 81.8 ± 15.4 82.9 ± 13.6 0.741 Total cholesterol (mg/dL) 162.4 ± 30.6 175.1 ± 42.0 0.226 LDL cholesterol (mg/dL) 95.5 ± 29.7 110.7 ± 36.1 0.055 HDL cholesterol (mg/dL) 45.3 ± 12.7 48.0 ± 10.0 0.076 Glucose (mg/dL) 116.0 ± 49.0 116.8 ± 42.8 0.888	Have quit smoking for >3 years	4 (14.8)	10 (12.2)		
Private (num Fg) Private (num Fg) Private (num Fg) Diastolic blood pressure (mm Hg) 80.2 ± 12.3 77.0 ± 11.0 0.238 Heart rate (beats per min) 81.8 ± 15.4 82.9 ± 13.6 0.741 Total cholesterol (mg/dL) 162.4 ± 30.6 175.1 ± 42.0 0.226 LDL cholesterol (mg/dL) 95.5 ± 29.7 110.7 ± 36.1 0.055 HDL cholesterol (mg/dL) 45.3 ± 12.7 48.0 ± 10.0 0.076 Glucose (mg/dL) 116.0 ± 49.0 116.8 ± 42.8 0.888 Hemoglobin Alc (%) 6.1 ± 1.2 6.1 ± 0.9 0.673	Systolic blood pressure (mm Hg)	136.0 + 20.3	132.5 + 16.7	0.491	
Heart rate (beats per min) 81.8 ± 15.4 82.9 ± 13.6 0.741 Total cholesterol (mg/dL) 162.4 ± 30.6 175.1 ± 42.0 0.226 LDL cholesterol (mg/dL) 95.5 ± 29.7 110.7 ± 36.1 0.055 HDL cholesterol (mg/dL) 45.3 ± 12.7 48.0 ± 10.0 0.076 Glucose (mg/dL) 116.0 ± 49.0 116.8 ± 42.8 0.888 Hemoglobin ALc (%) 6.1 ± 1.2 6.1 ± 0.9 0.673	Diastolic blood pressure (mm Hø)	80.2 + 12.3	77.0 + 11.0	0.238	
Total cholesterol (mg/dL) 162.4 ± 30.6 175.1 ± 42.0 0.226 LDL cholesterol (mg/dL) 95.5 ± 29.7 110.7 ± 36.1 0.055 HDL cholesterol (mg/dL) 45.3 ± 12.7 48.0 ± 10.0 0.076 Glucose (mg/dL) 116.0 ± 49.0 116.8 ± 42.8 0.888 Hemoglobin A1c (%) 6.1 ± 1.2 6.1 ± 0.9 0.673	Heart rate (beats per min)	81.8 + 15.4	82.9 + 13.6	0.741	
LDL cholesterol (mg/dL) 95.5 ± 29.7 110.7 ± 36.1 0.055 HDL cholesterol (mg/dL) 45.3 ± 12.7 48.0 ± 10.0 0.076 Glucose (mg/dL) 116.0 ± 49.0 116.8 ± 42.8 0.888 Hemoglobin ALc (%) 6.1 ± 1.2 6.1 ± 0.9 0.673	Total cholesterol (mg/dL)	162.4 + 30.6	175.1 + 42.0	0.226	
HDL cholesterol (mg/dL) 45.3 ± 12.7 48.0 ± 10.0 0.0376 Glucose (mg/dL) 116.0 ± 49.0 116.8 ± 42.8 0.888 Hemoglobin A1c (%) 6.1 ± 1.2 6.1 ± 0.9 0.623	LDL cholesterol (mg/dL)	95 5 + 29 7	110 7 + 36 1	0.055	
Glucose (mg/dL) 116.0 ± 49.0 116.8 ± 42.8 0.888 Hemoglobin A1c (%) 6.1 ± 1.2 6.1 ± 0.9 0.673	HDL cholesterol (mg/dL)	45 3 + 12 7	48.0 + 10.0	0.076	
Hemoglobin A1c (%) 6.1 + 1.2 6.1 + 0.9 0.603	Glucose (mg/dL)	116.0 + 49.0	116.8 + 42.8	0.888	
	Hemoglobin A1c (%)	6.1 ± 1.2	6.1 ± 0.9	0.623	

Baseline characteristics are indicated in mean \pm standard deviation for continuous variables and number (percentage, %) for categorical variables. Continuous variables were compared using the Mann–Whitney *U* test, and categorical variables were compared using Fisher's exact tests between decliners (n = 27) and non-decliners (n = 82). The * symbol indicates statistical significance at p < 0.05.

HDL, high-density lipoprotein K-MMSE, Korean-Mini Mental Status Examination; LDL, low-density lipoprotein; NIHSS, National Institutes of Health Stroke Scale.

TABLE 2 Baseline characteristics of training and test sets.

Baseline characteristics	Training set	Test set	p		
	(n = 87)	(n = 22)			
Age (years)	71.4 ± 8.6	72.3 ± 8.8	0.470		
Female	54 (62.1)	13 (59.1)	0.810		
Education years	4.9 ± 4.6	4.7 ± 4.2	0.694		
Months between stroke onset and baseline	4.7 ± 0.7	4.6 ± 0.5	0.895		
K-MMSE score at baseline	19.1 ± 4.5	19.2 ± 4.4	0.922		
NIHSS score on admission	2.2 ± 2.0	1.6 ± 1.6	0.190		
Lesion ratio in the whole brain (%)	3.2 ± 1.8	3.5 ± 1.9	0.464		
Ischemic events					
Ischemic stroke	86 (98.9)	21 (95.5)	0.364		
Transient ischemic event	1 (1.2)	1 (4.6)			
High-risk index intracerebral hemorrhage	High-risk index intracerebral hemorrhage				
History of intracerebral hemorrhage	16 (18.4)	3 (13.6)			
Radiological findings of intracerebral hemorrhage	19 (21.8)	7 (31.8)	0.637		
Multiple microbleeds	52 (59.8)	12 (54.6)	12 (54.6)		
Months between stroke onset and pharmacotherapy	1.0 ± 1.2	0.9 ± 1.1	0.991		
Cilostazol or aspirin therapy					
Cilostazol (100 mg/day)	50 (57.5)	12 (54.6)	0.814		
Aspirin (100 mg/day)	37 (42.5)	10 (45.5)			
Addition of probucol or none					
Probucol (250 mg/day)	46 (52.9)	12 (54.6)	200.05		
No probucol	41 (47.1)	10 (45.5)	<0.999		

Baseline characteristics are indicated in mean \pm standard deviation for continuous variables and number (percentage, %) for categorical variables. Continuous variables were compared using the Mann–Whitney *U* test, and categorical variables were compared using Fisher's exact tests between training (n = 87) and test (n = 22) sets.

K-MMSE, Korean-Mini Mental Status Examination; NIHSS, National Institutes of Health Stroke Scale.

K-MMSE received a negative label (cognitive non-decliners, n = 81, 74.3%).

To address class imbalance and prevent synthetic data biasing evaluation metrics, Synthetic Minority Over-sampling Technique with Tomek Links (SMOTETomek) was applied exclusively to the training set (n = 87) using a sampling strategy of 0.9. This method combines oversampling of the minority class (SMOTE) with undersampling of the majority class (Tomek Links) to improve class distribution while reducing noise (29). The test set (n = 22) preserved the original class distribution, ensuring a fair and unbiased performance assessment under real-world conditions. This separation allowed the model to learn from balanced data while maintaining external validity.

We selected the four machine learning algorithms commonly used in PSCI prediction, as identified in recent systematic reviews (20). The models trained for binary classification of cognitive decline included Categorical Boosting (CatBoost) (30), Adaptive Boosting (AdaBoost) (31), eXtreme Gradient Boosting (XGBoost) (32), and logistic regression (33). Boosting algorithms were chosen for their ability to aggregate weak learners and reduce overfitting, particularly in imbalanced datasets. Logistic regression was included for its simplicity and high interpretability.

A fivefold cross-validation scheme was implemented to evaluate model performance and optimize hyperparameters within the training set. In each fold, models were trained on four randomly selected subsets and validated on the remaining subset, known as the test set. GridSearchCV was used for AdaBoost and XGBoost, as their hyperparameter spaces are relatively small and consist of discrete values. In contrast, RandomizedSearchCV with 10 iterations was applied to CatBoost and logistic regression, which have broader or continuous hyperparameter spaces, to improve computational efficiency. StratifiedKFold was used to maintain class balance across folds, and all models were optimized based on the area under the curve (AUC). The optimal model was selected according to the average performance metrics obtained during cross-validation.

Classification performance for predicting cognitive decline at 9 months was assessed using multiple evaluation metrics, including AUC from the receiver operating characteristic (ROC) curve, accuracy, sensitivity, and specificity. Model evaluation was conducted separately on training set (n = 87) and independent test (n = 22) set. Optimal classification thresholds were determined using the Youden index (sensitivity + specificity – 1) to balance true positive and true negative rates.

All analyses were performed using Python 3.9. Key libraries included scikit-learn (v1.1.3) for training and evaluating models (logistic regression, AdaBoost); CatBoost (v1.1.1) and XGBoost (v1.4.2) for gradient boosting; and imbalanced-learn (v0.10.1) for resampling. A summary of the machine learning workflow and full package versions is provided in Supplementary Table 1.

2.5 Feature importance analysis

Feature importance was analyzed using SHapley Additive exPlanations (SHAP, v0.44.1) to interpret model predictions, identify key predictors of cognitive decline, and enhance overall model transparency (34). SHAP values were used to rank input variables according to their contribution to model output. Features that consistently exhibited low SHAP values across cross-validation folds were excluded to reduce overfitting and improve interpretability. Exclusion thresholds were determined based on both cross-validated model performance and the stability of feature rankings. The final set of input features used for each model is illustrated in the SHAP summary plots (Figure 2).

2.6 Statistical analysis

Data are expressed as mean \pm standard deviation (SD) for continuous variables and number of subjects (%, percentage) for categorical variables. The demographic and clinical characteristics were compared between training (n = 87) and test (n = 22) sets as well as between cognitive decliners (n = 27) and non-decliners (n = 82) using Mann–Whitney *U* tests for continuous variables and Fisher's exact tests for categorical variables. These analyses were conducted to identify baseline differences and confirm that model training and evaluation sets were balanced.



3 Results

3.1 Differential baseline characteristics between cognitive decliners and non-decliners

Table 1 presents baseline characteristics of cognitive decliners and non-decliners. Cognitive decliners had a mean age of 69.2 years, 6.4 years of completed education, were 55.6% male, and had a K-MMSE total score of 23.0, indicating very mild cognitive impairment. Cognitive non-decliners had a mean age of 72.3 years, 4.2 years of completed education, were 63.4% female, and had a K-MMSE total score of 17.9, indicating moderate cognitive impairment.

Regarding baseline characteristics, cognitive decliners had a longer duration between stroke onset and randomization into pharmacotherapy of the PICASSO trial (p = 0.033) than non-decliners. Although this difference did not reach statistical significance, cognitive decliners exhibited trends toward younger age (p = 0.070), more years of education (p = 0.080), longer interval between stroke onset and baseline MMSE (p = 0.089), and lower levels of low-density lipoprotein (LDL) (p = 0.055) and high-density lipoprotein (HDL) (p = 0.076) than non-decliners (Table 1).

Notably, the baseline K-MMSE total score was significantly higher in cognitive decliners compared to non-decliners (p < 0.001). All baseline K-MMSE subscores were also higher in the decliner group, including orientation to time (p = 0.004), orientation to place (p = 0.008), registration (p = 0.011), attention and calculation (p = 0.001), recall (p = 0.018), language (p < 0.001), and visuospatial ability (p = 0.006), compared to the non-decliner group.

3.2 Classification of cognitive decliners in the training set

Table 3 summarizes the classification performance of the four machine learning models on cognitive decliners in the training set (n = 87). In the training set, a fivefold cross-validation of CatBoost yielded most superior performance than the other three machine learning models in terms of accuracy, AUC, and sensitivity. The CatBoost algorithm achieved a mean AUC of

0.966, with an accuracy of 0.897, a sensitivity of 0.909, and a specificity of 0.888.

Mean AUC of the other three models ranked in the following descending order: XGBoost, 0.950; AdaBoost, 0.924; and logistic regression, 0.872. XGBoost algorithm achieved an accuracy of 0.880, a sensitivity of 0.836, and a specificity of 0.918, demonstrating the highest specificity score among the four models. AdaBoost algorithm achieved an accuracy of 0.838, a sensitivity of 0.855, and a specificity of 0.823. Logistic regression algorithm achieved an accuracy of 0.804, a sensitivity of 0.745, and a specificity of 0.855 (Table 3).

3.3 Classification of cognitive decliners in the test set

Classification results of the four machine learning models in the test set (n = 22) are summarized in Table 3 and Figure 3. In the test set, CatBoost outperformed the other three machine learning models in terms of AUC, accuracy, and sensitivity. It achieved a mean AUC of 0.897, an accuracy of 0.873, a sensitivity of 0.700, and a specificity of 0.911.

Mean AUC of the remaining three models ranked in the following descending order: logistic regression, 0.775; AdaBoost, 0.767; and XGBoost, 0.722. The logistic regression algorithm achieved an accuracy of 0.755, a sensitivity of 0.650, and a specificity of 0.778. The AdaBoost showed an accuracy of 0.845, a sensitivity of 0.550, and a specificity of 0.911. The XGBoost algorithm achieved an accuracy of 0.873, a sensitivity of 0.650, and a specificity of 0.650, and a specificity of 0.922, demonstrating the highest accuracy and specificity scores among the four models (Table 2).

3.4 Feature importance for classification of cognitive decliners

Feature importance was determined using the SHAP methodology, with the most crucial feature ranked at the top, as depicted in Figure 2, which shows the selected input features for each model. In the CatBoost model, K-MMSE total scores, language K-MMSE subscore, initial NIHSS score, orientation to place K-MMSE subscore, memory recall K-MMSE subscore, time between stroke onset and MMSE, and lesion volume ratio were the seven most

Dataset	Model	AUC	Accuracy	Sensitivity	Specificity
Training set (<i>n</i> = 87)	CatBoost	0.966	0.897	0.909	0.888
	AdaBoost	0.924	0.838	0.855	0.823
	XGBoost	0.950	0.880	0.836	0.918
	Logistic regression	0.872	0.804	0.745	0.855
Test set (<i>n</i> = 22)	CatBoost	0.897	0.873	0.700	0.911
	AdaBoost	0.767	0.845	0.550	0.911
	XGBoost	0.722	0.873	0.650	0.922
	Logistic Regression	0.775	0.755	0.650	0.778

TABLE 3 Performance of the four machine learning models in predicting cognitive decline.

Performance of the four machine learning models in predicting cognitive decline (\geq 3 points of changes in the total scores of K-MMSE for 9 months) was evaluated using area under the curve (AUC), accuracy, sensitivity, and specificity scores in training (n = 87) and test (n = 22) sets.

AdaBoost, Adaptive Boosting; AUC, area under the receiver operating characteristic curve; CatBoost, Categorical Boosting; XGBoost, Extreme Gradient Boosting.



Performance of the four machine learning models in predicting cognitive decline. The ROC curve indicates AUC of the four machine learning models in predicting cognitive decline (≥3 points of changes in the total scores of K-MMSE over 9 months) of patients with PSCI in the test set. The relationship between true positive rate and false positive rate is indicated for CatBoost (red), AdaBoost (green), XGBoost (blue), and logistic regression (purple). AdaBoost, Adaptive Boosting; AUC, area under the curve; CatBoost, Categorical Boosting; K-MMSE, Korean-Mini Mental Status Examination; ROC, receiver operating characteristic; XGBoost, Extreme Gradient Boosting.

important features at baseline evaluation in predicting cognitive decliners after 9 months.

Across the four machine learning models, K-MMSE total scores, language K-MMSE subscore, and orientation to time K-MMSE subscore were the top three influential features, with higher baseline values contributing to predicting cognitive worsening after 9 months. Higher diastolic blood pressure, longer time since stroke onset, and status of current smoking contributed to predicting cognitive worsening. Patient age, initial NIHSS score, lesion volume ratio, follow-up duration, and time between stroke onset and randomization into pharmacotherapy were also included as the top seven influential factors contributing to predicting cognitive worsening across the four models.

3.5 Sensitivity analysis excluding the baseline K-MMSE total score

To assess the influence of baseline cognitive status on model predictions, we conducted a sensitivity analysis by removing the K-MMSE total score from the input features. As shown in Supplementary Table 2, this led to decreased performance in AUC and sensitivity across most models. For instance, in the CatBoost model, the test AUC dropped from 0.897 to 0.739, accuracy from 0.873 to 0.755, and sensitivity from 0.700 to 0.500. Despite these reductions, the models retained moderate accuracy (0.736–0.845) and specificity (0.767–0.944), indicating preserved overall discriminative ability in the test set.

Notably, the AdaBoost model demonstrated improved performance in the training set (AUC increased from 0.924 to 0.964; sensitivity from 0.855 to 0.930), and its test specificity increased from 0.911 to 0.944. This suggests that, in the absence of the baseline K-MMSE total score, the model adopted a more conservative decision threshold—prioritizing the accurate classification of non-decliners over the detection of true decliners.

SHAP analysis revealed a corresponding shift in feature importance toward clinical and imaging variables, including initial NIHSS score, diastolic blood pressure, lesion volume ratio, and the time interval from stroke onset to pharmacotherapy or cognitive assessment (Supplementary Figure 1). These findings underscore the residual predictive value of non-cognitive features, even when global baseline cognitive measures are excluded.

3.6 Sensitivity analysis without SMOTETomek

To evaluate the impact of resampling, we conducted a sensitivity analysis comparing model performance with and without the application of SMOTETomek (Supplementary Table 3). When SMOTETomek was not applied to the training set (n = 87), sensitivity markedly declined across all models except AdaBoost—most notably in XGBoost and logistic regression, where test sensitivity dropped from 0.650 to 0.250. In contrast, specificity remained high (e.g., logistic regression: 0.989). These findings indicate that SMOTETomek substantially improved sensitivity, particularly for detecting the minority class, while having minimal effect on specificity. This comparison highlights the importance of resampling strategies in addressing class imbalance and enhancing model performance for detecting cognitive decline in test data.

3.7 Sensitivity analysis excluding education years

To evaluate the impact of cognitive reserve proxies on model performance, we conducted a sensitivity analysis by excluding education years from the input features. As shown in Supplementary Table 4, CatBoost and AdaBoost retained or improved performance in the test set despite the exclusion. CatBoost achieved higher test accuracy (from 0.873 to 0.891) and specificity (from 0.911 to 0.922), while AdaBoost maintained stable accuracy (from 0.845 to 0.882) and showed an increase in AUC (from 0.767 to 0.875). In contrast, performance declined in XGBoost and logistic regression, particularly in sensitivity, suggesting greater dependence on education-related input.

SHAP analysis (Figure 2 and Supplementary Figure 2) further supported these findings. Although education years contributed modestly to prediction, their exclusion did not substantially affect the relative importance of key features such as baseline K-MMSE subscores, NIHSS score, diastolic blood pressure, lesion volume ratio, and stroke-to-assessment intervals. These findings suggest that the high-performing models relied primarily on cognitive and clinical variables, reinforcing their robustness in predicting cognitive decline independent of educational attainment.

4 Discussion

This study presents four machine learning models that use clinical and imaging data to predict patients with PSCI at high risk of cerebral hemorrhage, who are likely to experience cognitive decline within 14 months after stroke onset, which is a 9-month follow-up. Specifically, CatBoost demonstrated the highest performance in terms of AUC, accuracy, and sensitivity in training and test sets. The most influential factors for predicting cognitive decline were higher baseline K-MMSE scores (total, language, orientation to place, and recall), a longer interval between stroke onset and baseline MMSE and initial NIHSS scores, and lesion volume ratio in CatBoost. Cognitive decliners who deteriorated after 9 months (mean time since stroke, 14.0 months) had a longer interval between stroke onset and pharmacotherapy, with trends of longer duration between stroke onset and MMSE, younger age, more education, and lower LDL and HDL levels, than non-decliners who showed deterioration earlier at baseline (mean time since stroke, 4.7 months). Moreover, cognitive decliners exhibited higher baseline K-MMSE total scores and subscores compared to non-decliners.

High AUC and accuracy of our CatBoost model emphasize its reliable prediction of cognitive decline in patients with PSCI and highrisk cerebral hemorrhage. Considering typical machine learning models that predict PSCI at a single time point and focus on poststroke functional outcomes, our model excelled in predicting PSCI worsening after 9 months. Previous machine learning models demonstrated comparable predictive performance ranges, with an AUC of 0.80-0.91, an accuracy of 0.74-0.80, a sensitivity of 0.70-0.90, and a specificity of 0.68–0.82 (20, 35–38). In classifying cognitive decliners, boosting models excelled by combining weak learners and preventing overfitting through hyperparameter tuning. CatBoost effectively handled imbalanced datasets and mixed data types, achieving an AUC of 0.897 (30). XGBoost, known for its high efficiency and flexibility, achieved the highest specificity of 0.922 but the lowest AUC of 0.722 (32), AdaBoost, which combines weak classifiers (31) and logistic regression, which is valued for its interpretability (33), showed AUC scores of 0.767 and 0.775, respectively.

In CatBoost that demonstrated the highest performance, a longer interval between stroke onset and baseline MMSE and higher baseline K-MMSE scores were key predictors of cognitive worsening as determined using the SHAP methodology. Statistically, cognitive decliners had higher baseline K-MMSE scores and tended to have a longer time since stroke, be younger, and have more years of education than cognitive non-decliners. This result suggests that cognitive decliners with less baseline impairment have protective factors, such as younger age and higher education, delaying cognitive decline to the 9-month follow-up, unlike cognitive non-decliners who began deteriorating at baseline (39, 40). In AdaBoost and XGBoost, older age importantly predicted cognitive deterioration, which is supported by previous findings that advanced age increases the PSCI odds ratio from 3.5 to 9.4, alongside greater brain plaque formation and reduced blood

vessel elasticity (41-43). An extended educational background, serving as a cognitive reserve, was associated with lower PSCI occurrence, lower dementia prevalence, and improved long-term survival after stroke (39, 40). Sensitivity analysis excluding education years indicated that model predictions were predominantly driven by clinically and cognitively salient features, such as baseline K-MMSE subscores, NIHSS scores, and lesion characteristics, rather than cognitive reserve proxies. This suggests that the model captures neurologically meaningful patterns and remains applicable across populations with varying educational backgrounds, supporting its generalizability in clinical contexts (17, 38, 44). Age and education, both strongly correlated with baseline MMSE scores, have been identified as key predictors in previous machine learning models of PSCI (20, 35). Additionally, higher baseline K-MMSE scores-particularly in the orientation and language domains-may reflect preserved cognitive integrity and serve as protective factors against subsequent decline. These domains are supported by temporoparietal and frontal cortical networks, which are commonly vulnerable to ischemic injury and play a central role in sustaining functional independence and cognitive resilience during post-stroke recovery (45).

In CatBoost, initial NIHSS scores and lesion volume ratio were important factors predicting cognitive decline at 14 months poststroke. These stroke severity indices, although not significantly different between cognitive decliners and non-decliners, importantly predicted PSCI after 9 months without a clear directionality. Higher initial NIHSS scores were associated with an increased risk of cognitive decline in patients with PSCI, particularly among those with subcortical stroke (10, 44, 46). Although the NIHSS was originally developed to quantify acute neurological deficits, elevated scores have also been linked to unfavorable long-term cognitive outcomes, especially in the presence of extensive white matter damage or impaired cerebral perfusion (47). Our finding that the lesion volume ratio (ischemic stroke volume relative to total brain volume) is a key predictor of PSCI aligns with previous research identifying both stroke volume and brain atrophy as major determinants of post-stroke cognition (45). White matter lesions have been identified as significant risk factors for PSCI, contributing to slowed processing speed and impaired executive function through disruption of the fronto-subcortical circuits (35, 37, 48). Larger subcortical infarcts can disrupt key networks involving the basal ganglia, thalamus, and prefrontal cortex-regions essential for working memory and cognitive control-and are strongly linked to PSCI, particularly in small vessel disease (37, 47).

In both the AdaBoost and XGBoost models, the time interval between stroke onset and pharmacotherapy initiation emerged as a significant predictor of cognitive deterioration. In the PICASSO trial, this interval was significantly longer among cognitive decliners than among non-decliners. The trial showed that pharmacotherapy with cilostazol or aspirin, with and without probucol reduced cardiovascular events in patients with ischemic stroke and a high risk of hemorrhage (12). Delayed treatment may compromise cerebrovascular integrity, prolong inflammation, and hinder neurovascular repair, thereby increasing the risk of post-stroke cognitive deterioration (49). These findings suggest that early pharmacologic intervention may play a protective role against cognitive deterioration by mitigating vascular events and supporting recovery mechanisms (50). Consistent with this, diastolic blood pressure and current smoking status were within the top five predictors of cognitive decline in AdaBoost, XGBoost, and logistic regression, which is supported by previous findings that PSCI is closely associated with vascular risk factors such as hypertension, smoking history, diabetes mellitus, and heart disease, all affecting inflammation and cerebral perfusion (20, 38, 51). A trend toward lower LDL and HDL levels, observed in cognitive decliners, suggests disturbed neural maintenance and antioxidant effects underlying cognitive worsening in patients with PSCI (38, 52).

These findings suggest that the model captures not only statistical associations but also underlying pathophysiological mechanisms contributing to post-stroke cognitive decline. The neurobiological relevance of prioritized features supports their clinical utility and mechanistic validity in predicting cognitive trajectories. A predictive model capable of identifying patients at high risk for PSCI within the first year post-stroke may facilitate timely and personalized interventions. High-risk individuals may benefit from early initiation of tailored and intensive cognitive rehabilitation, more frequent neurocognitive monitoring (e.g., every 3-6 months), and earlier use of cognitive-enhancing pharmacologic treatments such as cholinesterase inhibitors or memantine (53, 54). Clinicians may also re-evaluate secondary prevention strategies, including stricter control of vascular risk factors and adjustment of antiplatelet regimens to minimize the risk of hemorrhagic complications (55). Early identification allows for proactive caregiver involvement and planning for personalized support services, helping families prepare for potential cognitive deterioration. Collectively, these targeted interventions may attenuate decline, reduce long-term disability, and enhance functional recovery in patients with PSCI at high risk of cognitive deterioration.

To further assess model robustness and the influence of baseline cognitive status, we conducted a sensitivity analysis excluding the K-MMSE total score. This led to a reduction in AUC and sensitivity in most models, highlighting the strong predictive weight of baseline cognition. Interestingly, AdaBoost demonstrated improved training performance and increased test specificity, indicating a more conservative classification pattern prioritizing the accurate identification of non-decliners at the cost of missing true decliners. SHAP analysis revealed a shift in feature importance toward non-cognitive variables such as NIHSS score, diastolic blood pressure, lesion volume ratio, and timing of pharmacotherapy or cognitive evaluation. These results suggest that even in the absence of global cognitive scores, meaningful clinical and imaging predictors of cognitive decline can still be identified. However, the trade-off in reduced sensitivity underscores the need to balance predictor selection with intended clinical use-particularly in early detection versus diagnostic confirmation contexts.

Several limitations should be considered when interpreting these findings. The relatively small sample size (n = 109) and retrospective design may limit the generalizability of the results and introduce potential selection or information biases. The small size of the test set (n = 22) limits the stability of performance estimates, particularly sensitivity. While internal cross-validation and standardized data collection provide some reassurance, external validation with larger, prospective, and multicenter cohorts is essential to confirm the generalizability, robustness, and clinical applicability of our model.

Additionally, class imbalance—cognitive decliners made up only 24.7% of the sample—may have contributed to reduced sensitivity. To address this, we applied SMOTETomek, which combines oversampling of the minority class with the removal of borderline majority class samples. This method was applied only to the training set to prevent data leakage and preserve the original class distribution in the test set. As shown in our sensitivity analysis (Supplementary Table 3), removing

SMOTETomek led to a marked drop in sensitivity—particularly in XGBoost and logistic regression—while specificity remained high. This demonstrates its utility in improving minority class detection with minimal loss of specificity. However, synthetic sampling can introduce bias or overfitting, particularly in small datasets. To mitigate this, we used stratified k-fold cross-validation to maintain class proportions and ensure reliable performance estimates, and performed SHAP analysis to confirm the stability and clinical plausibility of feature importance. Future studies should incorporate external validation with larger and more diverse cohorts. In addition to SMOTETomek, methods such as bootstrap resampling and integration of multimodal clinical and imaging data may further enhance model performance.

Furthermore, the operational definition of cognitive decline as $a \ge 3$ -point decrease in K-MMSE scores over 9 months, while consistent with prior studies, may not fully capture the multidimensional nature of cognitive deterioration. Incorporating clinical assessments and a broader range of neuropsychological tools would provide a more comprehensive evaluation framework. Future research should also consider extending the follow-up period to better understand the long-term trajectory of PSCI beyond 9 months. Our sensitivity analysis demonstrated that excluding the baseline K-MMSE total score resulted in a reduction in AUC and sensitivity, highlighting its critical role in the early detection of cognitive decline. Nonetheless, the increased relative importance of non-cognitive features-such as lesion volume, blood pressure, stroke severity, stroke duration, and timing of pharmacotherapy-suggests that models omitting global cognitive scores may still capture clinically relevant predictors. Future studies should consider excluding baseline cognitive scores to better delineate the contributions of alternative features, while carefully addressing the inherent trade-off in predictive sensitivity.

In conclusion, machine learning models, particularly the CatBoost algorithm, may reliably predict patients with PSCI with high-risk cerebral hemorrhage, who may experience cognitive decline within 14 months after stroke onset. According to SHAP and statistical analyses, cognitive decliners had protective factors of younger age and extended education, which delayed deterioration till the 9-month follow-up, compared with cognitive non-decliners who showed cognitive worsening earlier at baseline. A longer interval between stroke onset and pharmacotherapy, along with smoking status and cholesterol levels, may contribute to predicting cognitive decline as risk factors.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the institutional review boards at each study site (list available as Supplementary material). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

EN: Formal analysis, Methodology, Writing – original draft, Funding acquisition. YK: Methodology, Validation, Visualization, Writing – review & editing. SK: Conceptualization, Supervision, Validation, Writing – review & editing. D-WK: Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This research was supported by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI) funded by the Ministry of Health & Welfare (HR18C0016) and a grant from the National Research Foundation of Korea (NRF) funded by the Korean government (MSIT) (2022R1F1A1060778), Republic of Korea.

Acknowledgments

We thank the PICASSO investigators.

Conflict of interest

YK was employed by Nunaps Inc. D-WK is the CEO of Nunaps Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2025.1569073/ full#supplementary-material

References

1. Gorelick PB. The global burden of stroke: persistent and disabling. *Lancet Neurol.* (2019) 18:417–8. doi: 10.1016/S1474-4422(19)30030-4

2. Jokinen H, Melkas S, Ylikoski R, Pohjasvaara T, Kaste M, Erkinjuntti T, et al. Poststroke cognitive impairment is common even after successful clinical recovery. *Eur J Neurol.* (2015) 22:1288–94. doi: 10.1111/ene.12743

3. Pendlebury ST, Rothwell PM. Prevalence, incidence, and factors associated with pre-stroke and post-stroke dementia: a systematic review and meta-analysis. *Lancet Neurol.* (2009) 8:1006–18. doi: 10.1016/S1474-4422(09)70236-4

4. Melkas S, Oksala NK, Jokinen H, Pohjasvaara T, Vataja R, Oksala A, et al. Poststroke dementia predicts poor survival in long-term follow-up: influence of prestroke cognitive decline and previous stroke. *J Neurol Neurosurg Psychiatry*. (2009) 80:865–70. doi: 10.1136/jnnp.2008.166603

5. Nys G, Van Zandvoort M, De Kort P, Van Der Worp H, Jansen B, Algra A, et al. The prognostic value of domain-specific cognitive abilities in acute first-ever stroke. *Neurology*. (2005) 64:821–7. doi: 10.1212/01.WNL.0000152984.28420.5A

6. Pendlebury ST. Dementia in patients hospitalized with stroke: rates, time course, and clinico-pathologic factors. *Int J Stroke*. (2012) 7:570–81. doi: 10.1111/j.1747-4949.2012.00837.x

7. Yu K-H, Cho S-J, Oh MS, Jung S, Lee J-H, Shin J-H, et al. Cognitive impairment evaluated with vascular cognitive impairment harmonization standards in a multicenter prospective stroke cohort in Korea. *Stroke.* (2013) 44:786–8. doi: 10.1161/STROKEAHA.112.668343

8. Martinez-Ramirez S, Greenberg SM, Viswanathan A. Cerebral microbleeds: overview and implications in cognitive impairment. *Alzheimers Res Ther*. (2014) 6:1–7. doi: 10.1186/alzrt263

9. Yamada M. Cerebral amyloid angiopathy: emerging concepts. J Stroke. (2015) 17:17–30. doi: 10.5853/jos.2015.17.1.17

10. Diener H-C, Bogousslavsky J, Brass LM, Cimminiello C, Csiba L, Kaste M, et al. Aspirin and clopidogrel compared with clopidogrel alone after recent ischaemic stroke or transient ischaemic attack in high-risk patients (MATCH): randomised, double-blind, placebo-controlled trial. *Lancet.* (2004) 364:331–7. doi: 10.1016/S0140-6736(04)16721-4

11. Johnston SC, Amarenco P, Albers GW, Denison H, Easton JD, Evans SR, et al. Ticagrelor versus aspirin in acute stroke or transient ischemic attack. *N Engl J Med.* (2016) 375:35–43. doi: 10.1056/NEJMoa1603060

12. Kim BJ, Lee EJ, Kwon SU, Park JH, Kim YJ, Hong KS, et al. Prevention of cardiovascular events in Asian patients with ischaemic stroke at high risk of cerebral haemorrhage (PICASSO): a multicentre, randomised controlled trial. *Lancet Neurol.* (2018) 17:509–18. doi: 10.1016/S1474-4422(18)30128-5

13. Shinohara Y, Katayama Y, Uchiyama S, Yamaguchi T, Handa S, Matsuoka K, et al. Cilostazol for prevention of secondary stroke (CSPS 2): an aspirin-controlled, doubleblind, randomised non-inferiority trial. *Lancet Neurol.* (2010) 9:959–68. doi: 10.1016/S1474-4422(10)70198-8

14. Tardif J-C, Côté G, Lespérance J, Bourassa M, Lambert J, Doucet S, et al. Probucol and multivitamins in the prevention of restenosis after coronary angioplasty. *N Engl J Med.* (1997) 337:365–72. doi: 10.1056/NEJM199708073370601

15. Kim BJ, Kim JS. Ischemic stroke subtype classification: an Asian viewpoint. J Stroke. (2014) 16:8–17. doi: 10.5853/jos.2014.16.1.8

16. Qureshi AI, Mendelow AD, Hanley DF. Intracerebral haemorrhage. *Lancet*. (2009) 373:1632-44. doi: 10.1016/S0140-6736(09)60371-8

17. Filler J, Georgakis MK, Dichgans M. Risk factors for cognitive impairment and dementia after stroke: a systematic review and meta-analysis. *Lancet Healthy Longev*. (2024) 5:e31–44. doi: 10.1016/S2666-7568(23)00217-9

18. Mijajlović MD, Pavlović A, Brainin M, Heiss W-D, Quinn TJ, Ihle-Hansen HB, et al. Post-stroke dementia-a comprehensive review. *BMC Med.* (2017) 15:1–12. doi: 10.1186/s12916-017-0779-7

19. Murphy TH, Corbett D. Plasticity during stroke recovery: from synapse to behaviour. *Nat Rev Neurosci.* (2009) 10:861–72. doi: 10.1038/nrn2735

20. Li X, Chen Z, Jiao H, Wang B, Yin H, Chen L, et al. Machine learning in the prediction of post-stroke cognitive impairment: a systematic review and meta-analysis. *Front Neurol.* (2023) 14:1211733. doi: 10.3389/fneur.2023.1211733

21. Rasquin S, Verhey F, Van Oostenbrugge R, Lousberg R, Lodder J. Demographic and CT scan features related to cognitive impairment in the first year after stroke. *J Neurol Neurosurg Psychiatry*. (2004) 75:1562–7. doi: 10.1136/jnnp.2003.024190

22. Yu K-H, Hong K-S, Oh M-S, Lee J, Lee JS, Kwon SU. Design and rationale for a cognitive outcome substudy in ischemic stroke patients with high risk of cerebral hemorrhage. *J Stroke Cerebrovasc Dis.* (2016) 25:2061–6. doi: 10.1016/j. jstrokecerebrovasdis.2016.04.028

23. Kang Y, Na D-L, Hahn S. A validity study on the Korean Mini-mental state examination (K-MMSE) in dementia patients. J Korean Neurol Assoc. (1997) 15:300–8.

24. Kwah LK, Diong J. National institutes of health stroke scale (NIHSS). J Physiother. (2014) 60:12. doi: 10.1016/j.jphys.2013.12.012

25. Fazekas F, Chawluk JB, Alavi A, Hurtig HI, Zimmerman RA. MR signal abnormalities at 1.5 T in Alzheimer's dementia and normal aging. *Am J Roentgenol.* (1987) 149:351-6. doi: 10.2214/ajr.149.2.351

26. Schmidt P, Wink L. (2017) LST: a lesion segmentation tool for SPM. Manual. Version 2.0. Jülich (Germany): Institute of Neuroscience and Medicine, Research Center Jülich.

27. Hensel A, Angermeyer MC, Riedel-Heller SG. Measuring cognitive change in older adults: reliable change indices for the Mini-mental state examination. *J Neurol Neurosurg Psychiatry*. (2007) 78:1298–303. doi: 10.1136/jnnp.2006.109074

28. Tombaugh TN. Test-retest reliable coefficients and 5-year change scores for the MMSE and 3MS. Arch Clin Neuropsychol. (2005) 20:485–503. doi: 10.1016/j.acn.2004.11.004

29. Wang Z, Wu C, Zheng K, Niu X, Wang X. Smotetomek-based resampling for personality recognition. *IEEE Access*. (2019) 7:129678–89. doi: 10.1109/ACCESS.2019.2940061

30. Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A. CatBoost: unbiased boosting with categorical features In: Bengio S, Wallach HM, Larochelle H, Grauman K, Cesa-Bianchi N, Garnett R, editors. Advances in neural information processing systems. Vol. 31. Curran Associates, Inc. (2018) p. 6639–49.

31. Freund Y, Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. *J Comput Syst Sci.* (1997) 55:119–39. doi: 10.1006/jcss.1997.1504

32. Chen T., Guestrin C. (2016). Xgboost: a scalable tree boosting system, in: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 785–794.

33. 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 models. *J Clin Epidemiol.* (2019) 110:12–22. doi: 10.1016/j.jclinepi.2019.02.004

34. Mangalathu S, Hwang S-H, Jeon J-S. Failure mode and effects analysis of RC members based on machine-learning-based SHapley additive exPlanations (SHAP) approach. *Eng Struct.* (2020) 219:110927. doi: 10.1016/j.engstruct.2020.110927

35. Chander RJ, Lam BYK, Lin X, Ng AYT, Wong APL, Mok VCT, et al. Development and validation of a risk score (CHANGE) for cognitive impairment after ischemic stroke. *Sci Rep.* (2017) 7:12441. doi: 10.1038/s41598-017-12755-z

36. Drozdowska BA, Mcgill K, Mckay M, Bartlam R, Langhorne P, Quinn TJ. Prognostic rules for predicting cognitive syndromes following stroke: a systematic review. *Eur Stroke J.* (2021) 6:18–27. doi: 10.1177/2396987321997045

37. Kandiah N, Wiryasaputra L, Narasimhalu K, Karandikar A, Marmin M, Chua EV, et al. Frontal subcortical ischemia is crucial for post stroke cognitive impairment. *J Neurol Sci.* (2011) 309:92–5. doi: 10.1016/j.jns.2011.07.013

38. Lee M, Yeo N-Y, Ahn H-J, Lim J-S, Kim Y, Lee S-H, et al. Prediction of post-stroke cognitive impairment after acute ischemic stroke using machine learning. *Alzheimers Res Ther.* (2023) 15:147. doi: 10.1186/s13195-023-01289-4

39. Brayne C., Ince P.G., Keage H.A., Mckeith I.G., Matthews F.E., Polvikoski T., et al. (2010) Education, the brain and dementia: neuroprotection or compensation? EClipSE Collaborative Members *Brain* 133 2210–2216. doi: 10.1093/brain/awq185

40. Ojala-Oksala J, Jokinen H, Kopsi V, Lehtonen K, Luukkonen L, Paukkunen A, et al. Educational history is an independent predictor of cognitive deficits and long-term survival in postacute patients with mild to moderate ischemic stroke. *Stroke.* (2012) 43:2931–5. doi: 10.1161/STROKEAHA.112.667618

41. Brainin M, Tuomilehto J, Heiss WD, Bornstein NM, Bath PM, Teuschl Y, et al. Post-stroke cognitive decline: an update and perspectives for clinical research. *Eur J Neurol.* (2015) 22:229–38. doi: 10.1111/ene.12626

42. Rasquin SM, Lodder J, Ponds RW, Winkens I, Jolles J, Verhey FR. Cognitive functioning after stroke: a one-year follow-up study. *Dementia Geriatr Cogn Disord*. (2004) 18:138–44. doi: 10.1159/000079193

43. Zimmerman B, Rypma B, Gratton G, Fabiani M. Age-related changes in cerebrovascular health and their effects on neural function and cognition: a comprehensive review. *Psychophysiology*. (2021) 58:e13796. doi: 10.1111/psyp.13796

44. Glymour MM, Berkman LF, Ertel KA, Fay ME, Glass TA, Furie KL. Lesion characteristics, NIH stroke scale, and functional recovery after stroke. *Am J Phys Med Rehabil.* (2007) 86:725–33. doi: 10.1097/PHM.0b013e31813e0a32

45. Puy L, Barbay M, Roussel M, Canaple S, Lamy C, Arnoux A, et al. Neuroimaging determinants of poststroke cognitive performance: the GRECogVASC study. *Stroke.* (2018) 49:2666–73. doi: 10.1161/STROKEAHA.118.021981

46. Liu Z, Liu Y, Tu X, Shen H, Qiu H, Chen H, et al. High serum levels of malondialdehyde and 8-OHdG are both associated with early cognitive impairment in patients with acute ischaemic stroke. *Sci Rep.* (2017) 7:9493. doi: 10.1038/s41598-017-09988-3

47. Munsch F, Sagnier S, Asselineau J, Bigourdan A, Guttmann CR, Debruxelles S, et al. Stroke location is an independent predictor of cognitive outcome. *Stroke*. (2016) 47:66–73. doi: 10.1161/STROKEAHA.115.011242

48. Kandiah N, Chander RJ, Lin X, Ng A, Poh YY, Cheong CY, et al. Cognitive impairment after mild stroke: development and validation of the SIGNAL2 risk score. *J Alzheimers Dis.* (2016) 49:1169–77. doi: 10.3233/JAD-150736

49. Iadecola C, Anrather J. The immunology of stroke: from mechanisms to translation. *Nat Med.* (2011) 17:796–808. doi: 10.1038/nm.2399

50. Moskowitz MA, Lo EH, Iadecola C. The science of stroke: mechanisms in search of treatments. *Neuron*. (2010) 67:181–98. doi: 10.1016/j.neuron.2010.07.002

51. Sahathevan R, Brodtmann A, Donnan GA. Dementia, stroke, and vascular risk factors; a review. *Int J Stroke.* (2012) 7:61–73. doi: 10.1111/j.1747-4949.2011. 00731.x

52. Kim KY, Shin KY, Chang K-A. Potential biomarkers for post-stroke cognitive impairment: a systematic review and meta-analysis. *Int J Mol Sci.* (2022) 23:602. doi: 10.3390/ijms23020602

53. Cicerone KD, Langenbahn DM, Braden C, Malec JF, Kalmar K, Fraas M, et al. Evidence-based cognitive rehabilitation: updated review of the literature from 2003 through 2008. *Arch Phys Med Rehabil.* (2011) 92:519–30. doi: 10.1016/j.apmr.2010.11.015

54. Román GC, Sachdev P, Royall DR, Bullock RA, Orgogozo J-M, López-Pousa S, et al. Vascular cognitive disorder: a new diagnostic category updating vascular cognitive impairment and vascular dementia. *J Neurol Sci.* (2004) 226:81–7. doi: 10.1016/j.jns.2004.09.016

55. Kernan WN, Ovbiagele B, Black HR, Bravata DM, Chimowitz MI, Ezekowitz MD, et al. Guidelines for the prevention of stroke in patients with stroke and transient ischemic attack: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke.* (2014) 45:2160–236. doi: 10.1161/STR.00000000000024