

Automated detection of new cerebral infarctions and prognostic implications using deep learning on serial MRI

Received: 20 September 2025

Accepted: 22 February 2026

Cite this article as: Cho, H.-h., Lee, J., Bae, J. *et al.* Automated detection of new cerebral infarctions and prognostic implications using deep learning on serial MRI. *npj Digit. Med.* (2026). <https://doi.org/10.1038/s41746-026-02511-x>

Hwan-ho Cho, Joonwon Lee, Jeonghoon Bae, Dongwhane Lee, Hyung Chan Kim, Suk Yoon Lee, Jung Hwa Seo, Woo-Keun Seo, Jin-Man Jung, Hyunjin Park & Seongho Park

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

Automated Detection of New Cerebral Infarctions and Prognostic Implications Using Deep Learning on Serial MRI

Hwan-ho Cho¹, Joonwon Lee², Bae Jeonghoon³, Dongwhane Lee⁴, Hyung Chan Kim⁵,
Suk Yoon Lee⁶, Jung Hwa Seo⁷, Woo-Keun Seo⁸, Jin-Man Jung⁹, Hyunjin Park^{10,11},
Seongho Park¹²

1. Department of Electronics Engineering, Incheon National University, Incheon, South Korea
2. Department of Neurology, Inje University Haeundae Paik Hospital, Busan, South Korea
3. Department of Neurology, Chung-Ang University Gwangmyeong hospital, Gwangmyeong, South Korea
4. Department of Neurology, Uijeongbu Eulji Medical Center, Eulji University School of Medicine, Uijeongbu, South Korea
5. Department of Neurology, Ulsan Hospital, Ulsan, South Korea
6. Department of Neurology, Busan Paik Hospital, Inje University College of Medicine, Busan, South Korea
7. Department of Neurology, Dong-A University College of Medicine, Busan, South Korea
8. Department of Neurology and Stroke Center, Samsung Medical Center, Seoul, South Korea
9. Department of Neurology, Korea University Ansan Hospital, Ansan, South Korea
10. Department of Electrical and Computer Engineering, Sungkyunkwan University, Suwon, South Korea
11. Center for Neuroscience Imaging Research, Institute for Basic Science, Suwon, South Korea
12. Department of Neurology, Hanyang University Guri Hospital, College of Medicine, Hanyang University, Guri, South Korea

Corresponding author:

Seongho Park, Prof., M.D.

Department of Neurology, Hanyang University Guri Hospital, College of Medicine, Hanyang University,

Kyongchun-ro 153, Guri-si, Gyeonggi-do, Republic of Korea

Postcode: 11923

Tel.: +82-31-560-2076

e-mail: risepsh@gmail.com

Abstract

We developed and externally validated a deep learning model to automatically detect new ischemic lesions on serial FLAIR MRI scans in patients with stroke. Manual interpretation of follow-up imaging is labor-intensive and variable, and silent brain infarctions (SBIs) are frequently missed despite their prognostic importance. Using 25,451 paired slices from 1,055 patients across two hospitals, we trained a convolutional neural network with supervised contrastive learning to classify new lesion occurrence. The model achieved an area under the receiver operating characteristic curve of 0.89 in both internal and external validation cohorts. To evaluate clinical relevance, we further analyzed an independent asymptomatic cohort of 307 patients with a median follow-up of two years. Patients classified as SBI-positive by the model showed a significantly higher risk of subsequent symptomatic stroke than those without SBI. In multivariable Cox regression adjusted for age and major vascular risk factors, model-positive patients had a 3.8-fold increased risk of stroke recurrence. These findings indicate that AI can identify clinically meaningful SBIs that are under-recognized in routine practice and independently associated with stroke recurrence. Automated lesion detection may provide a reproducible imaging biomarker for risk stratification, supporting standardized interpretation of follow-up MRI and informing secondary stroke prevention strategies.

Introduction

Accurate detection of new ischemic lesions on follow-up brain MRI is critical for monitoring disease progression and tailoring secondary prevention strategies in patients with cerebrovascular disease. Serial comparisons of baseline and follow-up scans provide essential information regarding recurrent infarction and treatment response.¹ A substantial proportion of new infarcts, however, occur without overt neurological symptoms and remain clinically unrecognized.² These silent brain infarctions (SBIs) are common, occurring in approximately 20% of older adults, and have been consistently associated with increased risk of future symptomatic stroke and cognitive decline.^{3,4} Reliable identification of SBIs therefore holds significant prognostic and therapeutic implications.

Despite their clinical importance, detection of new infarcts on serial MRI remains challenging in practice. Conventional assessment relies on manual slice-by-slice comparison by radiologists or stroke specialists, a process that is time-consuming, cognitively demanding, and prone to interobserver variability.^{5,6} Small or subtle lesions may be overlooked, particularly when no clinical suspicion is present, leading to inconsistent ascertainment across studies and clinical settings. Prior reports indicate disagreement among experienced readers in up to one quarter of patients when adjudicating new lesions, underscoring the limitations of current approaches.⁷ Moreover, the lack of standardized and reproducible detection methods has hindered the translation of imaging-based markers such as SBI into routine clinical decision-making.

Advances in deep learning have enabled automated analysis of complex neuroimaging data, offering potential solutions to the limitations of manual longitudinal image comparison. Convolutional neural networks (CNNs) can capture subtle differences between serial scans and facilitate standardized lesion detection, with visualization techniques providing additional interpretability.^{8,9}

Accordingly, the primary objective of this study was to develop and externally validate a CNN-based framework for automated detection of new ischemic lesions on serial FLAIR MRI. As a secondary objective, and as a downstream clinical application of the developed framework, we evaluated the prognostic significance of AI-detected silent brain infarctions in an asymptomatic cohort. We hypothesized that an automated, standardized detection framework could improve reliability of SBI identification and support individualized risk stratification in secondary stroke prevention.

Results

Study Population and Dataset Characteristics

Among 15,267 patients screened, 1,258 underwent at least 2 temporally distinct MRI examinations, of whom 1,055 met inclusion criteria after exclusions (633 from Hospital_A and 422 from Hospital_B). The mean (SD) age was 72.2 (13.0) years, and 579 participants (54.9%) were men. The median interval between baseline and follow-up MRI was 28 days (interquartile range [IQR], 5–295).

Annotation of paired FLAIR slices yielded 25,451 image pairs, with 3,784 (14.9%) labeled as “changed” and 21,667 (85.1%) as “unchanged.” In Hospital_A (internal dataset), 15,357 pairs were annotated (13.4% positive), while Hospital_B (external dataset) contributed 10,094 pairs (17.1% positive) (**Table 1**). At the patient level, at least one slice labeled as “change” was

identified in 281 patients (44.4%) in Hospital_A and 202 patients (47.9%) in Hospital_B.

Diagnostic Performance of the AI Model

At the slice level, the model achieved an AUC of 0.894 (95% CI, 0.880–0.907) in the internal validation set and 0.898 (95% CI, 0.890–0.907) in the external validation set. Using an optimal threshold derived from Youden's J statistic, sensitivity and specificity were 80.6% and 83.6% internally, and 78.5% and 87.7% externally (**Figure 1A; Supplementary Table 1**). Calibration analysis demonstrated good agreement between predicted and observed probabilities (Brier score, 0.078 internally and 0.082 externally) (**Figure 1B**).

At the patient level, the AUC was 0.861 (95% CI, 0.809–0.912) in the internal cohort and 0.892 (95% CI, 0.862–0.923) in the external cohort, with corresponding sensitivities of 77.7% and 85.1% and specificities of 80.2% and 81.4%, respectively (**Supplementary Figure 6, and Supplementary Table 1**). Grad-CAM visualizations highlighted regions corresponding to true new infarcts on follow-up imaging, confirming that the model localized clinically relevant lesions (**Figure 2**).

Clinical Prognostic Analysis in the Asymptomatic Cohort

Of the 422 patients in the external validation cohort, 307 were asymptomatic at the index follow-up MRI and included in prognostic analyses. Among these, 182 (59.3%) were classified as AI-detected SBI-positive and 125 (40.7%) as AI-detected SBI-negative. Baseline characteristics were generally similar, though diabetes, dyslipidemia, and atrial fibrillation were more frequent in the AI-positive group (**Table 2**).

Over a median follow-up of 678 days (IQR, 42–1521), 24 symptomatic ischemic strokes occurred: 20 in the AI-positive group and 4 in the AI-negative group. Stroke incidence rates were 62.0 per 1,000 person-years in the AI-positive group and 14.3 per 1,000 person-years in the AI-negative group. Kaplan–Meier analysis demonstrated a significantly higher cumulative incidence of stroke in the AI-positive group (log-rank $P=.005$). At 5 years, cumulative stroke incidence was 24.9% (95% CI, 13.0%–35.2%) in the AI-positive group vs 5.5% (95% CI, 0.0%–10.8%) in the AI-negative group (**Figure 3**).

In unadjusted Cox regression, AI-detected SBI was associated with a 4.11-fold increased risk of symptomatic stroke (HR, 4.11; 95% CI, 1.40–12.1; $P=.01$). After adjustment for age, diabetes, and atrial fibrillation, AI-detected SBI remained independently associated with increased risk (adjusted HR, 3.83; 95% CI, 1.30–11.30; $P=.02$). Diabetes was also independently associated with stroke recurrence (adjusted HR, 2.51; 95% CI, 1.10–5.75; $P=.03$) (**Table 3**).

Discussion

In this multicenter cohort study, we developed and externally validated a deep learning model for the automated detection of new ischemic lesions on sequential FLAIR MRI. The model demonstrated high diagnostic accuracy at both the slice and patient levels, with consistent performance in the external validation cohort. Beyond its technical performance, the model identified silent brain infarctions (SBIs) in patients without new neurological symptoms, and the presence of these AI-detected SBIs was independently associated with an increased risk of subsequent symptomatic stroke. These findings underscore the potential of automated image analysis to improve the reproducibility of lesion detection and to provide prognostic information not readily available from routine clinical assessment.

Our model's identification of SBIs as a potent predictor of future stroke aligns with and reinforces findings from prior epidemiologic studies. The nearly 4-fold increased risk of stroke associated with AI-detected SBIs is comparable in magnitude to the hazard ratio (approximately 3.8) for incident ischemic lesions reported by Fang et al,¹ supporting the notion that lesions identified by our algorithm carry similar clinical significance to those defined in prospective cohorts. However, the clinical utility of SBIs has been hampered by the challenges of manual detection. Conventional slice-by-slice visual comparison is labor-intensive, time-consuming, and subject to considerable inter-reader variability, particularly for small or subtle lesions.⁷ Our automated framework addresses this critical gap by providing a standardized and reproducible method for identifying these prognostically important lesions, which might otherwise be overlooked in clinical practice.

The strong prognostic association observed in this study suggests that AI-detected SBIs should not be regarded as mere incidental findings. Rather, their detection presents a critical opportunity to refine secondary prevention strategies. Identifying a patient with a new SBI could prompt more aggressive management, such as stricter blood pressure control, optimization of antithrombotic therapy, or a more extensive workup for potential cardioembolic sources. Furthermore, as SBIs are increasingly recognized as a key manifestation of cerebral small vessel disease linked to both recurrent stroke and cognitive decline,^{4,15–18} our model may help identify patients at risk for broader vascular and cognitive complications, facilitating more comprehensive long-term care planning.

A key strength of our approach is the model's inherent interpretability. Through Grad-CAM visualizations, the model highlights specific image regions that drive its predictions with heatmaps. In our validation, these heatmaps consistently corresponded with radiologically confirmed new infarctions, confirming that the model learned clinically relevant features. This transparency helps mitigate the "black box" concern often associated with medical AI and fosters clinical trust,¹⁹ positioning the tool as a robust assistive aid that can reduce cognitive burden and improve the consistency of radiologic interpretation, rather than as a replacement for clinician expertise.

Beyond individual patient care, our automated tool holds considerable potential for clinical research. Symptomatic recurrent strokes are relatively infrequent events, often necessitating large cohorts and long follow-up periods in clinical trials. Given that SBIs occur more frequently, their standardized and reliable detection could enable their use as a surrogate end point for assessing the efficacy of novel preventive therapies,^{20,21} potentially facilitating more efficient and timely clinical trials. This approach is well-established in other neurologic conditions, such as multiple sclerosis, where new MRI lesions are an accepted marker of disease activity.

In this study, we proposed an approach for detecting newly developed lesions based on longitudinal FLAIR imaging; however, several additional directions warrant further investigation. First, because our analysis focused on lesions that were relatively well delineated on FLAIR images, it is possible that newly developed ischemic lesions occurring in the presence of extensive leukoaraiosis may have been partially obscured by background signal abnormalities. In particular, in cases with severe white matter hyperintensities, newly emerging ischemic changes may overlap with pre-existing lesions and be difficult to distinguish visually, highlighting an inherent limitation of approaches relying on a single imaging sequence.

In this context, future studies may benefit from extending beyond FLAIR-only analyses by integrating additional imaging sequences. For example, the incorporation of diffusion-weighted imaging (DWI) or other quantitative MRI techniques could enable more precise characterization of lesion evolution over time and improve discrimination between newly developed and pre-existing abnormalities. Furthermore, given that leukoaraiosis itself is considered an imaging marker of underlying cerebral small vessel vulnerability, systematic quantification of its burden and spatial distribution, together with the occurrence of new ischemic lesions and subsequent clinical events, may represent an important avenue for future research.

In addition, while the present study primarily focused on detecting imaging changes over time, future work should explore how such imaging-based information can be translated into clinical decision-making. For instance, prospective studies could evaluate whether the identification of newly emerging lesions on follow-up imaging can inform treatment strategies, such as adjustment of antiplatelet or anticoagulant therapy or modification of surveillance imaging intervals. Such investigations would help clarify the clinical utility of longitudinal imaging-based approaches and further define their role in guiding patient management.

This study has several limitations. First, the retrospective design of this study limits causal inference and is inherently susceptible to selection bias. In particular, the analysis of the asymptomatic cohort included only patients who underwent brain MRI because of suspected cerebrovascular symptoms or who received serial MRI examinations, which may reflect specific clinical contexts rather than an unselected population. Therefore, the findings should be interpreted not as an estimate of disease prevalence in the general population, but rather as an assessment of risk stratification within patients undergoing follow-up MRI in real-world clinical settings. Further prospective clinical studies are warranted to validate these findings.

Second, the requirement for serial MRI acquisitions poses a limitation in terms of clinical applicability, particularly in population-based screening settings where assessments are typically based on single time-point imaging. However, the proposed model was not designed to determine lesion presence at a single time point, but rather to detect temporal imaging changes and stratify risk in clinical scenarios where follow-up MRI is available. Thus, the requirement for serial imaging represents an inherent constraint of longitudinal change detection rather than a flaw of model implementation. Future studies should explore complementary strategies that enable application of this framework in clinical settings without serial imaging, thereby expanding its potential clinical utility.

Third, although the ground truth was established by expert consensus with reference to multiparametric imaging, human annotation is inherently imperfect. Finally, the number of symptomatic recurrent events in the asymptomatic cohort was modest, leading to wide confidence intervals in the survival analysis and limiting the number of covariates included in the adjusted models. Nevertheless, the consistent association observed in both unadjusted and adjusted analyses supports the robustness of our findings.

In conclusion, we show that a deep learning-based approach enables reliable detection of new

ischemic lesions on serial FLAIR MRI. The independent association between AI-detected silent brain infarctions and subsequent stroke underscores the potential clinical relevance of automated longitudinal imaging analysis as a standardized risk marker, warranting further prospective evaluation.

Methods

Study Design and Setting

This was a multicenter retrospective cohort study conducted at 2 tertiary referral hospitals in South Korea. The study consisted of 3 phases: (1) development of a deep learning model for detecting new ischemic lesions on serial FLAIR MRI, (2) internal and external validation of diagnostic performance, and (3) prognostic analysis of AI-detected silent brain infarctions (SBIs) in an asymptomatic cohort. Imaging data were collected between August 2008 and March 2023. The study was approved by the Institutional Review Boards of Haeundae Paik Hospital (IRB No. 2021-09-025) and Busan Paik Hospital (IRB No. 2023-05-023). The requirement for informed consent was waived by the institutional review boards due to the retrospective nature of the study and the use of deidentified data.

Participants

All consecutive adult patients (aged ≥ 18 years) who underwent brain MRI, including DWI and FLAIR sequences, for suspected stroke during the study period were screened to construct a dataset for AI model development and validation. Patients with at least 2 temporally distinct MRI examinations (baseline and follow-up) were eligible. Exclusion criteria included presence of alternative dominant pathology on imaging (eg, severe leukoaraiosis, primary brain tumor, or intracranial hemorrhage) that precluded reliable assessment of ischemic lesions.

Of 15,267 patients screened across the 2 hospitals, 1,258 underwent at least 2 MRI sessions. After applying exclusion criteria, 1,055 patients were included in the final analysis (Hospital_A: $n=633$; Hospital_B: $n=422$) (**Supplementary Figure 1**). Data from Hospital_A were used for model training and internal validation, while Hospital_B served as an external validation cohort and provided the subset for prognostic analyses. Patient demographics, vascular risk factors, and clinical outcomes were obtained from institutional stroke registries and electronic health records.

Image Data Preparation and Annotation

For each patient, paired FLAIR scans (baseline and follow-up) were spatially coregistered using Advanced Normalization Tools to align anatomic slices.¹⁰ Registration was performed intra-patient using a rigid-body (linear) transformation, with the follow-up scan treated as the fixed image and the baseline scan as the moving image. Each follow-up slice was matched to the corresponding baseline slice to generate slice pairs. A fixed set of 24 axial slice pairs covering the brain was used per scan pair after skull stripping.

Preprocessing included skull stripping using a pretrained 3D U-Net–based brain extraction model implemented in ANTsPyNet, resampling to a uniform in-plane resolution of 256×256 .

pixels, and histogram matching to minimize intensity variation across sessions. Input images were constructed as 2-channel arrays (baseline and follow-up) (**Supplementary Figure 2**).

Ground truth labels were established at the slice-pair level. Five stroke neurologists independently reviewed images to determine whether a new ischemic lesion had appeared on follow-up compared with baseline, classifying each pair as “changed” or “unchanged” (**Supplementary Figure 3**). Consensus meetings were held to adjudicate ambiguous cases. Radiology reports incorporating multiparametric sequences (DWI, ADC, SWI, T1, and T2) were also reviewed to minimize misclassification. Conservative annotation rules were applied, excluding subtle nonspecific changes and labeling only lesions with clear imaging evidence of new infarction (**Supplementary Method 1 and Supplementary Figure 4**).

To prevent data leakage, data partitioning was performed at the patient level, such that all slice pairs from a given patient were assigned exclusively to a single subset (training, validation, or test). Model training and evaluation were conducted at the slice level, with each slice pair treated as an independent sample.

All preprocessing steps were implemented in Python using open-source libraries, including Advanced Normalization Tools Python (ANTsPy) and its deep-learning extension, ANTsPyNet.

Deep Learning Model Development

The deep learning framework was designed as a binary classifier to determine whether a new ischemic lesion was present on paired FLAIR slices. A supervised contrastive learning (SupCon) approach was adopted to enhance the learning of discriminative representations.¹¹ Model development process comprised two stages. In the first stage, a 2-dimensional convolutional neural network encoder was optimized to bring embeddings of pairs with identical labels closer while separating embeddings of pairs with different labels. The encoder architecture consisted of 3 convolutional layers (128 filters, kernel size 3×3, stride 2, rectified linear unit activation), followed by global average pooling and L2 normalization. Data augmentation included random rotations ($\pm 5^\circ$) and translations (up to 10% of image dimensions). The encoder was pretrained using SupCon loss with a temperature parameter of 0.1 and the Adam optimizer (learning rate 0.001) for 100 epochs.

In the second stage, the pretrained encoder was frozen and used to extract features. A sigmoid output layer was appended to the pretrained encoder and trained with binary cross-entropy loss to predict the probability of new infarction for each slice pair. Training was performed for 100 epochs with Adam optimization. Model development was implemented in Python 3.9 using TensorFlow, and training was conducted on NVIDIA RTX A6000 GPUs. To facilitate interpretability, Gradient-weighted Class Activation Mapping (Grad-CAM) was applied to visualize regions that most contributed to the model's predictions, thereby providing heatmaps superimposed on FLAIR images.⁹ The code for the entire preprocessing pipeline and model implementation is publicly available on GitHub (<https://github.com/Hwan-ho/SupConFLAIRChange>).

Model Performance Evaluation

Model performance was assessed at both slice and patient levels. For slice-level evaluation, predicted probabilities were dichotomized using an optimal threshold derived from Youden's J

statistic on the internal validation set.¹² Diagnostic metrics included sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and the area under the receiver operating characteristic curve (AUC), with 95% confidence intervals estimated by bootstrap resampling. Calibration was evaluated using Brier scores and calibration plots.^{13,14}

For patient-level analysis, the average predicted probability across the 3 highest-scoring slices was used as the representative score for each patient. A patient was considered positive if at least 3 slices were annotated as new lesions. Patient-level AUC, sensitivity, and specificity were then calculated for internal and external validation cohorts.

Clinical Outcome Analysis in the Asymptomatic Cohort

To evaluate the prognostic utility of the model, we analyzed a subset of the external validation cohort. This asymptomatic cohort included patients who, at the time of follow-up MRI (index date), exhibited no new or worsening neurological symptoms compared with baseline. Patients with stable sequelae of prior stroke were eligible provided no acute clinical deterioration was documented.

Using the deep learning model, each patient was classified as AI-detected SBI-positive if at least 1 slice exceeded the predefined probability threshold for new infarction, or AI-detected SBI-negative otherwise. Patients were followed longitudinally for the occurrence of symptomatic ischemic stroke, defined as the first new neurological deficit confirmed by treating physicians and supported by imaging evidence. Outcomes were ascertained through linkage of institutional stroke registries, electronic health records, and national mortality data. Follow-up extended up to 7 years after the index MRI.

Statistical Analysis

Cumulative incidence of symptomatic stroke was estimated using the Kaplan–Meier method, with group comparisons performed by log-rank test. Cox proportional hazards regression was used to estimate hazard ratios (HRs) with 95% confidence intervals (CIs) for the association between AI-detected SBI and subsequent stroke. In the primary multivariable model, covariates included age, diabetes mellitus, and atrial fibrillation, selected a priori as established vascular risk factors and to minimize overfitting given the limited number of outcome events. Proportional hazards assumptions were verified using Schoenfeld residuals. For clinical variables in the asymptomatic cohort, missing values were not imputed; analyses were performed using all available data. Because this was a retrospective study with no prior data on automated SBI detection, no formal sample size calculation was performed; all eligible patients during the study period were included.

All analyses were performed with R version 4.2 (R Foundation for Statistical Computing). Two-sided P values <.05 were considered statistically significant.

Data Availability

The datasets generated and/or analyzed during the current study are not publicly available because they contain individual-level patient imaging data and are subject to institutional data use and privacy regulations. Deidentified participant data and analytic code used in this study are available from the corresponding author upon reasonable request.

Code Availability

All code for data preprocessing, model training, and evaluation is available at <https://github.com/Hwan-ho/SupConFLAIRChange> under an open-source license.

Acknowledgment

This research was supported by the K-Brain Project of the National Research Foundation (NRF) funded by the Korean government (MSIT) (No. RS-2023-00265393). This research has been supported by the HANDOK JESEOK FOUNDATION. This work was supported by the research fund of Hanyang University (HY-202500000001209). This work was supported by Incheon National University Research Grant in 2024.

Author Contributions

S.P. and H.-h.C. conceived and designed the study. J.L., H.C.K., S.Y.L., and J.H.S. collected and curated the data. H.-h.C. and S.P. performed imaging analysis. H.-h.C., H.P., and S.P. developed the deep learning model. S.P. conducted the statistical analysis. S.P. and H.-h.C. drafted the manuscript. S.P., B.J., D.L., W.-K.S., and J.-M.J. critically revised the manuscript for important intellectual content. S.P. supervised the study. All authors reviewed and approved the final version of the manuscript.

Competing Interests

All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Seongho Park is an inventor on a pending South Korean patent application (Application No. 10-2025-0037382) owned by the Inje University Industry–Academic Cooperation Foundation related to automated analysis of serial brain MRI. All other authors declare no competing financial or non-financial interests.

Reference

1. Fang, R. *et al.* Risk factors and clinical significance of post-stroke incident ischemic lesions. *Alzheimers Dement.* **20**, 8412–8428 (2024).
2. Longstreth, W. T. *et al.* Incidence, Manifestations, and Predictors of Brain Infarcts Defined by Serial Cranial Magnetic Resonance Imaging in the Elderly. *Stroke* **33**, 2376–2382 (2002).
3. Vermeer, S. E., Longstreth, W. T. & Koudstaal, P. J. Silent brain infarcts: a systematic review. *Lancet Neurol.* **6**, 611–619 (2007).
4. Sacco, R. L. *et al.* An updated definition of stroke for the 21st century: a statement for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* **44**, 2064–2089 (2013).
5. Rudick, R. *et al.* Inter-Rater Reliability among Clinical Raters for New MRI Lesions in MS Patients (P03.068). *Neurology* **78**, P03.068-P03.068 (2012).
6. Cimflova, P. *et al.* MRI Diffusion-Weighted Imaging to Measure Infarct Volume: Assessment of Manual Segmentation Variability. *J. Neuroimaging Off. J. Am. Soc. Neuroimaging* **31**, 541–550 (2021).
7. Altay, E. E. *et al.* Reliability of Classifying Multiple Sclerosis Disease Activity Using Magnetic Resonance Imaging in a Multiple Sclerosis Clinic. *JAMA Neurol.* **70**, 338–344 (2013).
8. Chen, X. *et al.* Recent advances and clinical applications of deep learning in medical image analysis. *Med. Image Anal.* **79**, 102444 (2022).
9. Selvaraju, R. R. *et al.* Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. *Int. J. Comput. Vis.* **128**, 336–359 (2020).
10. Tustison, N. J. *et al.* The ANTsX ecosystem for quantitative biological and medical imaging. *Sci. Rep.* **11**, 9068 (2021).
11. Khosla, P. *et al.* Supervised Contrastive Learning. Preprint at <https://doi.org/10.48550/arXiv.2004.11362> (2021).

12. Hassanzad, M. & Hajian-Tilaki, K. Methods of determining optimal cut-point of diagnostic biomarkers with application of clinical data in ROC analysis: an update review. *BMC Med. Res. Methodol.* **24**, 84 (2024).
13. Steyerberg, E. W. *et al.* Assessing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiol. Camb. Mass* **21**, 128–138 (2010).
14. tutorial on calibration measurements and calibration models for clinical prediction models | Journal of the American Medical Informatics Association | Oxford Academic.
https://academic.oup.com/jamia/article/27/4/621/5762806?utm_source=chatgpt.com&login=true.
15. Vermeer, S. E. *et al.* Silent brain infarcts and white matter lesions increase stroke risk in the general population: the Rotterdam Scan Study. *Stroke* **34**, 1126–1129 (2003).
16. Feng, C., Bai, X., Xu, Y., Hua, T. & Liu, X.-Y. The ‘silence’ of silent brain infarctions may be related to chronic ischemic preconditioning and nonstrategic locations rather than to a small infarction size. *Clinics* **68**, 365–369 (2013).
17. Sigurdsson, S. *et al.* Incidence of Brain Infarcts, Cognitive Change, and Risk of Dementia in the General Population. *Stroke* **48**, 2353–2360 (2017).
18. Wardlaw, J. M. *et al.* Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. *Lancet Neurol.* **12**, 822–838 (2013).
19. Kim, C., Gadgil, S. U. & Lee, S.-I. Transparency of medical artificial intelligence systems. *Nat. Rev. Bioeng.* 1–19 (2025) doi:10.1038/s44222-025-00363-w.
20. Kang, D.-W. *et al.* Silent new ischemic lesions after index stroke and the risk of future clinical recurrent stroke. *Neurology* **86**, 277–285 (2016).
21. Lee, E.-J., Kang, D.-W. & Warach, S. Silent New Brain Lesions: Innocent Bystander or Guilty Party? *J. Stroke* **18**, 38–49 (2015).

Table 1. Dataset Characteristics for Model Development and Slice-Level Annotation Distribution

Dataset	Internal Dataset (Hospital_A)	External Dataset (Hospital_B)	Total
Patients, n	633	422	1,055
Male (%)	353 (55.8)	226 (53.6)	579 (54.9)
Age (SD)	72.1 (12.8)	72.4 (13.4)	72.2 (13.0)
Time interval between scans, Median in days [IQR]	25 [5-277]	40 [5-330]	28 [5-295]
Total number of paired imaging slices	15,357	10,094	25,451
New infarct lesion (%)	2,063 (13.4)	1,721 (17.1)	3,784 (14.9)
Acute to subacute	1,568 (10.2)	1,127 (11.2)	2,695 (10.6)
Chronic	495 (3.2)	594 (5.9)	1,089 (4.3)
No new lesion (%)	13,294 (86.6)	8,373 (83.0)	21,667 (85.1)

Values are presented as n (%), mean (SD), or median (IQR) as appropriate. SD, standard deviation; IQR, interquartile range. Patient characteristics are reported at the patient level, whereas lesion counts are reported at the slice level, corresponding to the unit of analysis used for model training.

Table 2. Clinical Characteristics of the Asymptomatic Cohort according to AI-based Detection of New Silent Brain Infarction

	Total (n=307)	AI-positive (n=182)	AI-negative (n=125)
Age, y	71.5 (13.5)	71.4 (13.3)	71.6 (13.7)
Sex, male	174 (56.7%)	102 (56.0%)	72 (57.6%)
BMI, kg/m ²	25.0 (26.0)	24.9 (24.9)	25.1 (27.6)
Hypertension	159 (57.2%)	94 (57.0%)	65 (57.5%)
Diabetes	89 (29.0%)	52 (41.6%)	37 (20.3%)
Dyslipidemia	32 (11.5%)	17 (15.0%)	15 (9.1%)
Smoking	68 (24.5%)	36 (21.8%)	32 (28.3%)
Atrial fibrillation	54 (19.4%)	40 (24.2%)	14 (12.4%)
CHD	38 (13.7%)	26 (15.8%)	12 (10.6%)
Malignancy	10 (3.6%)	7 (4.3%)	3 (2.7%)
NIHSS, median	3.0 [2.0, 7.0]	4.0 [2.0, 10.0]	2.0 [1.0, 5.0]
TOAST			
LAA	94 (33.9%)	71 (43.0%)	23 (20.5%)
SVO	66 (23.8%)	26 (15.8%)	40 (35.7%)
CE	56 (20.2%)	43 (26.1%)	13 (11.6%)
OD	11 (4.0%)	7 (4.2%)	4 (3.6%)
UD	50 (16.3%)	18 (14.4%)	32 (17.6%)
SBP, mmHg	155.4 (29.9)	155.7 (30.0)	155.0 (29.8)
DBP, mmHg	85.6 (17.5)	84.6 (17.3)	87.1 (17.9)
Glucose, mg/dL	154.1 (71.9)	155.7 (74.5)	151.8 (68.2)
HbA1c, %	6.44 (1.53)	6.46 (1.60)	6.40 (1.41)
LDL, mg/dL	108.8 (42.2)	103.9 (42.1)	116.2 (41.4)
HDL, mg/dL	45.7 (12.9)	44.9 (13.3)	46.9 (12.3)
TG, mg/dL	137.4 (98.1)	128.6 (89.5)	150.6 (108.8)
CRP, mg/L	0.84 (2.40)	0.97 (2.69)	0.65 (1.88)
PT (INR)	1.05 (0.20)	1.07 (0.25)	1.02 (0.09)
Medication at discharge			
Antiplatelet agent	223 (80.2%)	122 (73.9%)	101 (89.4%)
Anticoagulant	54 (19.4%)	41 (24.8%)	13 (11.5%)

Values are presented as mean (standard deviation), median [interquartile range], or number (%). All clinical characteristics were measured at the time of base (initial) scan, and each statistic was calculated among patients with available data for the respective variable. The AI risk stratification (AI-positive vs. AI-negative) was determined at the index date (follow-up scan) based on the AI model's detection of new infarctions between base and follow-up scans.

BMI indicates body mass index; CHD, coronary heart disease; NIHSS, National Institutes of Health Stroke Scale; TOAST, Trial of ORG 10172 in Acute Stroke Treatment; LAA, large artery atherosclerosis; SVO, small vessel occlusion; CE, Cardioembolism; OD, other determined etiology; UD, undetermined etiology; SBP, systolic blood pressure; DBP, diastolic blood pressure; LDL, low-density lipoprotein; HDL, high-density lipoprotein; TG, triglycerides; CRP, c-reactive protein; PT (INR), prothrombin time (international normalized ratio).

Table 3. Cox Proportional Hazards Analysis for Symptomatic Ischemic Stroke Recurrence

	Hazard Ratio (95% CI)	P-value
Univariable Analysis		
AI-detected SBI (Positive vs. Negative)	4.11 (1.40–12.04)	0.010
Multivariable Analysis *		
AI-detected SBI	3.83 (1.30–11.30)	0.015
Age (per 1-year increase)	1.03 (0.985–1.07)	0.216
Diabetes Mellitus	2.51 (1.10–5.75)	0.030
Atrial Fibrillation	1.55 (0.58–4.14)	0.380

SBI, Silent Brain Infarct.

* Multivariable model was adjusted for age, diabetes mellitus, and atrial fibrillation.

Figure 1. Slice-Level Performance of the deep learning model in internal and external dataset

(A) Discrimination of the model at the slice level, as assessed by receiver operating characteristic curves in the internal validation and external test sets. (B) Calibration of predicted probabilities in the internal validation and external test sets.

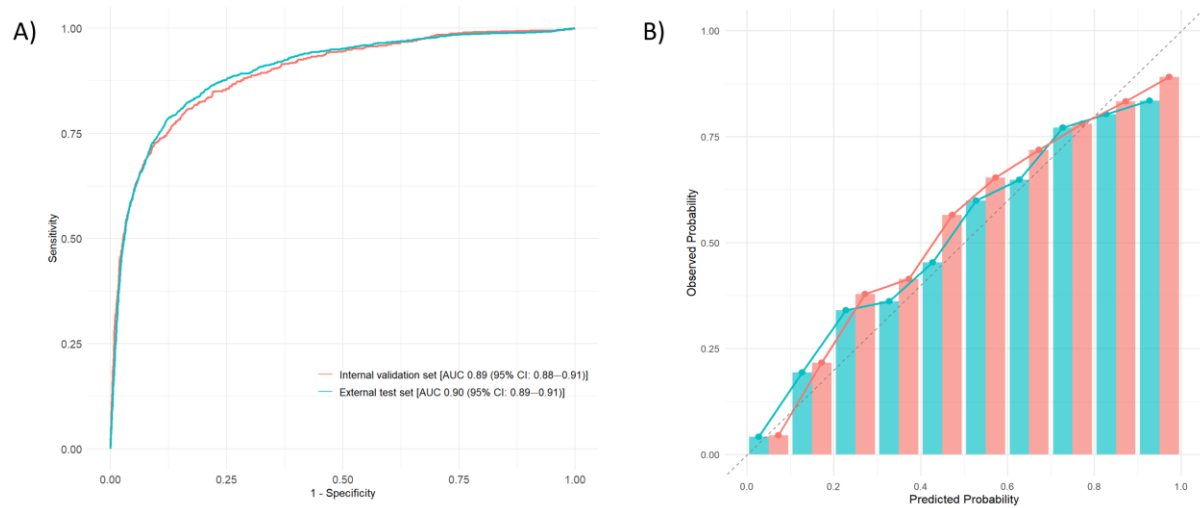


Figure 2. Representative cases of newly developed infarcts identified by the model

Panels A–D show examples where newly developed infarcts (red circles on follow-up FLAIR scans) were correctly localized by the Grad-CAM activation maps, with highlighted regions corresponding to the true lesion sites. Panel D illustrates a case with both a true new infarct (red circle) and a pre-existing signal change (yellow circle). The model selectively activated the new lesion while appropriately ignoring the pre-existing region, indicating robust specificity in lesion detection.

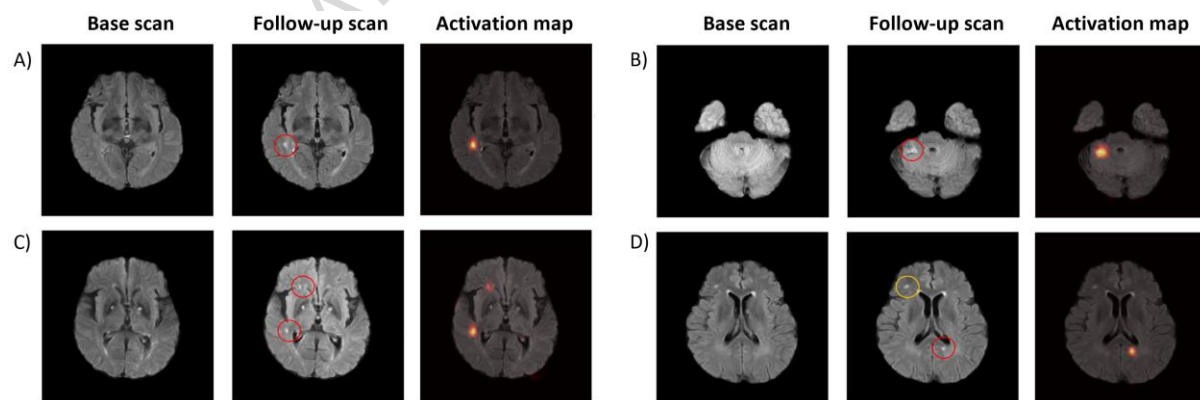


Figure 3. Kaplan-Meier Analysis for Recurrent Symptomatic Ischemic Stroke

Kaplan-Meier curves show the cumulative incidence of recurrent symptomatic ischemic stroke in the asymptomatic cohort. The AI-positive group (red line), defined as patients with AI-detected new silent brain infarction, had a significantly higher cumulative stroke incidence compared to the AI-negative group (blue line) (log-rank test, $p = 0.0051$).

