



OPEN ACCESS

Original research

Automated detection of large vessel occlusion using deep learning: a pivotal multicenter study and reader performance study

Jae Guk Kim,¹ Sue Young Ha ,² You-Ri Kang,³ Hotak Hong ,² Dongmin Kim,² Myungjae Lee,² Leonard Sunwoo ,⁴ Wi-Sun Ryu ,² Joon-Tae Kim ³

► Additional supplemental material is published online only. To view, please visit the journal online (<https://doi.org/10.1136/jnis-2024-022254>).

¹Department of Neurology, Daejeon Eulji University Hospital, Daejeon, Daejeon, Korea

²Artificial Intelligence Research Center, JIL Inc, Seoul, Korea

³Department of Neurology, Chonnam National University Medical School, Gwangju, Korea

⁴Department of Radiology, Seoul National University Bundang Hospital, Seongnam, Gyeonggi-do, Korea

Correspondence to

Dr Wi-Sun Ryu; wisunryu@gmail.com and Professor Joon-Tae Kim; alldelight2@jnu.ac.kr

JGK, SYH and Y-RK contributed equally.

Received 17 July 2024

Accepted 6 September 2024

ABSTRACT

Background To evaluate the stand-alone efficacy and improvements in diagnostic accuracy of early-career physicians of the artificial intelligence (AI) software to detect large vessel occlusion (LVO) in CT angiography (CTA).

Methods This multicenter study included 595 ischemic stroke patients from January 2021 to September 2023. Standard references and LVO locations were determined by consensus among three experts. The efficacy of the AI software was benchmarked against standard references, and its impact on the diagnostic accuracy of four residents involved in stroke care was assessed. The area under the receiver operating characteristic curve (AUROC), sensitivity, and specificity of the software and readers with versus without AI assistance were calculated.

Results Among the 595 patients (mean age 68.5 ± 13.4 years, 56% male), 275 (46.2%) had LVO. The median time interval from the last known well time to the CTA was 46.0 hours (IQR 11.8–64.4). For LVO detection, the software demonstrated a sensitivity of 0.858 (95% CI 0.811 to 0.897) and a specificity of 0.969 (95% CI 0.943 to 0.985). In subjects whose symptom onset to imaging was within 24 hours ($n=195$), the software exhibited an AUROC of 0.973 (95% CI 0.939 to 0.991), a sensitivity of 0.890 (95% CI 0.817 to 0.936), and a specificity of 0.965 (95% CI 0.902 to 0.991). Reading with AI assistance improved sensitivity by 4.0% (2.17 to 5.84%) and AUROC by 0.024 (0.015 to 0.033) (all $P<0.001$) compared with readings without AI assistance.

Conclusions The AI software demonstrated a high detection rate for LVO. In addition, the software improved diagnostic accuracy of early-career physicians in detecting LVO, streamlining stroke workflow in the emergency room.

INTRODUCTION

Advancements in stroke imaging and procedural devices have extended the window for endovascular therapy (EVT) in patients with large vessel occlusion (LVO).¹ Recent randomized clinical trials have established a new standard of care for patients with LVO who arrive at the hospital within 6 to 24 hours of their last known well time.^{2,3} The triage process for these clinical trials primarily relies on magnetic resonance (MR) perfusion or CT perfusion to identify clinical or tissue mismatches.^{2,3} However, the majority of primary stroke facilities

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Advancements in stroke imaging have extended the endovascular treatment window for patients with large vessel occlusion (LVO), with CT angiography and artificial intelligence (AI) tools improving detection and treatment efficiency.

WHAT THIS STUDY ADDS

⇒ We proved the efficacy of the AI algorithm using multicenter datasets. Additionally, we demonstrated that the AI algorithm improved the sensitivity of LVO detection in early-career physicians.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ AI software may streamline stroke workflow by improving the diagnostic accuracy of early-career physicians.

worldwide lack widespread access to these advanced imaging techniques.⁴ Recent studies have emphasized using more accessible imaging methods like CT angiography (CTA). The CT for Late Endovascular Reperfusion (CLEAR) trial demonstrated comparable clinical outcomes between patients selected using non-contrast CT with CTA and those selected with CT or MR perfusion.⁵ Furthermore, a sub-study conducted by the HERMES (Highly Effective Reperfusion Evaluated in Multiple Endovascular Stroke Trials) collaboration has expanded on this notion within the early time frame (0–6 hours), showing that the rates of favorable functional outcomes were similar between patients who underwent CT perfusion and those who did not.⁶

Initially, 66% of EVT candidates were directed to centers incapable of performing EVT,⁷ despite having better chances of favorable outcomes at EVT-capable sites. Therefore, it is imperative for non-EVT-capable centers to reliably and promptly detect LVOs at all times—24 hours a day, 7 days a week—facilitating the rapid transfer of patients to EVT-capable facilities. However, the scarcity of vascular experts poses a challenge for many non-EVT-capable centers. Even in EVT-capable centers, an automated tool for screening CTA to detect the presence of LVOs could enhance operational efficiency, optimize staffing, and reduce the time from



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Kim JG, Ha SY, Kang Y-R, et al. *J NeuroIntervent Surg* Epub ahead of print: [please include Day Month Year]. doi:10.1136/jnis-2024-022254

patient arrival to the initiation of the procedure by facilitating prompt LVO detection.

This multicenter study aimed to demonstrate the efficacy of a fully automated, deep learning-based software (JLK-LVO, JLK Inc., Republic of Korea) in detecting LVO in patients with acute ischemic stroke. The algorithm's performance was evaluated against a standard reference established through consensus among stroke experts. Furthermore, we examined whether the implementation of artificial intelligence (AI) software enhances the diagnostic accuracy of early-career physicians compared with their performance without AI assistance.

METHODS

Study design and data source

This study is reported in accordance with the STARD (Standards for Reporting of Diagnostic Accuracy Studies) reporting guidelines. From January 2021 to September 2023, we retrospectively included patients with ischemic stroke who were admitted to two university hospitals within 7 days of symptom onset and underwent CTA for evaluation of vessel status. Among 603 eligible patients, we excluded subjects with poor image quality ($n=8$), severe metallic artifacts ($n=4$), insufficient contrast filling ($n=1$), and source image thickness >2 mm ($n=3$), leaving 595 patients for analysis (online supplemental figure 1). Stroke subtypes were determined using an MR-based stroke classification algorithm by the attending physician at each hospital, as previously described.⁸

Sample size estimates

The primary objective is to test whether the sensitivity and specificity of the software are comparable to prespecified criteria based on prior studies.^{9–13} For the sensitivity analysis, an LVO-positive sample size of 270 patients will have 90% power to detect, with a one-sided test at a 2.5% significance level, a prespecified lower bound of 0.7489. For the specificity analysis, an LVO-negative sample size of 320 patients will have 90% power to detect, with a one-sided test at a 2.5% significance level, a prespecified lower bound of 0.9155. This sample size accounts for a 2% dropout rate.

Definition of large vessel occlusion

In this study, LVO was defined as arterial occlusion involving the intracranial segment of the internal carotid artery (ICA), the M1 segment of the middle cerebral artery (MCA-M1), and the M2 segment of the MCA (MCA-M2). The intracranial ICA refers to the segment of the ICA extending from the petrous part to the bifurcation of the MCA and the anterior cerebral artery (ACA).¹⁴ The MCA-M1 segment is defined as the portion of the MCA from the MCA-ACA bifurcation to the MCA branching point. The MCA-M2 segment refers to the portion of the MCA ascending vertically along the Sylvian fissure from the MCA branching point.¹⁴ In our study, occlusions at the intracranial ICA or MCA-M1 were categorized as intracranial LVO, whereas occlusions at the MCA-M2 without proximal arterial involvement were designated as isolated MCA-M2 occlusions. In cases where the MCA divided early, we employed a functional rather than a traditional definition: the artery segment closest to its origin was designated as the M1 segment, and branches further downstream were classified as M2 segments.¹⁵ To ascertain the presence of LVO, two experienced vascular neurologists, each with at least 5 years of experience, meticulously reviewed the CTA source images, maximum intensity projection (MIP) images, and three-dimensional (3D) rendering

images, in addition to patients' MRI scans and symptom data. In instances of labeling disagreement, a third reviewer made the final decision.

Deep learning-based software

Source images of CTA with a slice thickness between 0.5 and 2 mm were processed using commercially available deep learning-based software (JLK-LVO, JLK Inc., online supplemental figure 2).^{16–17} In brief, an automated algorithm selects slices from source images to construct MIP images. The vessel segmentation involves a two-dimensional U-Net based on the Inception Module,¹⁸ which is trained to segment vessels in axial MIP images. Following this, an LVO detection algorithm combines the vessel masks into a compressed image, which is used to train an EfficientNetV2 model.¹⁹ The saliency map is visualized when the probability of LVO exceeds 10%. The side of the LVO is determined by the software based on the saliency map.

Study design

Stand-alone performance

LVO probability scores predicted by the software were used to evaluate the standalone performance metrics of the software, including the area under the receiver operating characteristic curve (AUROC), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). To construct the confusion matrix, we set the threshold for LVO probability at 50%. We defined a true positive when the predicted side of the LVO from the AI software matched the expert consensus, except in cases of bilateral occlusion. If the diagnosis was correct but the side of the LVO was discordant between the AI software and the expert consensus, we designated the case as a false negative. For bilateral occlusions, a true positive was defined when the AI software-generated heatmap was present on both sides, with the smaller side being at least 50% of the larger side.

Reader assessment study

A retrospective, crossover-designed trial was conducted to assess the efficacy of software in aiding diagnostic decisions for detecting LVO in CTA. The multi-reader study involved four residents specializing in the care of ischemic stroke from the fields of radiology, neurology, neurosurgery, and emergency medicine. Before the evaluation, the reviewers were divided into groups A and B. Group A was provided with original CTA images, including MIP and 3D rendering images, along with the AI software's results (with AI assistance; online supplemental figures 1 and 3). Conversely, group B received only the original CTA images, including MIP and 3D rendering images (without AI assistance). The software presented segmented and merged vessel images with a heatmap and an LVO probability score. A custom image viewer (MEDIHUB Stroke, JLK Inc., stroke.medihub.ai) was utilized to visualize the CTA and the software's output. The reviewers were blinded to the standard reference confirmed by expert consensus. A second assessment was conducted 4–5 weeks after a washout period,²⁰ during which study subjects were shuffled and randomly allocated new study numbers. In this second assessment, group A reassessed the CTA images without the aid of the software's findings (without AI assistance), while group B reassessed the CTA images with the software's results (with AI assistance). Readers determined the presence and the side of LVO and assigned a confidence score for the presence of LVO using a 5-point scale.

Statistical analysis

Using the t-test or rank sum test for continuous variables, and the χ^2 test for categorical variables as appropriate, we compared baseline characteristics stratified by the presence of LVO. To evaluate the accuracy of the software in diagnosing LVO, we computed the AUROC, as well as sensitivity, specificity, PPV, and NPV. A 1000-repeat bootstrap analysis was employed to calculate the 95% confidence intervals (95% CI) for all parameters. The AUROC was used in combination with the DeLong method²¹ to compute the standard error of the AUROC. The cut-off for the LVO probability score used in the analysis was set at 0.5. We conducted an additional analysis to determine the optimal threshold that would yield the maximum Youden index (sensitivity+specificity–1). Given that the software is primarily intended for screening LVO, we also computed specificity, PPV, and NPV at the fixed sensitivity level of 0.90.

After dividing the subjects into isolated MCA-M2 occlusion and intracranial LVO groups, we reran the analysis for subgroup analysis. In this analysis, patients without LVO were included as the control group for both subgroups. Considering the EVT time window, we repeated the analysis after excluding patients whose onset to CTA time exceeded 24 hours. After stratifying subjects by stroke subtype, we reran the analysis.

The Obuchowski-Rockette method was used for analyzing multireader multicase (MRMC) studies, along with the MRMCaov library,²² for all analyses of improvement of diagnostic performance with versus without AI assistance. This method tested the null hypothesis that the average AUROC of the readers without AI assistance was equal to that with AI assistance. The Obuchowski-Rockette method accounts for the fact that, in an MRMC study, the same cases are evaluated by each reader. Consequently, the error terms are assumed to be equivariant among readers and cases, rather than independent. We also calculated sensitivity, specificity, accuracy, and their differences between with and without AI assistance. A P value <0.05 was considered statistically significant. The analyses were conducted using R version 4.2.3, STATA software (version 16.0, College Station, TX) and MedCalc (version 17.2, MedCalc Software, Ostend, Belgium, 2017).

RESULTS

Patient characteristics

Among 595 patients, 275 (46.2%) were diagnosed with LVO. Specifically, 213 patients had intracranial LVO, and 62 patients exhibited isolated MCA-M2 occlusion. Details regarding the occlusion sites are presented in online supplemental table 1. The mean±SD age was 68.5±13.4 years and 332 (56%) were male. The median time interval between the last known well time to CTA was 46 hours (IQR 11.8–64.4). Patients with LVO were older, had shorter time intervals between the onset of symptoms and imaging, and presented with more severe strokes compared with those without LVO (table 1).

Stand-alone performance of AI software

The software achieved an AUROC of 0.961 (95% CI 0.945 to 0.976) (online supplemental figure 4A) at a cut-off point of 0.50. The sensitivity, specificity, PPV, and NPV were 0.858 (95% CI 0.811 to 0.897), 0.969 (95% CI 0.943 to 0.985), 0.959 (95% CI 0.927 to 0.980), and 0.888 (95% CI 0.850 to 0.919), respectively (table 2). The highest Youden index was observed at the optimal cut-off point of 0.362, yielding sensitivity, specificity, PPV, and NPV values of 0.880 (95% CI 0.836 to 0.916), 0.953 (95% CI 0.924 to 0.974), 0.942 (95% CI 0.906 to 0.967), and

Table 1 Baseline characteristics stratified by the presence of large vessel occlusion

	No LVO (n=320)	LVO (n=275)	P value
Age	66.4±13.1	71.2±13.3	<0.001
Sex, men	180 (56.3%)	152 (55.3%)	0.81
Onset to admission, hours	6.8 (2.5–25.8)	2.7 (1.1–9.3)	<0.001*
Onset to imaging, hours	49.2 (18.2–68.2)	41.7 (4.7–60.3)	<0.001*
Initial NIHSS	2 (1–4)	11 (6–15)	<0.001*
Previous stroke	51 (15.9%)	50 (18.5%)	0.42
Hypertension	181 (56.6%)	158 (58.3%)	0.67
Diabetes	98 (30.6%)	70 (25.8%)	0.20
Atrial fibrillation	53 (16.6%)	111 (41.0%)	<0.001
Revascularization therapy	40 (12.5%)	120 (43.6%)	<0.001
Intravenous only	27 (8.4%)	22 (8.1%)	
Endovascular only	8 (2.5%)	43 (15.9%)	
Combined	5 (1.6%)	55 (20.3%)	

Data presented as mean±SD, median (IQR), and number (%).

*Rank-sum test was used.

LVO, large vessel occlusion; NIHSS, National Institutes of Health Stroke Scale.

0.902 (95% CI 0.866 to 0.932), respectively. At a given sensitivity of 0.90, the specificity was 0.916 (95% CI 0.825 to 0.968).

When limiting the analysis to intracranial LVO, the AUROC was 0.961 (95% CI 0.940 to 0.975) (online supplemental figure 4B). The sensitivity and specificity at the cut-off point of 0.5 were 0.873 (95% CI 0.821 to 0.915) and 0.969 (95% CI 0.943 to 0.985), respectively (online supplemental table 2). Restricting the analysis to isolated MCA-M2 occlusion, the software demonstrated an AUROC of 0.928 (95% CI 0.895 to 0.961) (online

Table 2 Diagnostic performance of software detecting large vessel occlusion

	Confusion matrix	Prediction	
		LVO	No LVO
Threshold of 0.50	Ground truth, LVO	236	39
	Ground truth, no LVO	10	310
	Sensitivity (95% CI)	0.858 (0.811 to 0.897)	
	Specificity (95% CI)	0.969 (0.943 to 0.985)	
	PPV (95% CI)	0.959 (0.927 to 0.980)	
	NPV (95% CI)	0.888 (0.850 to 0.919)	
Optimal threshold	Youden (J) index (95% CI)	0.833 (0.784 to 0.869)	
	J _{max} cut-off point	0.362	
	J _{max} sensitivity (95% CI)	0.880 (0.836 to 0.916)	
	J _{max} specificity (95% CI)	0.953 (0.924 to 0.974)	
	J _{max} PPV (95% CI)	0.942 (0.906 to 0.967)	
	J _{max} NPV (95% CI)	0.902 (0.866 to 0.932)	
Fixed sensitivity of 0.90	Sens ₉₀ specificity (95% CI)	0.916 (0.825 to 0.968)	
	Sens ₉₀ PPV (95% CI)	0.902 (0.860 to 0.934)	
	Sens ₉₀ NPV (95% CI)	0.916 (0.880 to 0.944)	
	Sens ₉₀ cut-off point	0.221	

Jmax represents, across all thresholds, the maximum Youden index (sensitivity+specificity–1). As a secondary reference point, Jmax provides an optimality criterion with equal weighting for sensitivity and specificity. LVO, large vessel occlusion; NPV, negative predictive value; PPV, positive predictive value.

supplemental figure 4C), with a sensitivity of 0.692 (95% CI 0.578 to 0.792) and a specificity of 0.960 (95% CI 0.932 to 0.978).

Considering the EVT time window, the analysis was restricted to those whose symptom onset to imaging was within 24 hours ($n=195$), of which 109 (55.9%) demonstrated LVO. The software exhibited an AUROC of 0.973 (95% CI 0.939 to 0.991) (online supplemental figure 5). With the cut-off point set at 0.5, the sensitivity and specificity were 0.890 (95% CI 0.817 to 0.936) and 0.965 (95% CI 0.902 to 0.991), respectively.

After excluding 14 patients (2.4%) who lacked information on stroke subtype, 70 patients (11.8%) were classified as having large artery atherosclerosis, 161 (27.1%) as having cardioembolism, 205 (34.5%) as being undetermined, 139 (23.4%) as having small vessel occlusion, and six (1.0%) as having other-determined strokes. In the large artery atherosclerosis group, where 42 (60%) had LVO, the software achieved an AUROC of 0.953 (95% CI 0.874 to 0.989), with a sensitivity of 0.833 (95% CI 0.686 to 0.930) and a specificity of 1.00 (0.877 to 1.000) at a threshold of 0.5 (online supplemental figure 6). In the cardioembolism group, where 110 (68.3%) had LVO, the software achieved an AUROC of 0.965 (95% CI 0.924 to 0.988), with a sensitivity of 0.909 (95% CI 0.839 to 0.956) and a specificity of 0.980 (95% CI 0.896 to 1.000) at a threshold of 0.5. In the undetermined stroke group, where 108 (52.7%) had LVO, the software exhibited an AUROC of 0.954 (95% CI 0.916 to 0.978), with a sensitivity of 0.815 (95% CI 0.729 to 0.883) and a specificity of 0.969 (95% CI 0.912 to 0.994).

Reader assessment study

For all readers, the sensitivities of reading with AI assistance versus without were 91.82% (95% CI 90.04% to 93.37%) and 87.81% (95% CI 85.73% to 89.68%), respectively, with a mean difference of 4.00% (95% CI 2.17% to 5.84%, $P<0.001$) (table 3). The specificities were 95.70% and 96.30%, respectively,

with no statistical difference observed. Reading with AI assistance yielded higher accuracy, with a mean difference of 0.76% (95% CI 0.01% to 1.50%, $P=0.049$). The average AUROC with AI assistance (0.967, 95% CI 0.956 to 0.983) was significantly higher compared with that without AI assistance (0.944, 95% CI 0.906 to 0.977, $P<0.001$), with the difference of 0.024 (0.015 to 0.033) (figure 1).

DISCUSSION

In this multicenter study, fully automated AI software for detecting LVO achieved a sensitivity of 86% and a specificity of 97%. For isolated MCA-M2 occlusion, the AI software attained a sensitivity of 69% and a specificity of 96%. Additionally, the reader assessment study demonstrated that AI assistance significantly improved the sensitivity of LVO detection among early-career physicians in stroke care. To the best of our knowledge, this is the first study to prove the stand-alone efficacy and the enhancement in reader performance using AI assistance across multicenter datasets in detecting LVO on CTA.

A few AI software packages have been implemented in clinical practice. RAPID LVO (iSchemaView, Menlo Park, CA), that primarily relies on vessel density threshold assessment, showed higher performance in a pooled cohort from two stroke trials, with a sensitivity of 95% and specificity of 79%.²³ However, RAPID LVO requires different threshold settings based on the occlusion site (ICA, MCA-M1, or MCA-M2), which may pose challenges for early-career physicians in interpreting the results. Furthermore, recent studies have demonstrated a lower PPV for RAPID LVO, which may contribute to alarm desensitization, leading to missed alarms or delayed responses.²⁴ Viz LVO (Viz. AI, San Francisco, CA) and CINA LVO (Avicenna.ai, La Ciotat, France), using an end-to-end deep learning algorithm to detect LVO, showed high sensitivity and specificity in recent studies, with sensitivity ranging from 76% to 94%.^{11 13 25} In the present study, as a stand-alone tool, the software achieved a sensitivity

Table 3 Performance of readers with versus without artificial intelligence assistance

	With AI	Without AI	Difference	P value
Sensitivity				
All reviewers	91.82 (90.04 to 93.37)	87.81 (85.73 to 89.68)	4.00 (2.17 to 5.84)	<0.001
Reader 1	92.00 (88.14 to 94.92)	80.00 (74.78 to 84.56)	12.00 (7.43 to 16.57)	<0.001
Reader 2	86.55 (81.94 to 90.35)	86.50 (81.87 to 90.31)	0.00 (−3.50 to 3.50)	1.00
Reader 3	94.55 (91.16 to 96.92)	92.00 (88.14 to 94.92)	2.55 (−0.55 to 5.64)	0.11
Reader 4	94.18 (90.72 to 96.64)	92.73 (88.99 to 95.50)	1.45 (−1.73 to 4.64)	0.37
Specificity				
All reviewers	95.70 (94.44 to 96.74)	96.30 (95.11 to 97.27)	−0.63 (−1.94 to 0.69)	0.42
Reader 1	96.56 (93.93 to 98.27)	98.44 (96.39 to 99.49)	−1.87 (−4.16 to 0.41)	0.11
Reader 2	99.38 (97.76 to 99.92)	98.13 (95.96 to 99.31)	1.25 (−0.48 to 2.98)	0.15
Reader 3	93.08 (89.71 to 95.61)	96.56 (93.93 to 98.27)	−3.46 (−6.26 to −0.66)	0.016
Reader 4	93.75 (90.51 to 96.14)	92.26 (88.79 to 94.93)	1.56 (−1.84 to 4.97)	0.37
Accuracy				
All reviewers	93.90 (92.86 to 94.83)	92.36 (91.22 to 93.40)	0.76 (0.01 to 1.50)	0.049
Reader 1	94.45 (92.30 to 96.15)	89.92 (87.21 to 92.22)	2.27 (0.69 to 3.85)	0.007
Reader 2	93.45 (91.15 to 95.30)	92.76 (90.37 to 94.71)	0.42 (−1.08 to 1.92)	0.66
Reader 3	93.76 (91.50 to 95.57)	94.45 (92.30 to 96.15)	−0.34 (−1.72 to 1.04)	0.72
Reader 4	93.95 (91.72 to 95.73)	92.48 (90.06 to 94.46)	0.76 (−0.73 to 2.24)	0.37

Data in parentheses are 95% CI.
AI, artificial intelligence.

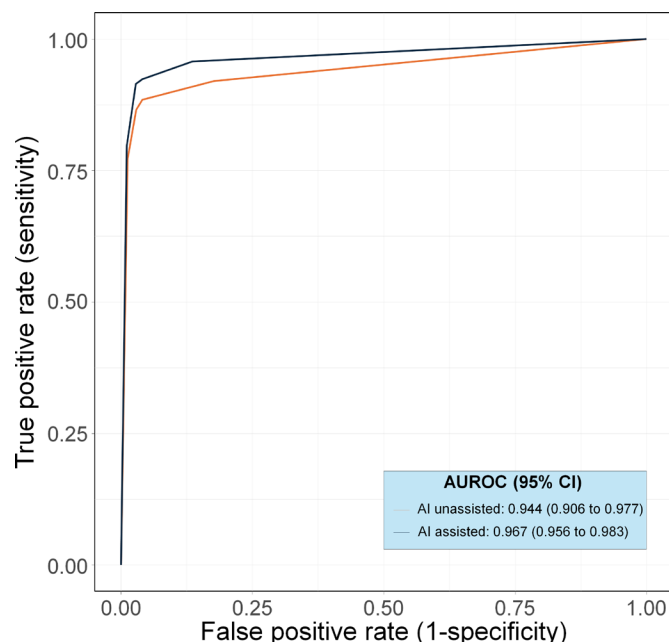


Figure 1 Diagnostic performance of readers without versus with artificial intelligence (AI) assist. Average reader receiver operating characteristic curves for detecting large vessel occlusion under two reading conditions: with and without AI assistance. Average area under the receiver operating characteristic curve (AUROC) was computed across four readers participating in the study using the Obuchowski-Rockette method, which accounts for the multireader multicase study design.

of 87% and a specificity of 97% in detecting intracranial LVO largely comparable or superior to other software packages.

Recent studies have begun to shed light on the potential benefits of EVT for patients with MCA-M2 segment occlusions, broadening the traditional focus from proximal LVO to include more distal vessels. A meta-analysis revealed that patients with MCA-M2 occlusions treated with EVT showed improved functional outcomes at 90 days compared with those receiving standard medical therapy alone.²⁶ Nevertheless, detecting isolated MCA-M2 occlusions remains challenging, even for experienced clinicians. In a study involving 520 patients with ischemic stroke, of which 16% had LVO and 40 patients had isolated MCA-M2 occlusion, experienced neuroradiologists missed 26% of MCA-M2 occlusions at initial CTA evaluation.²⁷ Additionally, the study highlighted that non-neuroradiologists had a fivefold higher risk of missing LVO in a multivariable analysis.²⁷ In our study, the software achieved a sensitivity of 69% in detecting isolated MCA-M2 occlusion, which is higher than that of other automated LVO detection software packages in patients with isolated MCA-M2 occlusion.^{11 13 23–25} This disparity may have resulted from the inclusion of isolated MCA-M2 occlusion as LVO in the training dataset for the software.

The reader assessment study of our research highlights another critical aspect of AI in stroke care: the improvement of diagnostic accuracy in early-career physicians. The observed improvement in sensitivity among early-career physicians when using AI assistance not only validates AI's role as a diagnostic aid but also addresses a significant challenge in stroke care—the scarcity of vascular experts, especially in areas with limited resources. By offering a high degree of sensitivity and specificity in LVO detection, the software can democratize access to high-quality stroke diagnostics, ensuring more patients are correctly

identified for EVT, regardless of their geographical location or the immediate availability of stroke specialists.

Our study is subject to limitations inherent in its retrospective design and the potential for selection bias. Additionally, the real-world efficacy of the software may vary due to differences in imaging equipment, protocols, and patient demographics across healthcare settings. Furthermore, the prevalence of LVO associated with intracranial arterial stenosis is notably higher in Asian populations,²⁸ potentially introducing complexity into our dataset for deep learning analysis. This complexity arises because LVOs associated with intracranial arterial stenosis typically exhibit an increased number of collateral vessels²⁹ and present with less distinct LVO characteristics when contrasted with LVOs resulting from cardioembolic occlusions. These factors underscore the need for prospective, multicenter studies to further validate our findings and explore the integration of the software into diverse clinical workflows. Despite the lack of clinical evidence supporting EVT in basilar artery occlusion, the next version of the software should address EVT-amenable arterial occlusions in the posterior circulation, given the severity of these cases.

In conclusion, we have suggested the clinical efficacy of the software for detecting LVO in this multicenter study in patients with acute ischemic stroke. Furthermore, we demonstrated that AI-assisted reading significantly increases sensitivity of LVO detection in early-career physicians. By facilitating the early identification of patients eligible for EVT, the software has the potential role to improve clinical outcomes of stroke patients.

Contributors J-TK and W-SR had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. JGK, SYH and Y-RK contributed equally to this manuscript. Concept and design: W-SR, DK, J-TK. Acquisition, analysis, or interpretation of data: W-SR, J-TK, JGK, Y-RK. Drafting of the manuscript: W-SR, JGK, SYH, Y-RK, ML, LS, J-TK. Critical review of the manuscript for important intellectual content: W-SR, LS, J-TK, ML. Statistical analysis: SYH, W-SR. Obtained funding: DK. Administrative, technical, or material support: HH, DK, ML. Guarantor: J-TK.

Funding This study was supported by the Multimistry Grant for Medical Device Development (KMDF_PR_20200901_0098), funded by the Korean government and JLK Inc.

Disclaimer The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Competing interests SYH, HH, DK, ML and W-SR are employees of JLK Inc, Seoul, Republic of Korea.

Patient consent for publication Not applicable.

Ethics approval For this diagnostic study, the recruitment, use, analysis, and prospective testing of radiographic images were approved by the institutional review board of Chonnam National University Hospital (CNUH-2023-311) and Daejeon Eulji University Hospital (2023-08-007).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. The data that support the findings of this study are available upon reasonable request and in compliance with local and international ethical guidelines.

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

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is

properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Sue Young Ha <http://orcid.org/0000-0003-2214-1106>

Hotak Hong <http://orcid.org/0009-0009-5159-4914>

Leonard Sunwoo <http://orcid.org/0000-0003-0374-8658>

Wi-Sun Ryu <http://orcid.org/0000-0002-2823-5253>

Joon-Tae Kim <http://orcid.org/0000-0003-4028-8339>

REFERENCES

- Morsi RZ, Elfil M, Ghaith HS, *et al*. Endovascular Thrombectomy for Large Ischemic Strokes: A Living Systematic Review and Meta-Analysis of Randomized Trials. *J Stroke* 2023;25:214–22.
- Jovin TG, Nogueira RG, Investigators D. Thrombectomy 6 to 24 Hours after Stroke. *N Engl J Med* 2018;378:1161–2.
- Albers GW, Marks MP, Kemp S, *et al*. Thrombectomy for Stroke at 6 to 16 Hours with Selection by Perfusion Imaging. *N Engl J Med* 2018;378:708–18.
- Hill MD, Warach S, Rostanski SK. Should Primary Stroke Centers Perform Advanced Imaging? *Stroke* 2022;53:1423–30.
- Nguyen TN, Abdalkader M, Nagel S, *et al*. Noncontrast Computed Tomography vs Computed Tomography Perfusion or Magnetic Resonance Imaging Selection in Late Presentation of Stroke With Large-Vessel Occlusion. *JAMA Neurol* 2022;79:22–31.
- Jadhav AP, Goyal M, Ospel J, *et al*. Thrombectomy With and Without Computed Tomography Perfusion Imaging in the Early Time Window: A Pooled Analysis of Patient-Level Data. *Stroke* 2022;53:1348–53.
- Kang J, Kim SE, Park HK, *et al*. Routing to Endovascular Treatment of Ischemic Stroke in Korea: Recognition of Need for Process Improvement. *J Korean Med Sci* 2020;35:e347.
- Ko Y, Lee S, Chung J-W, *et al*. MRI-based Algorithm for Acute Ischemic Stroke Subtype Classification. *J Stroke* 2014;16:161–72.
- Fasen BACM, Berendsen RCM, Kwee RM. Artificial intelligence software for diagnosing intracranial arterial occlusion in patients with acute ischemic stroke. *Neuroradiology* 2022;64:1579–83.
- Rava RA, Peterson BA, Seymour SE, *et al*. Validation of an artificial intelligence-driven large vessel occlusion detection algorithm for acute ischemic stroke patients. *Neuroradiol J* 2021;34:408–17.
- Matsoukas S, Morey J, Lock G, *et al*. AI software detection of large vessel occlusion stroke on CT angiography: a real-world prospective diagnostic test accuracy study. *J Neurointerv Surg* 2023;15:52–6.
- Rodrigues G, Barreira CM, Bouslama M, *et al*. Automated Large Artery Occlusion Detection in Stroke: A Single-Center Validation Study of an Artificial Intelligence Algorithm. *Cerebrovasc Dis* 2022;51:259–64.
- Yahav-Dovrat A, Saban M, Merhav G, *et al*. Evaluation of Artificial Intelligence-Powered Identification of Large-Vessel Occlusions in a Comprehensive Stroke Center. *AJNR Am J Neuroradiol* 2021;42:247–54.
- Giotta Lucifero A, Baldoncini M, Bruno N, *et al*. Microsurgical Neurovascular Anatomy of the Brain: The Anterior Circulation (Part I). *Acta Biomed* 2021;92:e2021412.
- Brzegowy P, Polak J, Wnuk J, *et al*. Middle cerebral artery anatomical variations and aneurysms: a retrospective study based on computed tomography angiography findings. *Folia Morphol (Warsz)* 2018;77:434–40.
- Han JH, Ha SY, Lee H, *et al*. Automated identification of thrombectomy amenable vessel occlusion on computed tomography angiography using deep learning. *Front Neurol* 2024;15:1442025.
- Yum KS, Chung J-W, Ha S, *et al*. Automated detection of large vessel occlusion: a multicenter study validating efficacy and proving clinical implications. *medRxiv* 05.08.24307038 [Preprint] 2024.
- Szegedy C, Ioffe S, Vanhoucke V, *et al*. Inception-v4, inception-ResNet and the impact of residual connections on learning. *AAAI* 2017;31.
- Tan M, Le Q. Efficientnetv2: smaller models and faster training. *PMLR Proc Mach Learn Res* 2021;139:10096–106.
- Ahn JS, Ebrahimian S, McDermott S, *et al*. Association of Artificial Intelligence-Aided Chest Radiograph Interpretation With Reader Performance and Efficiency. *JAMA Netw Open* 2022;5:e2229289.
- DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837–45.
- Smith B, Hillis S, Pesce L. *MRMCaov: multi-reader multi-case analysis of variance. R package version 0.1*. 2021:16.
- Amukotuwa SA, Straka M, Dehkharghani S, *et al*. Fast Automatic Detection of Large Vessel Occlusions on CT Angiography. *Stroke* 2019;50:3431–8.
- Delora A, Hadjialiakbari C, Perenti E, *et al*. Viz LVO versus rapid LVO in detection of large vessel occlusion on CT angiography for acute stroke. *J Neurointerv Surg* 2024;16:599–602.
- Schlossman J, Ro D, Salehi S, *et al*. Head-to-head comparison of commercial artificial intelligence solutions for detection of large vessel occlusion at a comprehensive stroke center. *Front Neurol* 2022;13:1026609.
- Sarraj A, Hassan AE, Savitz S, *et al*. Outcomes of Endovascular Thrombectomy vs Medical Management Alone in Patients With Large Ischemic Cores: A Secondary Analysis of the Optimizing Patient's Selection for Endovascular Treatment in Acute Ischemic Stroke (SELECT) Study. *JAMA Neurol* 2019;76:1147–56.
- Fasen BACM, Heijboer RJJ, Hulsmans F-JH, *et al*. CT Angiography in Evaluating Large-Vessel Occlusion in Acute Anterior Circulation Ischemic Stroke: Factors Associated with Diagnostic Error in Clinical Practice. *AJNR Am J Neuroradiol* 2020;41:607–11.
- Tong X, Wang Y, Fiehler J, *et al*. Thrombectomy Versus Combined Thrombolysis and Thrombectomy in Patients With Acute Stroke: A Matched-Control Study. *Stroke* 2021;52:1589–600.
- Guglielmi V, LeCouffe NE, Zinkstok SM, *et al*. Collateral Circulation and Outcome in Atherosclerotic Versus Cardioembolic Cerebral Large Vessel Occlusion. *Stroke* 2019;50:3360–8.