

Artificial intelligence assisted detection of large vessel occlusion on CT angiography in acute stroke patients: a multi-reader multi-case study

Kiruba Nagaratnam, MD^{1,*}, Phil Mathieson, MBChB, MBA², Anna Podlasek, MD, PhD^{3,4}, Peter Slade, MBBCh⁵, Gary A. Ford, MBBChir^{2,6,*}, Anthony Cox, MBBS⁷, James H. Briggs, MBChB^{1,8}, Zoe V. J. Woodhead, PhD⁸, George Harston, DPhil^{2,8}

¹Royal Berkshire NHS Foundation Trust, Reading, RG1 5AN, United Kingdom

²Oxford University Hospitals NHS Trust, Stroke Medicine, Oxford, OX3 9DU, United Kingdom

³Division of Radiological and Imaging Sciences, University of Nottingham, Nottingham, NG7 2UH, United Kingdom

⁴Taysie Innovation MedTech Ecosystem (TIME), University of Dundee, Dundee, DD2 1FD, United Kingdom

⁵Stroke Medicine Morriston, Morriston Hospital, Swansea, SA6 6NL, United Kingdom

⁶Radcliffe Department of Medicine, University of Oxford, Oxford, OX3 9DU, United Kingdom

⁷North Bristol NHS Trust, BS10 5NB, United Kingdom

⁸Brainomix Limited, Oxford, OX2 7HN, United Kingdom

*Corresponding authors: Kiruba Nagaratnam, MD, Acute Stroke Unit, Battle Block, Level 1, Royal Berkshire Hospital, London Road, Reading, RG1 5AN, United Kingdom (kiruba.nagaratnam@nhs.net); Gary A. Ford, MBBChir, Radcliffe Department of Medicine, University of Oxford, Level 6, West Wing, John Radcliffe Hospital, Headington, Oxford, OX3 9DU, United Kingdom (gary.ford@rdm.ox.ac.uk)

Abstract

Objectives: We assessed the impact of artificial intelligence (AI) software (e-CTA, Brainomix) on clinical decision-making in patients with suspected acute ischemic stroke.

Methods: A retrospective, multi-reader-multi-case crossover design compared readers' performance with vs without software support. Twenty cases were included, 10 with large vessel occlusion (LVO) and 10 without LVO. Twenty-one NHS clinicians, representing intended software users ranging in experience, conducted 2 sessions (washout period > 2 weeks). In session 1, software support was provided for 10 randomly selected cases. In session 2, support allocation was reversed. Outcome measures included LVO detection, collateral scoring, diagnosis, treatment decision, time taken and confidence.

Results: Sensitivity, specificity, and accuracy of LVO detection improved with imaging software for LVO detection, with increased confidence and reduced time taken. There was no significant difference in collateral scoring or diagnoses.

Conclusion: e-CTA can improve performance of NHS clinicians when interpreting acute stroke imaging.

Advances in knowledge: This paper provides new evidence that AI decision support software has the capacity to improve the performance of representative users in the NHS when interpreting imaging to identify patients for acute stroke treatments.

Keywords: artificial intelligence; AI; acute ischemic stroke; CT angiography; CTA; large vessel occlusion; LVO; collateral score.

Introduction

In current clinical practice, cervical and intracranial CT angiography (CTA) is the primary method for detecting large vessel occlusion (LVO) in patients with suspected stroke. CTA determines mechanical thrombectomy (MT) eligibility by identifying LVO in the anterior and posterior circulation.¹⁻³ Collateral blood flow is an important predictor of outcome from LVO stroke and can be used to approximate the salvageable brain tissue and gauge the suitability for transport between hospitals; hence, collateral status derived from CTA is an important component of the final decision on whether to proceed with MT for the particular patient.^{1,3-5} The speed of treatment delivery is critically important to maximize the chance of good patient outcomes⁶⁻⁸; therefore, there is a need for real-time, fast and reliable imaging interpretation to identify and localize the occlusion and guide treatment decisions.

However, access to neuroimaging expertise is not readily available in all hospitals and even in specialist centres may not be available out of hours.⁹ Artificial intelligence (AI) decision support tools are increasingly used to support clinicians in the timely identification of patients eligible for MT.¹⁰⁻¹² AI decision support software is recommended to be used in routine clinical practice by both UK (Stroke—Getting It Right First Time—GIRFT) and US guidelines.¹³

Regulatory and clinical evidence for AI imaging software is typically based on standalone performance testing against expert consensus. A growing number of papers have examined the standalone performance of software devices for automated detection of proximal anterior circulation LVOs¹⁴⁻²⁰ (typically defined as intracranial internal carotid artery [ICA] or middle cerebral artery [MCA] M1 occlusions), including some with very large datasets²¹ or with challenging, real-world

datasets.^{22,23} Performance of automated tools for assessment of collaterals has also been explored.²⁴⁻²⁶ In practice, however, these software devices are designed to generate a triage notification or to be used as a decision-support tool. While stand-alone performance is important, so too is the impact that the software outputs have on clinical decision making. There is much more limited literature evaluating the impact of such software impact on clinicians' imaging interpretation,²⁷ or on clinical outcomes.²⁸⁻³¹

This multiple reader multiple case (MRMC) study aims to measure the impact of decision support software for CTA ("Brainomix 360 e-CTA," Brainomix Limited) on clinical decision-making in the context of patients presenting to acute hospitals with a clinically suspected stroke in the UK. A range of outcome measures were assessed (reflecting the different features of the software), relating to LVO detection, collateral scoring, diagnosis, treatment decisions, time taken, and confidence. The modest sample size (20 cases and 21 readers) was justified pragmatically, positioning the study as a relatively small discovery study with exploratory hypothesis testing. Stroke clinicians with a range of experience from different UK institutions were included to provide a representative sample of the software users across the NHS. The study's overall objective was to assess whether AI imaging improves accuracy, speed and confidence in CTA image interpretation in NHS stroke clinicians.

Methods

Procedure

This clinical performance assessment utilized a MRMC approach to evaluate the performance of a set of clinical readers when interpreting CTA imaging for a selection of patients with LVO of the MCA or intracranial ICA ($N=10$) or without LVO ($N=10$). All cases were randomly selected from a historical research registry of cases with suspected stroke undergoing CTA.²⁵

A viewing platform was created to view CTA images alongside representative clinical vignettes for each case and to collect readers' responses. When images were presented without decision support ("without e-CTA"), only the unannotated CTA images (and vignette) were presented. When images were presented with decision support from e-CTA ("with e-CTA"), the annotated output from the automated e-CTA analysis was also presented in the form of an axial maximum intensity project and details of the estimated CTA collateral score (CS), vessel density and occlusion detection. [Figure 1](#) below shows screenshot examples of the viewing platform *with* (top) and *without* (bottom) decision support from e-CTA.

Readers viewed the 20 images in 2 sessions, spaced by at least 2 weeks. [Figure 2](#) shows a schematic depiction of the study procedure. At the first session, half of the cases were randomly selected to be presented with e-CTA and the remaining half were presented without e-CTA. The selection of cases randomized separately for each reader to avoid bias. At the second session, the allocation of decision support was reversed so that by the end of the second session, all cases had been reviewed once with and once without decision support from e-CTA. The readers' ability and confidence to detect key diagnostic features (LVO and CS), and to arrive at a diagnosis and treatment decision was recorded. A radiological ground truth was set by an expert neuroradiologist with access to follow-up information. Questions that readers were asked are outlined in [Table 1](#).

Case selection

Twenty cases were randomly selected from a historical research registry of cases with suspected stroke undergoing CTA.²⁵ A balanced mix of cases with and without LVO (10 of each) were selected.

All cases met the following criteria: (1) age 18 years or over; (2) availability of baseline CTA with adequate quality (eg, free from excessive motion artefact); (3) availability of demographic information and follow-up imaging; and (4) no evidence of intracranial haemorrhage.

Selection of LVO cases was enriched to ensure that occlusion locations matched the distribution in the HERMES meta-analysis that informed the guidelines for patient selection for MT.³² LVO was defined as an occlusion of either the intracranial ICA or proximal segment of the MCA. Patients with distal occlusions or posterior circulation stroke were not included, as these are beyond the scope of the e-CTA software.

The CT images were acquired from a range of scanners with different protocols, representative of the range of data acquired in clinical practice. An experienced stroke physician (G.H.), independent of the test reader group, generated representative clinical vignettes. All image headers and clinical vignettes were anonymized.

Readers

Twelve readers completed all 20 cases at both sessions; a further 9 completed only the first session. Data from all 21 were included in the final analysis.

The readers were representative UK clinicians, specialising in either stroke, geriatric medicine, or radiology, at a specialist trainee (ST) or consultant grade (see [Table 2](#)).

Readers were familiarized with the purpose of the study and received training in the use of e-CTA and the viewing portal prior to beginning the study. The training consisted of written materials and a webinar video describing the purpose and function of the software, as well as the CS grading system. The design of the study was described, including instruction that occlusions (if present) would be restricted to the anterior circulation.

Ground truth

Ground truth for presence or absence of LVO and diagnosis was set by an expert neuroradiologist with access to follow up clinical information, final diagnosis, and follow up imaging. Due to the subjective nature of collateral scoring, a consensus ground truth was taken for CSs between 3 expert clinicians: 2 neuroradiologists, and 1 stroke physician.

Automated CTA analysis

Images on the reader platform were processed with Brainomix 360 e-CTA, CE-marked decision support software. The software provides analysis and viewing capabilities of CTA datasets in stroke patients, including assessment of collateral status, detection of intracranial ICA or MCA LVOs, and visualization of intracranial blood vessels. Image analysis involves the following: (1) image preprocessing, including brain extraction and normalization (registration); (2) blood vessel segmentation and generation of a maximum intensity projection (MIP) image or multiplanar reconstructions (MPRs); (3) detection of occlusions in the patent MCA large vessels through use of a convolutional neural network (CNN); (4) generation of a probabilistic heatmap of reduced vessel density through comparison of affected and unaffected

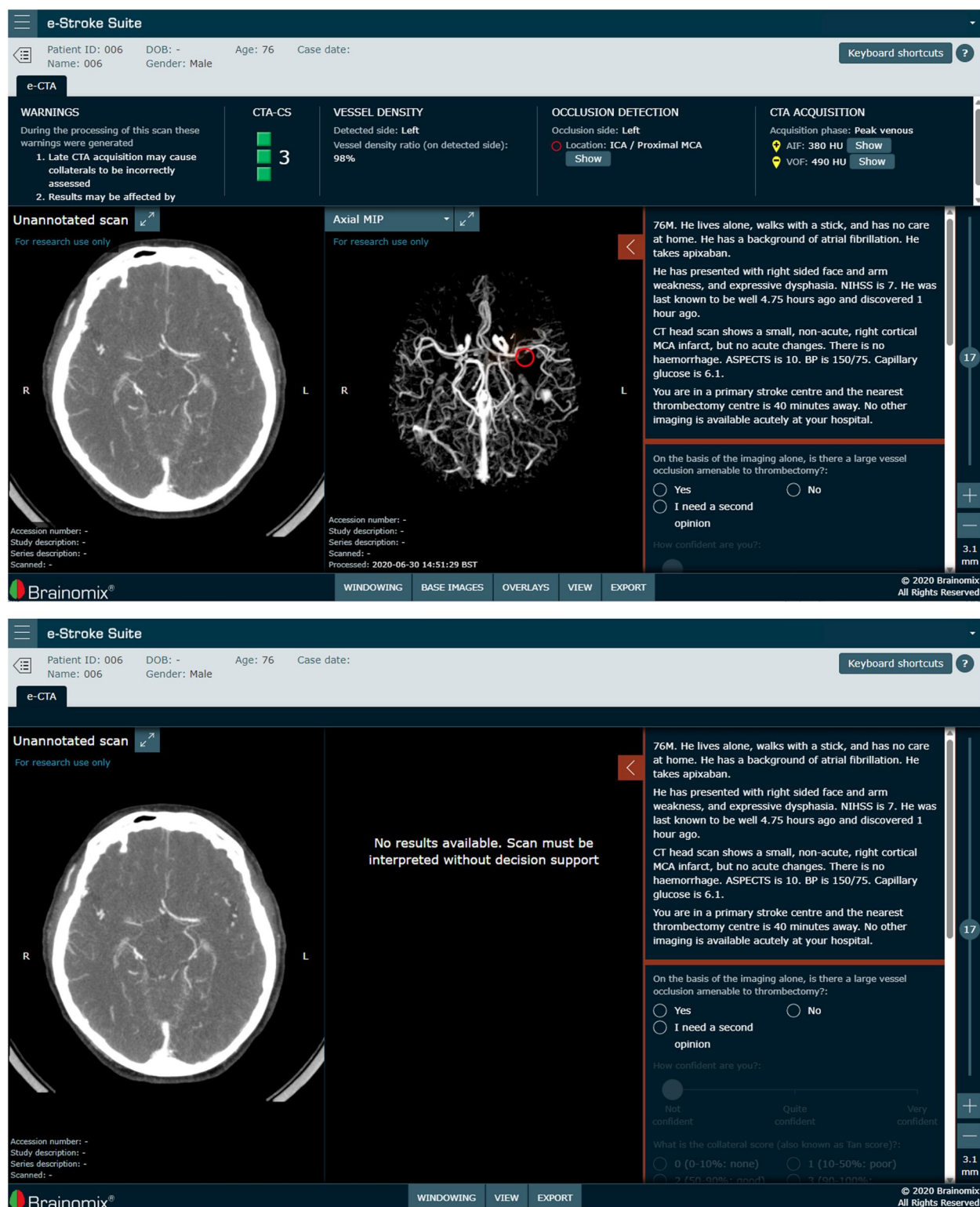


Figure 1. Screenshot examples of the study viewing platform with (top) and without (bottom) decision support from e-CTA.

hemispheres; and (5) quantification of a CTA CS³³ (CTA-CS) on the basis of the vessel density map.

On the reader platform, the outputs of e-CTA were presented, including the preprocessed CTA image and the MIP, overlaid with the vessel density heatmap, the CTA-CS and the location of any LVO detected by the software (see Figure 1). Information on the acquisition phase (as detected

by the software) is also shown to facilitate user interpretation of image quality.

Statistical analysis

A number of outcome measures were collected for each reader and analysed as described in Table 3. Each analysis

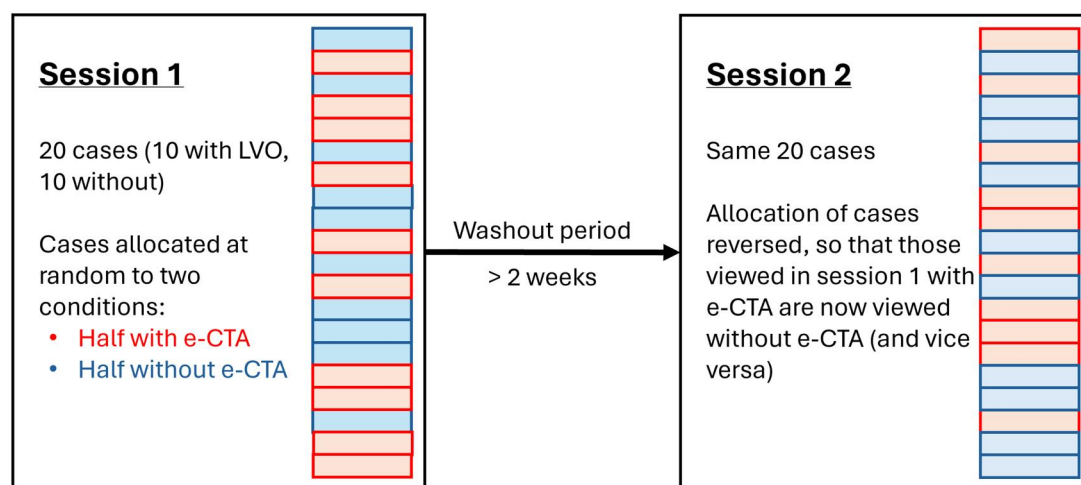


Figure 2. Schematic depiction of the study procedure, showing the 2 sessions, spaced by at least 2 weeks, each containing 20 cases.

Table 1. Questions and response options for readers.

Question	Response options
On the basis of the imaging alone, is there a large vessel occlusion amenable to thrombectomy?	Yes No I need a second opinion
How confident are you regarding LVO?	Sliding scale from 0 (Not confident) to 100 (Very confident)
What is the collateral score (Tan score)?	0 (0%-10%; none) 1 (10%-50%; poor) 2 (50%-90%; good) 3 (90%-100%; complete)
How confident are you about the collateral score?	Sliding scale from 0 to 100
On the basis of the imaging and the vignette, what is the most likely diagnosis?	Stroke Seizure Migraine Other
How confident are you about the diagnosis?	Sliding scale from 0 to 100
What treatment would you offer (in the absence of other contraindications)? (select one or two answers)	IV Thrombolysis IV Thrombolysis; only after additional imaging Mechanical thrombectomy Mechanical thrombectomy; only after additional imaging None

compared the impact of e-CTA support on the outcome measure.

The outcome measures for LVO detection (sensitivity, specificity and area under the receiver operating characteristic curve [AUC]) were analysed using the unified Obuchowski and Rockette/Dorman, Berbaum and Metz (OR-DBM) MRMC methods devised by Hillis et al.³⁴ as implemented in the MCMCao package in R.³⁵ This analysis computed the overall area AUC, sensitivity and specificity for LVO detection at the group level.

For the remaining measures, percentages (for accuracy) or averages (for confidence and time taken) were calculated for each reader in the with and without e-CTA conditions. The results for each condition were then compared using Mann-Whitney U tests (a non-parametric equivalent of an independent samples t-test). A non-parametric test was chosen as the outcome measures were either bounded variables (eg, percentages), or because data exploration showed that they were not normally distributed. An independent samples test was chosen as not all cases were reviewed twice (once with and once without) by each reader.

In addition to the outcome measures listed above, the agreement between readers and ground truth (with and without e-CTA) was assessed using Cohen's Kappa.

An alpha criterion of $P < .05$ was used to indicate statistical significance.

LVO detection sensitivity analysis

For LVO detection, readers were given the options “Yes,” “No” or “I need a second opinion.” If the reader selected the latter option, the response was coded as an incorrect response; that is, either a false negative (for positive cases) or a false positive (for negative cases). In order to evaluate the impact of this coding scheme, a sensitivity analysis was performed whereby the LVO analyses (for LVO AUC, sensitivity, specificity, accuracy, confidence, and Cohen's kappa) were repeated, but excluding any cases where a response of “I need a second opinion” was given.

Treatment decision

The frequencies of readers' decisions to recommend treatment with IV thrombolysis and/or MT (either with or without further imaging) were evaluated.

Results

Table 4 and Figure 3 show the total number of cases reviewed with or without e-CTA, the mean of each outcome measure

with and without e-CTA. A ROC curve for the LVO detection measure is shown in Figure 4.

The results demonstrate that all outcome measures were numerically higher (ie, better performance) when the readers

were given e-CTA decision support, than when they were not. Using the pre-specified alpha criterion of $p < .05$, a statistically significant difference was observed for LVO AUC, LVO sensitivity, LVO specificity, LVO accuracy, LVO confidence,

Table 2. Clinical role and specialty of the study readers (where available), and number of cases read with and without e-CTA decision support.

Reader	Cases With e-CTA	Without e-CTA	Role	Specialty
Reader01	20	20	ST	Radiology
Reader02	20	20	Consultant	Stroke
Reader03	20	20	Consultant	Stroke
Reader04	20	20	Consultant	Stroke
Reader05	20	20	Consultant	Geriatric medicine
Reader06	20	20	Consultant	Stroke
Reader07	20	20	Consultant	Geriatric medicine
Reader08	20	20	Consultant	Stroke
Reader09	20	20	Consultant	Stroke
Reader10	20	20	Consultant	Geriatric medicine
Reader11	20	20	Consultant	Stroke
Reader12	20	20	Consultant	Stroke
Reader13	8	12	Consultant	Stroke
Reader14	7	13	Consultant	Stroke
Reader15	11	9	ST	Stroke
Reader16	10	10	Consultant	Radiology
Reader17	8	12	ST	Stroke
Reader18	9	11	Consultant	Stroke
Reader19	11	9	ST	Stroke
Reader20	8	12	Consultant	Stroke
Reader21	10	10	Consultant	Stroke

ST = specialist trainee.

Table 3. Outcome measures and analysis methods.

Measure	Description	Analysis
LVO AUC	Area under the ROC curve for LVO detection	OR-DBM analysis
LVO sensitivity	Sensitivity of LVO detection (TP/(TP+FN))	OR-DBM analysis
LVO specificity	Specificity of LVO detection (TN/(TN+FP))	OR-DBM analysis
LVO accuracy	Percentage accuracy of LVO detection	Mann Whitney U
LVO confidence	Mean confidence in LVO detection	Mann Whitney U
CS accuracy	Percentage accuracy of collateral scores	Mann Whitney U
CS confidence	Mean confidence in collateral scoring	Mann Whitney U
Diagnosis accuracy	Percentage accuracy in diagnosis	Mann Whitney U
Diagnosis confidence	Mean confidence in diagnosis	Mann Whitney U
Treatment confidence	Mean confidence in treatment decision	Mann Whitney U
Time taken	Time taken to complete all scores per case	Mann Whitney U

LVO = large vessel occlusion; CS = collateral score; TP = true positive; FN = false negative; TN = true negative; FP = false positive; OR-DBM = Obuchowski-Rockette/Dormann Berbaum and Metz.

Table 4. Summary statistics for each outcome measure are reported for cases reviewed with and without e-CTA decision support.

Measure	With e-CTA	Without e-CTA	<i>p</i>
N Cases	322	338	NA
LVO TP/FP/TN/FN	149/15/146/12	135/41/128/34	NA
LVO AUC	0.92 [CI: 0.86, 0.98]	0.79 [CI: 0.69, 0.88]	$t = 4.35, p < .001$
LVO sensitivity	0.94 [CI: 0.88, 1.00]	0.82 [CI: 0.68, 0.96]	$t = 2.95, p < .001$
LVO specificity	0.91 [CI: 0.81, 1.00]	0.75 [CI: 0.63, 0.88]	$t = 2.75, p = .012$
LVO accuracy	0.95 [IQR: 0.90, 1.00]	0.85 [IQR: 0.65, 0.90]	$U = 350, p < .001$
LVO confidence	0.75 [IQR: 0.70, 0.86]	0.62 [IQR: 0.54, 0.70]	$U = 327, p = .007$
CS accuracy (%)	0.60 [IQR: 0.50, 0.70]	0.50 [IQR: 0.40, 0.60]	$U = 282, p = .124$
CS confidence	0.59 [IQR: 0.50, 0.68]	0.47 [IQR: 0.34, 0.55]	$U = 303, p = .038$
Diagnosis accuracy (%)	0.88 [IQR: 0.85, 0.90]	0.85 [IQR: 0.75, 0.90]	$U = 262, p = .303$
Diagnosis confidence	0.65 [IQR: 0.48, 0.75]	0.58 [IQR: 0.47, 0.63]	$U = 266, p = .261$
Treatment confidence	0.57 [IQR: 0.40, 0.67]	0.48 [IQR: 0.42, 0.57]	$U = 35, p < .001$
Time taken (s)	125 [IQR: 106, 146]	135 [IQR: 113, 164]	$U = 441, p < .001$

For LVO AUC, sensitivity and specificity, the results are the estimated group means, shown with 95% CI in brackets. For all other outcome measures, medians are reported with IQR in brackets. Results with and without e-CTA are compared using either t-tests or Mann Whitney U tests as appropriate. LVO = large vessel occlusion; CS = collateral score; TP = true positive; FN = false negative; TN = true negative; FP = false positive.

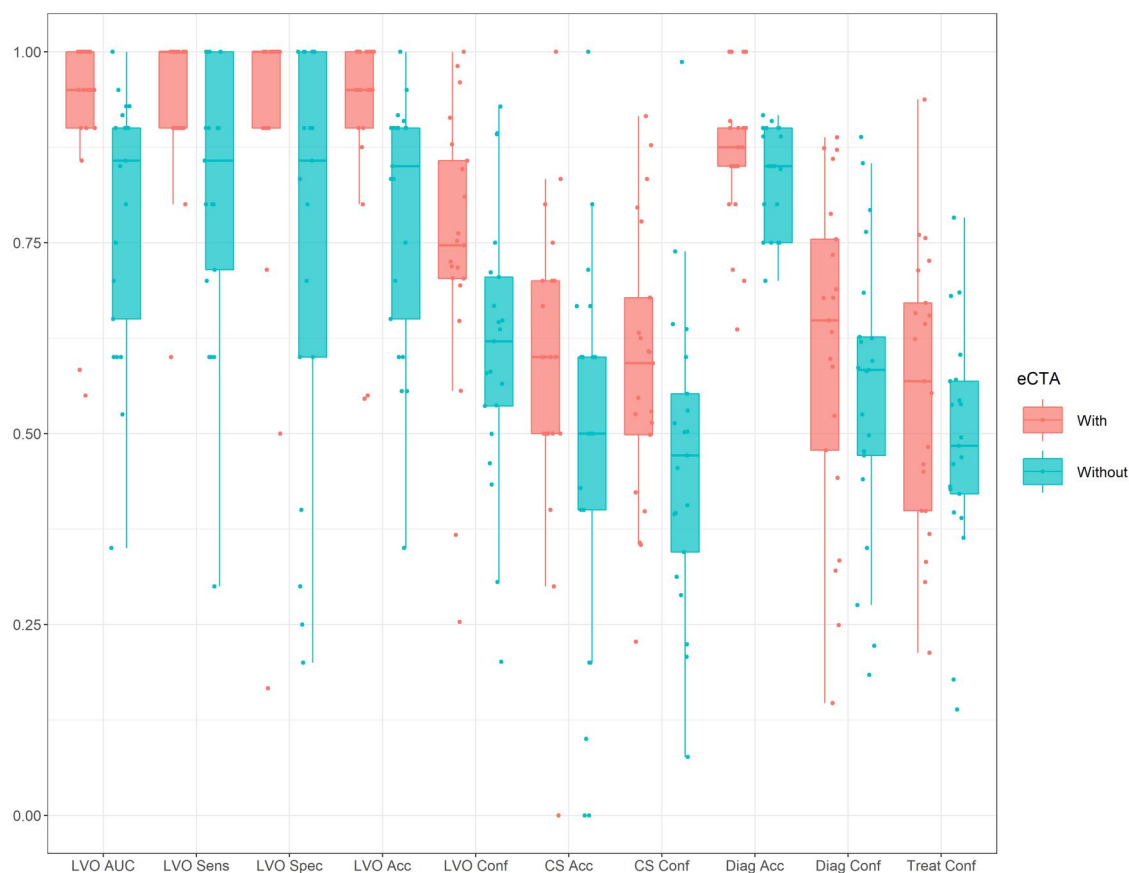


Figure 3. Boxplot showing group level summary statistics (median and interquartile range) for each outcome measure (LVO AUC, LVO sensitivity, LVO specificity, LVO accuracy, LVO confidence, CS accuracy, Collateral Score confidence, Diagnosis accuracy, Diagnosis confidence, and Treatment confidence; for a description of each outcome measure, see Table 3), when readers viewed cases with (red) or Without (blue) e-CTA decision support. Outcome measures for each reader are shown in red and blue points.

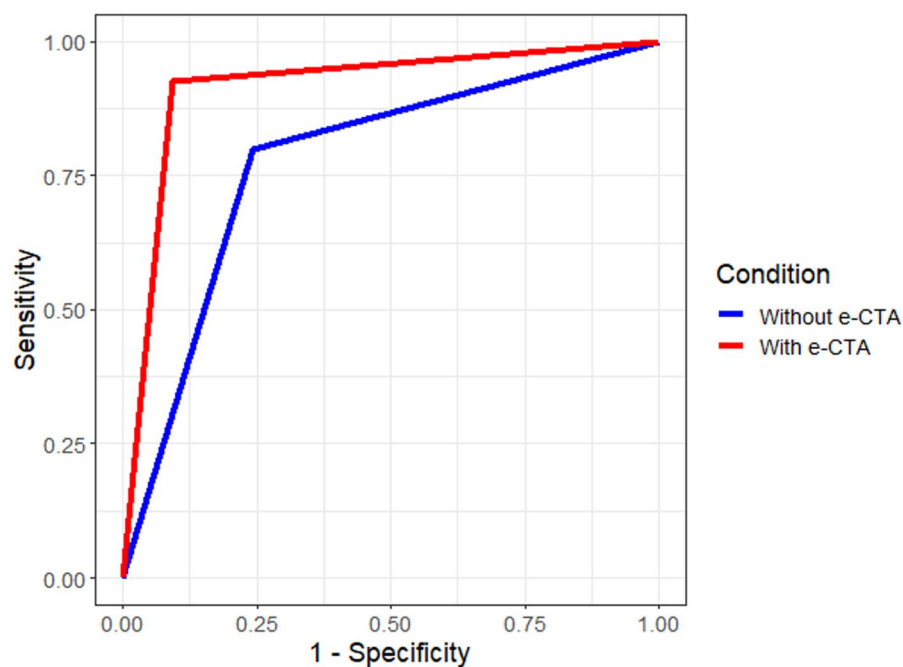


Figure 4. ROC curves for LVO detection of readers in the with e-CTA (red) and without e-CTA (blue) conditions.

Table 5. Alternative results for LVO detection sensitivity and specificity, where cases with the response of “I need a second opinion” were omitted from the analysis.

Measure	With e-CTA	Without e-CTA	<i>p</i>
N cases included	304	296	NA
N cases omitted	18	42	NA
LVO TP/FP/TN/FN	149/6/146/3	135/13/128/20	NA
LVO AUC	0.96 [CI: 0.93, 1.00]	0.88 [CI: 0.79, 0.97]	$t = 2.21, p = .03$
LVO sensitivity	0.98 [CI: 0.96, 1.00]	0.88 [CI: 0.76, 0.99]	$t = 2.22, p = .03$
LVO specificity	0.94 [CI: 0.87, 1.00]	0.88 [CI: 0.75, 1.00]	$t = 1.01, p = .32$
LVO accuracy	1.00 [IQR: 0.95, 1.00]	0.90 [IQR: 0.85, 0.95]	$U = 346, p < .001$
LVO confidence	0.76 [IQR: 0.70, 0.88]	0.65 [IQR: 0.54, 0.71]	$U = 328, p = .006$

TP = true positive; FN = false negative; TN = true negative; FP = false positive.

CS confidence, treatment confidence and time taken. Confidence levels for LVO detection, collateral scoring and treatment decisions all improved by around 10% when e-CTA support was available. The time taken to review the cases also showed a small but significant improvement of ~10 s per case. The results for CS accuracy, diagnosis accuracy and diagnosis confidence were not statistically significant.

For LVO detection *with* e-CTA, Cohen's Kappa was 0.83 (95% confidence interval [CI]: 0.77, 0.89), indicating very good agreement between readers and ground truth. For LVO detection *without* e-CTA, Cohen's Kappa was 0.56 (95% CI: 0.47, 0.64), indicating moderate agreement. Overall accuracy improved by around 10% with use of e-CTA although this varied between readers (range: -10% to 44%).

LVO detection sensitivity analysis

A sensitivity analysis was conducted to evaluate the effect of the way LVO detection responses were coded when readers selected “I need a second opinion.” This response was selected 18 times when e-CTA support was available (6% of responses), and 42 times when e-CTA support was not available (12% of responses). In the main analysis, “I need a second opinion” responses were treated as incorrect responses.

A second analysis was conducted where the “I need a second opinion” cases were excluded from the analysis. This changed the results as shown in Table 5. As shown in the table, excluding these cases had the effect of reducing the number of false positive (FP) and false negative (FN) counts.

The results of the second analysis were broadly similar to the main analysis. As in the main analysis, the difference between with and without e-CTA conditions were significant for LVO AUC, sensitivity, accuracy and confidence; however, the difference was not significant for LVO specificity in this analysis. For the second analysis, Cohen's Kappa was 0.94 (95% CI: 0.9, 0.98) with e-CTA, indicating very good agreement. Without e-CTA, Cohen's Kappa was 0.78 (95% CI: 0.71, 0.85), indicating good agreement.

Treatment decision

The readers' decisions to recommend treatment with IV thrombolysis and/or MT (either with or without further imaging, in the context of the imaging and clinical information provided in the vignette) was also examined (Table 6). This showed that treatment recommendations were higher when e-CTA was available than when it was not.

Table 6. Rates of treatment recommendations (in percentage) for ground truth data, and for cases viewed with e-CTA vs without e-CTA.

Condition	IVT	IVT after imaging	MT	MT after imaging
With e-CTA	38%	21%	26%	33%
Without e-CTA	36%	17%	20%	33%

IVT = intravenous thrombolysis; IVT after imaging = intravenous thrombolysis only after further imaging; MT = mechanical thrombectomy; MT after imaging = mechanical thrombectomy only after further imaging.

Discussion

This study demonstrates a significant improvement in performance of readers in the detection of LVO on CTA with AI decision support software compared to without. As well as showing an improvement in AUC, sensitivity, specificity, and accuracy, the decision support software also resulted in a higher degree of confidence in reader findings.

This increase in reader confidence is also reflected in the reduction in the proportion of cases where readers required a second opinion to determine whether LVO was present, occurring twice as often when e-CTA support was not available. Moreover, readers were more likely to recommend treatment, both in the form of IV thrombolysis and MT, when they reviewed the scans with decision support than without.

Use of e-CTA also resulted in improvements in readers' confidence in collateral scoring. An improvement in collateral scoring accuracy was observed, but this was non-significant likely due to the modest sample size (10 cases with LVO) and distribution of CSs. Readers were significantly faster to assess cases when e-CTA support was available—the time taken measure encompassed the total time to assess the presence or absence of LVOs, the CS, the diagnosis and treatment decision. Diagnosis accuracy did not change significantly, but performance on this measure was high in both with and without conditions (88% and 85%, respectively).

This study included a range of readers from different clinical specialty backgrounds, and of varying levels of experience, reflecting the different models delivering acute services for stroke patients at different sites, and at different times, across the NHS. These results suggest that in addition to improving individual readers' performance, decision support software reduces the range of performance across the sample user group. This may have particular benefit in reducing the variability of care between different hospitals, and at different times.

As well as assessing the performance of users in a research setting, this study also investigated the relationship between

decision support software and reader confidence and the time taken for clinical assessment and decision making. These are important outcomes for any healthcare technology when considering any real-world impact, as clinician confidence is an important element in determining the speed at which treatment can be delivered. This is especially important in care of acute stroke patients, in whom a reduced time to treatment has been shown to improve outcomes. Treatment with MT is highly efficacious, and research has shown that treating more acute stroke patients with MT will result in better long term outcomes (ie, fewer patients with serious disability) and lower long term healthcare costs to the economy.³⁶

It is important to note that in order for AI software to be integrated into routine clinical practice, users need to be trained in the correct use of the software, and to be made aware of the opportunities and limitations of its use. Reader studies such as this one can help build the understanding of both intended use and impact. Ultimately, the responsibility for clinical decisions lies with the clinician using the software. Nevertheless, it is important that the software has been appropriately evaluated and any risks identified before deployment that might influence patient care, hence the need for a robust regulatory framework to guide the use of AI decision support software.

This study is consistent with and complementary to the growing body of research demonstrating the benefits of AI-driven decision support software for interpretation of acute stroke imaging, both in terms of MRM reader (with-or-without) studies,^{27,37-39} and studies evaluating the impact of AI software on clinical outcomes.²⁸⁻³¹ Future studies will evaluate the impact of e-CTA in real-world prospective patient cohorts. Once such example has now been published, assessing the LVO detection performance of e-CTA in a prospective study.⁴⁰

Limitations of this study include the relatively small number of scans assessed, and incomplete assessment by some readers. The study was not designed or powered to allow evaluation of the standalone performance of the software (ie, the accuracy of the outputs of the software relative to ground truth), but details of this can be found in 2 recent studies.^{16,17} The proportion of cases with LVO is higher than typically seen in an unselected population of patients presenting with acute stroke symptoms where 20%-25% patients have LVO. The LVO cases were limited to MCA or ICA occlusions only, and did not include cases with distal or posterior circulation occlusions. Together with the synthetic vignettes and the low numbers, the population does not reflect real-world prevalence of the conditions, but was designed to balance statistical power against the pragmatic consideration of time required for clinicians to participate in the study.

Conclusions

AI decision support software has the capacity to improve the performance of representative users in the NHS when interpreting imaging to identify patients for acute stroke treatments. When using the software, clinicians were more likely to make recommend treatment with either systemic thrombolysis or MT. Prospective studies are ongoing to capture the impact of e-CTA technology in a real-world setting, without the bias inherent in a retrospective, enriched dataset. Future algorithm development could focus on broadening the

applicability of the software to distal or posterior circulation strokes.

The results of this real-world study demonstrate the potential for AI based decision support software to improve accuracy and confidence of UK clinicians in detection of LVO, as well as reducing variability of performance between individuals and increasing clinician confidence. These outcomes could help deliver more consistent, expeditious, and accurate selection of patient for MT, with the eventual aim of improving long term clinical outcomes for patients.

Author contributions

Conceptualization: G.H. and K.N.; Data curation and formal analysis: Z.V.J.W.; Investigation, Writing (Review & Editing): all authors; Writing (Original Draft Preparation): G.H. and K.N. G.F. and Z.V.J.W. are considered as co-corresponding authors of this work.

Funding

None declared.

Conflicts of interest

J.H.B., Z.V.J.W., and G.H. were employed by Brainomix Limited. K.N. and P.M. have received consultancy fees from Brainomix Limited. G.F. reports receiving consulting fees from AstraZeneca for management of stroke due to intracerebral haemorrhage (payment to his employer), Bayer for lecture on models of NHS industry working, CSL Behring for stroke trial consultancy, and being Chief Executive of Health Innovation Oxford and Thames Valley which has multiple joint working agreements and medical education grants with industry partners that are contracts with Oxford University Hospitals NHS Trust the host organisation for Health Innovation Oxford and Thames Valley. 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.

References

1. Nice. Stroke and transient ischaemic attack in over 16s: diagnosis and initial management (NG128). 2019. Accessed February 12, 2025. www.nice.org.uk/guidance/ng128
2. Turc G, Bhogal P, Fischer U, et al. European Stroke Organisation (ESO) - European Society for Minimally Invasive Neurological Therapy (ESMINT) guidelines on mechanical thrombectomy in acute ischaemic stroke endorsed by Stroke Alliance for Europe (SAFE). *Eur Stroke J*. 2019;4:6-12. <https://doi.org/10.1177/2396987319832140>
3. Powers WJ, Rabinstein AA, Ackerson T, et al. Guidelines for the early management of patients with acute ischemic stroke: 2019 update to the 2018 guidelines for the early management of acute ischemic stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke*. 2019;50:e344-e418. <https://doi.org/10.1161/STR.0000000000000211>
4. Tan BY, Wan-Yee K, Paliwal P, et al. Good intracranial collaterals trump poor ASPECTS (Alberta Stroke Program Early CT Score) for intravenous thrombolysis in anterior circulation acute ischemic stroke. *Stroke*. 2016;47:2292-2298. <https://doi.org/10.1161/STROKEAHA.116.013879>

5. Goyal M, Demchuk AM, Menon BK, et al. Randomized assessment of rapid endovascular treatment of ischemic stroke. *N Engl J Med*. 2015;372:1019-1030. <https://doi.org/10.1056/NEJMoa1414905>
6. Saver JL, Goyal M, van der Lugt A, et al. Time to treatment with endovascular thrombectomy and outcomes from ischemic stroke: a meta-analysis. *JAMA*. 2016;316:1279-1288. <https://doi.org/10.1001/jama.2016.13647>
7. Khatri P, Yeatts SD, Mazighi M, et al. Time to angiographic reperfusion and clinical outcome after acute ischaemic stroke: an analysis of data from the Interventional Management of Stroke (IMS III) phase 3 trial. *Lancet Neurol*. 2014;13:567-574. [https://doi.org/10.1016/S1474-4422\(14\)70066-3](https://doi.org/10.1016/S1474-4422(14)70066-3)
8. Ota T, Nishiyama Y, Koizumi S, Saito T, Ueda M, Saito N. Impact of onset-to-groin puncture time within three hours on functional outcomes in mechanical thrombectomy for acute large-vessel occlusion. *Interv Neuroradiol*. 2018;24:162-167. <https://doi.org/10.1177/1591019917747247>
9. Fasen B, Heijboer RJJ, Hulsmans FH, Kwee RM. 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;41:607-611. <https://doi.org/10.3174/ajnr.A6469>
10. Mokli Y, Pfaff J, Dos Santos DP, Herweh C, Nagel S. Computer-aided imaging analysis in acute ischemic stroke—background and clinical applications. *Neurol Res Pract*. 2019;1:23. <https://doi.org/10.1186/s42466-019-0028-y>
11. Murray NM, Unberath M, Hager GD, Hui FK. Artificial intelligence to diagnose ischemic stroke and identify large vessel occlusions: a systematic review. *J Neurointerv Surg*. 2020;12:156-164. <https://doi.org/10.1136/neurintsurg-2019-015135>
12. Wardlaw JM, Mair G, von Kummer R, et al. Accuracy of automated computer-aided diagnosis for stroke imaging: a critical evaluation of current evidence. *Stroke*. 2022;53:2393-2403. <https://doi.org/10.1161/STROKEAHA.121.036204>
13. Dusenbury W, Mathiesen C, Whaley M, et al.; American Heart Association Council on Cardiovascular and Stroke Nursing and the Stroke Council. Ideal foundational requirements for stroke program development and growth: a scientific statement from the American Heart Association. *Stroke*. 2023;54:e175-e187. <https://doi.org/10.1161/STR.00000000000000424>
14. Cimflova P, Golan R, Ospel JM, et al. Validation of a machine learning software tool for automated large vessel occlusion detection in patients with suspected acute stroke. *Neuroradiology*. 2022; 64:2245-2255. <https://doi.org/10.1007/s00234-022-02978-x>
15. 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;51:259-264. <https://doi.org/10.1159/000519125>
16. Seker F, Pfaff JAR, Mokli Y, et al. Diagnostic accuracy of automated occlusion detection in CT angiography using e-CTA. *Int J Stroke*. 2022;17(1):77-82. <https://doi.org/10.1177/1747493021992592>
17. Mallon DH, Taylor EJR, Vittay OI, Sheeka A, Doig D, Lobotesis K. Comparison of automated ASPECTS, large vessel occlusion detection and CTP analysis provided by Brainomix and RapidAI in patients with suspected ischaemic stroke. *J Stroke Cerebrovasc Dis*. 2022;31:106702. <https://doi.org/10.1016/j.jstrokecerebrovasdis.2022.106702>
18. Dehkharghani S, Lansberg M, Venkatsubramanian C, et al. High-performance automated anterior circulation CT angiographic clot detection in acute stroke: a multireader comparison. *Radiology*. 2021;298:665-670. <https://doi.org/10.1148/radiol.202102734>
19. Karamchandani RR, Helms AM, Satyanarayana S, et al. Automated detection of intracranial large vessel occlusions using Viz.ai software: experience in a large, integrated stroke network. *Brain Behav*. 2023;13:e2808. <https://doi.org/10.1002/brb3.2808>
20. Chan N, Sibtain N, Booth T, et al. Machine-learning algorithm in acute stroke: real-world experience. *Clin Radiol*. 2023;78:e45-e51. <https://doi.org/10.1016/j.crad.2022.10.007>
21. Luijten SPR, Wolff L, Duvekot MHC, et al. Diagnostic performance of an algorithm for automated large vessel occlusion detection on CT angiography. *J Neurointerv Surg*. 2022;14:794-798. <https://doi.org/10.1136/neurintsurg-2021-017842>
22. Werdiger F, Gotla S, Visser M, et al. Automated occlusion detection for the diagnosis of acute ischemic stroke: a detailed performance review. *Eur J Radiol*. 2023;164:110845. <https://doi.org/10.1016/j.ejrad.2023.110845>
23. Andralojc LE, Kim DH, Edwards AJ. Diagnostic accuracy of a decision-support software for the detection of intracranial large-vessel occlusion in CT angiography. *Clin Radiol*. 2023;78:e313-e318. <https://doi.org/10.1016/j.crad.2022.10.017>
24. Wolff L, Uniken Venema SM, Luijten SPR, et al. Diagnostic performance of an algorithm for automated collateral scoring on computed tomography angiography. *Eur Radiol*. 2022;32:5711-5718. <https://doi.org/10.1007/s00330-022-08627-4>
25. Grunwald IQ, Kulikovski J, Reith W, et al. Collateral automation for triage in stroke: evaluating automated scoring of collaterals in acute stroke on computed tomography scans. *Cerebrovasc Dis*. 2019;47:217-222. <https://doi.org/10.1159/000500076>
26. Verdolotti T, Pilato F, Cottonaro S, et al. ColorViz, a new and rapid tool for assessing collateral circulation during stroke. *Brain Sci*. 2020;10:882. <https://doi.org/10.1007/s00330-022-08627-4>
27. Jabal MS, Kallmes DF, Harston G, et al. Automated CT angiography collateral scoring in anterior large vessel occlusion stroke: a multi-reader study. *Interv Neuroradiol*. 2023;15910199221150470. <https://doi.org/10.1177/15910199221150470>
28. Nagaratnam K, Neuhaus A, Briggs JH, et al. Artificial intelligence-based decision support software to improve the efficacy of acute stroke pathway in the NHS: an observational study. *Front Neurol*. 2023;14:1329643. <https://doi.org/10.3389/fneur.2023.1329643>
29. Martinez-Gutierrez JC, Kim Y, Salazar-Marioni S, et al. Automated large vessel occlusion detection software and thrombectomy treatment times: a cluster randomized clinical trial. *JAMA Neurol*. 2023;80:1182-1190. <https://doi.org/10.1001/jama.neuro.2023.3206>
30. Al-Kawaz M, Primiani C, Urrutia V, Hui F. Impact of RapidAI mobile application on treatment times in patients with large vessel occlusion. *J Neurointerv Surg*. 2022;14:233-236. <https://doi.org/10.1136/neurintsurg-2021-017365>
31. Hassan AE, Ringheanu VM, Rabah RR, Preston L, Tekle WG, Qureshi AI. Early experience utilizing artificial intelligence shows significant reduction in transfer times and length of stay in a hub and spoke model. *Interv Neuroradiol*. 2020;26:615-622. <https://doi.org/10.1177/1591019920953055>
32. Goyal M, Menon BK, van Zwam WH, et al. Endovascular thrombectomy after large-vessel ischaemic stroke: a meta-analysis of individual patient data from five randomised trials. *Lancet*. 2016;387:1723-1731. [https://doi.org/10.1016/S0140-6736\(16\)00163-X](https://doi.org/10.1016/S0140-6736(16)00163-X)
33. Tan IY, Demchuk AM, Hopyan J, et al. CT angiography clot burden score and collateral score: correlation with clinical and radiologic outcomes in acute middle cerebral artery infarct. *AJNR Am J Neuroradiol*. 2009;30:525-531. <https://doi.org/10.3174/ajnr.A1408>
34. Hillis SL, Obuchowski NA, Schartz KM, Berbaum KS. A comparison of the Dorfman–Berbaum–Metz and Obuchowski–Rockette methods for receiver operating characteristic (ROC) data. *Stat Med*. 2005;24:1579-1607. <https://doi.org/10.1002/sim.2024>
35. Smith BJ, Hillis SL, Pesce LL. *MRMCAov: Multi-Reader Multi-Case Analysis of Variance*. R Package Version 0.3.0. 2023. Accessed February 12, 2025. <https://cran.r-project.org/package=MRMCAov>
36. Lobotesis K, Veltkamp R, Carpenter IH, Claxton LM, Saver JL, Hodgson R. Cost-effectiveness of stent-retriever thrombectomy in combination with IV t-PA compared with IV t-PA alone for acute ischemic stroke in the UK. *J Med Econ*. 2016;19:785-794. <https://doi.org/10.1080/13696998.2016.1174868>

37. Brinjikji W, Abbasi M, Arnold C, et al. e-ASPECTS software improves interobserver agreement and accuracy of interpretation of aspects score. *Interv Neuroradiol.* 2021;27:781-787. <https://doi.org/10.1177/15910199211011861>
38. Chen W, Wu J, Wei R, et al. Improving the diagnosis of acute ischemic stroke on non-contrast CT using deep learning: a multicenter study. *Insights Imaging.* 2022;13:184. <https://doi.org/10.1186/s13244-022-01331-3>
39. Kobeissi H, Kallmes DF, Benson J, et al. Impact of e-ASPECTS software on the performance of physicians compared to a consensus ground truth: a multi-reader, multi-case study. *Front Neurol.* 2023;14:1221255. <https://doi.org/10.3389/fneur.2023.1221255>
40. Mallon D, Fallon M, Blana E, et al. Real-world evaluation of Brainomix e-Stroke software. *Stroke Vasc Neurol.* 2024;9: 497-504. <https://doi.org/10.1136/svn-2023-002859>