Artificial intelligence to diagnose ischemic stroke and identify large vessel occlusions: a systematic review

Nick M Murray, Mathias Unberath, Gregory D Hager, Ferdinand K Hui

ABSTRACT

Background and purpose. Acute stroke caused by large vessel occlusions (LVOs) requires emergent detection and treatment by endovascular thrombectomy. However, radiologic LVO detection and treatment is subject to variable delays and human expertise, resulting in morbidity. Imaging software using artificial intelligence (AI) and machine learning (ML), a branch of AI, may improve rapid frontline detection of LVO strokes. This report is a systematic review of AI in acute LVO stroke identification and triage, and characterizes LVO detection software.

Methods. A systematic review of acute stroke diagnostic-focused AI studies from January 2014 to February 2019 in PubMed, Medline, and Embase using terms: ‘artificial intelligence’ or ‘machine learning or deep learning’ and ‘ischemic stroke’ or ‘large vessel occlusion’ was performed.

Results. Variations of AI, including ML methods of random forest learning (RFL) and convolutional neural networks (CNNs), are used to detect LVO strokes. Twenty studies were identified that use ML. Alberta Stroke Program Early CT Score (ASPECTS) commonly used RFL, while LVO detection typically used CNNs. Image feature detection had greater sensitivity with CNN than with RFL, 85% versus 68%. However, AI algorithm performance metrics use different standards, precluding ideal objective comparison. Four current software platforms incorporate ML: Brainomix (greatest validation of AI for ASPECTS, uses CNNs to automatically detect LVOs), General Electric, ischemiaView (largest number of perfusion study validations for thrombectomy), and Viz.ai (uses CNNs to automatically detect LVOs, then automatically activates emergency stroke treatment systems).

Conclusions. AI may improve LVO stroke detection and rapid triage necessary for expedited treatment. Standardization of performance assessment is needed in future studies.

INTRODUCTION: ISCHEMIC STROKE AND ARTIFICIAL INTELLIGENCE

Acute ischemic strokes caused by large vessel occlusions (LVOs) require emergent detection in pre-hospital triage, similar to the analogous detection of acute myocardial infarcts by electrocardiograms. The subsequent in-hospital verification of LVOs and treatment by endovascular thrombectomy (ET) are critical for reducing morbidity and mortality. The number needed to treat to reduce disability by at least one level on the modified Rankin Scale for one patient with an LVO is 2.0–2.6. LVOs are detected from ASPECTS (Alberta Stroke Programme Early CT Score) or non-invasive cerebral angiography and perfusion studies that also measure infarct core and penumbral volume mismatch. The results of these studies inform candidacy for acute ET.

However, clinically salient interpretation of perfusion, angiographic, and ASPECTS is subject to inconsistent local expertise, time delays, and varies between institutions. Furthermore, activation of interhospital communication for LVO triage and transport to a thrombectomy center can be operationally challenging. There is frequent time delay, of nearly 100 min, even in experienced systems, which may result in loss of neurons, increased morbidity, or even exclusion as a candidate for an ET intervention. For every 15 min of time reduction to ET recanalization of an LVO, 34 per 1000 treated patients have an improved disability outcome. Despite stroke system-optimization efforts, there remains an unmet need for more immediate and standardized time-sensitive stroke detection and triage.

Artificial intelligence (AI) is proposed as a tool to deal with this need. Recently, AI has been applied to standardize accurate LVO stroke detection. AI technology varies among stroke imaging software platforms and remains mostly proprietary. A systematic review of stroke AI software, including AI methodology and validation, has not yet been performed. Before doing so in this study, an introduction to AI fundamentals is reviewed.

We focus on a brief review of machine learning (ML) for image processing since diagnostic stroke imaging modalities fall under this category. ML is an area of AI-related research that provides tools to develop or discover decision-making rules from data. Algorithms used for LVO stroke detection use pure ML on highly specific data and fall into the category of ‘narrow AI’ that performs well only on a well-circumscribed task. In the case of stroke, the primary goal of any AI or ML algorithm is to reliably identify the presence or absence of an LVO from three-dimensional tomographic images.

However, many different approaches to achieving this goal using ML methods can be envisioned. These differ in the choice of input abstraction (ie, image input manipulation, either as manually engineered image abstract feature representations or non-abstract: pixel intensities), function approximation (ie, an algorithmic method of estimating or fitting the value of different image data states that have similar features; this includes methods such as a support vector machine, random forest learning (RFL), or deep neural network), final algorithm...
Table 1  Overview of the most commonly used machine learning (ML) design components and potential choices together with the associated implications for the ML system

<table>
<thead>
<tr>
<th>Design component (definition)</th>
<th>Common choices</th>
<th>Implications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input abstraction (Raw or modified data that are the input for the ML algorithm)</td>
<td>Hand-crafted features</td>
<td>Enables introduction of domain expertise by forcing inference from low-dimensional signals derived from high-dimensional volumes. This approach is generally considered to be more ‘interpretable’ since features are hand-crafted, but the performance of ML algorithms operating on this type of input is usually also more limited</td>
</tr>
<tr>
<td>No abstraction</td>
<td></td>
<td>Training ML algorithms directly on pixel intensities is difficult owing to the high dimensionality of the inputs, which makes it difficult to discover semantic correlations between intensity patterns and desired output during training. If successful, this approach gives rise to state-of-the-art performance at the cost of reduced interpretability</td>
</tr>
<tr>
<td>Function approximation (ML algorithm that computes significance of the input variables)</td>
<td>Support vector machine (SVM)</td>
<td>A linear learning model and associated training algorithm that can be made non-linear via a process termed ‘kernel trick’ that implicitly maps the inputs to a higher dimensional feature domain. SVM operates on feature vectors, which are n-dimensional vectors of numerical features that represent the object of interest in the images. All images represented as feature vector then define an n-dimensional vector space, also called feature space. The feature values are assigned weights to construct a linear predictor function SVM, which requires n to be small—that is, the feature space to be low dimensional (abstraction); it performs well in practice</td>
</tr>
<tr>
<td>Random forest learning (RFL)</td>
<td></td>
<td>Random forests are an ensemble learning method that aggregates predictions of multiple decision trees that process a feature vector extracted from the input sample. Random forests are appealing since they possess a natural way of ranking the importance of features for the classification or regression problem</td>
</tr>
<tr>
<td>Deep neural networks (DNN)</td>
<td></td>
<td>A non-linear learning method and associated training algorithm that consists of multiple, stacked layers. Values in every layer are computed by weighted summation of all inputs from the previous layer, the result of which is passed through a non-linear activation function. Neural networks are universal function approximators but in the fully connected approach (any value depends on all values in the previous layer) require input abstraction to be practical</td>
</tr>
<tr>
<td>Deep convolutional neural networks (CNN or DCNN)</td>
<td></td>
<td>A variant of deep neural networks that assumes translational invariance of input patterns, and therefore, enables weight sharing across the spatial domain, which yields substantially more compact models (online supplementary figure 1). Consequently, these models do not require input abstraction and can be applied directly to pixel intensities. Training of these models is complicated, often via backpropagation learning, and requires large amounts of data since, while more compact than deep neural networks, these models still have multiple millions of trainable parameters</td>
</tr>
<tr>
<td>Output format (Fitting of the input data to an estimated outcome)</td>
<td>Instance-level classification/ regression</td>
<td>Since the output is constrained to the instance level, the ML algorithm performs heavy dimensionality reduction, which may make the learning algorithm susceptible to overfitting</td>
</tr>
<tr>
<td>Supervisory signal (Amount of human guidance and training given to the function approximation algorithm during its development and/or validation study)</td>
<td>Un-/Self-supervised</td>
<td>If no ground truth labels are available for training, learning models can potentially still be trained by optimizing for encodings that, for example, form clusters in the encoded domain or can be decoded to reproduce the input. While such approaches may be able to extract good representations of the input data, they will not be able to correlate these representations with semantic labels unless provided by the user retrospectively</td>
</tr>
<tr>
<td>Supervised, structured</td>
<td>In supervised learning, every instance has an associated ground truth label (outcome) that is to be predicted using the learning model. If the annotation is not perfectly accurate, one usually refers to gold (or reference) standard rather than ground truth. Supervised learning generally yields models with the highest performance, but reliable annotations are usually expensive to obtain</td>
<td></td>
</tr>
<tr>
<td>Supervised, unstructured</td>
<td>In some cases, gold standard annotations are available in unstructured form—for example, as medical records. In this case, methods must be used that process unstructured data to enable model training. In the context of medical records, natural language processing can be used to extract structured annotations</td>
<td></td>
</tr>
</tbody>
</table>

Table 1 lists further definitions and the most commonly used setups for the aforementioned design choices.

In the case of ML used for LVO stroke detection, the input abstraction may be a hybrid of methods. Most function approximations commonly applied to the stroke image input are support vector machine, RFL, or deep neural network, using a classification scheme of the input features. To do so, the function approximation receives a supervisory signal, commonly structured supervised, to the ‘gold standard’ of the human’s radiological scan interpretation—the presence or absence of an LVO stroke.

A full understanding of which imaging platforms use AI, how accurately the AI performs, and how the clinician can use AI in real time to interpret images is needed. This understanding is necessary for informed interpretation of the scans of any individual patient with a stroke that use AI software. Previous studies have described some forms of AI in stroke. ML has been well described in ASPECTS and perfusion to diffusion mismatch, but there are no published studies on ML use in the diagnosis of ischemic stroke by directly identifying LVOs from source images. Although abstracts describing this have
been published,\(^2\)\(^3\) one retrospective study applying AI to vessel densities as an indirect measure of LVO presence,\(^4\)\(^7\) and a single retrospective study using AI analysis of patient demographics and clinical examination characteristics to predict LVO,\(^2\)\(^8\) to our knowledge there are no studies describing AI for LVO detection from source CT or MR image files.

Our study is a systematic review of the current landscape of AI in ischemic stroke diagnostics by characterizing: (1) the current literature and (2) new ML diagnostic technology for non-contract CT, CT angiography (CTA), and CT perfusion (CTP) images that has only recently been made available to clinicians, but not yet fully described yet in the literature. The diagnostic utility of different types of ML in patients with a stroke, with respect to acute diagnosis, triage, and prognosis is compared.

**METHODS**

A systematic review was performed in accordance with PRISMA guidelines using PubMed, Medline, and Embase for all peer-reviewed articles in English using predetermined search terms: ‘artificial intelligence’ or ‘machine learning’ or ‘deep learning’ and ‘ischemic stroke’ or ‘large vessel occlusion’. Studies were included if the primary aim was describing acute ischemic LVO diagnosis or prediction of acute outcome (including recanalization, symptomatic intracerebral hemorrhage, or acute in-hospital infarct growth), but not long-term outcomes. Studies that resulted using this search criteria were excluded only if they did not meet inclusion criteria or their primary aim or outcome was any of the following: association with stroke severity as measured by the National Institutes of Health Stroke Scale (NIHSS), or prognostic or effect of treatment study (eg, modified Rankin Scale score at ≥3 months), an AI development-only study (ie, without AI validation), a chronic stroke burden assessment or stroke MRI volume quantification study, or did not apply AI to raw imaging files. Only the past 5 years’ literature, published after January 1, 2014, was included, which is a timeframe that was felt to optimally capture the emerging technologies without redescribing well-known image analysis and AI technology.

Studies were extracted by a single author and reviewed by a second author independently. Unpublished emerging stroke diagnostic and LVO detection technology using both CTA and CTP was identified by worldwide web searches for the same search terms as used in the above databases. This technology assessment was verified and cross-checked by review of FDA 510(k) and European CE mark documents, manufacturer websites, and conference posters. The present review protocol is not a registered protocol. The last date for which the sources were searched was February 28, 2019. Bias in individual studies was assessed but wide variability in design and statistical calculations of the reviewed studies prevented calculation of principal summary measures; further details are contained in the online supplementary Methods.

**RESULTS**

**Systematic review of AI algorithms for acute stroke diagnostics**

A total of 20 studies met all inclusion and no exclusion criteria (figure 1). AI use in acute LVO stroke diagnostics and triage falls under three categories: automatic stroke core and penumbra size and mismatch quantification, detection of vascular thrombi or occlusions, and prediction of acute complications (described below and in table 2).

The majority of current studies describe AI-mediated detection of core stroke lesion volume on CT and MRI scans, and some use these features to predict acute stroke growth.\(^16\)\(^19\) For non-contract CT analysis specifically, automated ASPECTS from ML used in conjunction with clinical presentation accurately associates NIHSS score and can be used to select patients for ET.\(^20\)\(^21\) ML algorithms for ASPECTS most commonly employ RFL. The 10 studies reviewed that use RFL show that the AI often outperforms single radiologist ASPECTS and is non-inferior or even better than consensus ASPECTS. The reported sensitivity of AI algorithm ASPECTS range from 45% to 98%, mean 68%, and specificity ranges from 57% to 95%, mean 81%.\(^9\)\(^10\)\(^14\)\(^16\)\(^17\)\(^22\)\(^27\)

Use of a convolutional neural network (CNN) for a combined asymmetric middle cerebral artery territory hypodensity and dense vessel detection may have higher performance; however, only area under the curve (AUC) metrics are reported (receiver operating characteristic AUC 92–96%).\(^28\) Collectively among these studies, there is significant heterogeneity in study data, including cohort size, gold standard comparison, and time to initial head CT acquisition (table 2, online supplementary table I). Several commercial software platforms offer a ML algorithm for ASPECTS within their software: Brainomix e-ASPECTS (Oxford, UK), Siemens Frontier (Erlangen, Germany), iSchema-View ASPECTS (Menlo Park, California, USA), and others that are in development.

No peer-reviewed studies to date have described AI applied to direct LVO detection from CT angiographic studies. However, abstracts describing this have been published.\(^12\)\(^29\) We note that abstract extraction partially limits the systematic review below for bias, assessment of heterogeneity, characterization of methods, and other desired comparisons.

Many AI algorithms for different types of LVO characterization have been described. These include detection of vessel density asymmetry on CTA, presence of middle cerebral artery dot sign, biompedance asymmetry from phase shift spectroscopy, and true intravascular thrombus versus chronic

---

**Figure 1** Systematic literature screening and review protocol for studies included and excluded. Inclusion dates were January 1, 2014 to February 28, 2019.
<table>
<thead>
<tr>
<th>Study</th>
<th>Aim</th>
<th>Inclusion and exclusion criteria</th>
<th>Study type</th>
<th>AI method and Software</th>
<th>Secondary validation,</th>
<th>Accuracy metrics</th>
<th>n</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maegerlein et al 2019</td>
<td>AI vs neuroradiologist ASPECTS</td>
<td>Acute MCA with or without an LVO</td>
<td>SCRV</td>
<td>RFL, RAPID</td>
<td>Two neuroradiologists</td>
<td>Kappa human consensus: ~57%, AI 90%</td>
<td>152</td>
<td>ASPECTS by AI has better agreement to consensus score than independent neuroradiologists</td>
</tr>
<tr>
<td>Chatterjee et al 2019</td>
<td>CNN to detect LVOs from CTA images</td>
<td>Anterior circulation stroke syndromes</td>
<td>SCRV</td>
<td>CNN, Viz.ai</td>
<td>Radiology report without AI</td>
<td>Sensitivity 82%, specificity 94%, PPV 77%, NPV 95%</td>
<td>650</td>
<td>CNNs detect LVOs with good accuracy</td>
</tr>
<tr>
<td>Öman et al 2019</td>
<td>CNN to detect MCA strokes from CTA source images</td>
<td>Stroke code patients: MCA cases and final diagnosis of non-stroke controls</td>
<td>SCRV; two-part AI development and validation</td>
<td>CNN, 3D slicer, DeepMedic</td>
<td>Two radiologists' manual segmentation (ASPECTS) and quantitative voxel score</td>
<td>Sensitivity 67–74%, specificity 93–96%, AUC 91–93%</td>
<td>60</td>
<td>CNNs can use CTA source images to detect hypodensities correlated to acute ischemic stroke</td>
</tr>
<tr>
<td>Kuang et al 2019</td>
<td>ASPECTS vs MRI DWI</td>
<td>Acute ischemic stroke CT head and then MRI DWI brain</td>
<td>SCRV</td>
<td>RFL</td>
<td>MRI DWI sequences</td>
<td>Kappa 88%, sensitivity 98%, specificity 80%</td>
<td>157</td>
<td>RFL reliably assigns ASPECTS</td>
</tr>
<tr>
<td>Seker et al 2019</td>
<td>AI vs neuroradiologist ASPECTS with different CT postprocessing reconstruction</td>
<td>Acute proximal MCA occlusion</td>
<td>SCRV</td>
<td>RFL, Brainomix e-ASPECTS</td>
<td>Four radiologists</td>
<td>e-ASPECTS performed better than residents: kappa 92% vs ~75%</td>
<td>43</td>
<td>e-ASPECTS performs better than radiologists with varied reconstructions of CT scans from the same patient</td>
</tr>
<tr>
<td>Olive-Gadea et al 2019</td>
<td>AI vs radiologist ASPECTS association with CTP infarct core</td>
<td>Acute proximal MCA or ICA terminus occlusions that received ET</td>
<td>SCRV</td>
<td>RFL, Brainomix e-ASPECTS, RAPID CTP</td>
<td>RAPID CTP core</td>
<td>Radiologist vs AI ASPECTS correlation 44%, both scores correlated with CBV (r = −0.41 vs −0.54)</td>
<td>184</td>
<td>Radiologist and AI ASPECTS compared with CBV value similarly predict infarct volume</td>
</tr>
<tr>
<td>Chriashkova et al 2019</td>
<td>ASPECTS and time to score</td>
<td>Acute ischemic stroke syndromes</td>
<td>SCRV</td>
<td>RFL, Brainomix e-ASPECTS</td>
<td>26 Clinicians (including 11 radiologists) with access to clinical symptoms and follow-up scans</td>
<td>AI reduced scoring time by 34%, kappa improved from 26% to 38%</td>
<td>64</td>
<td>AI in e-ASPECTS improved efficiency and accuracy</td>
</tr>
<tr>
<td>Barreira et al 2018</td>
<td>CNN vs neurologist to detect LVOs from CTA</td>
<td>Acute ischemic stroke with or without LVO</td>
<td>Multicenter retrospective study</td>
<td>CNN, Viz.ai</td>
<td>One stroke neurologist</td>
<td>Sensitivity of 97% and specificity of 52%, PPV 74%, NPV 91%; overall accuracy 78%</td>
<td>152</td>
<td>CNN have high positive predictive value for presence of an LVO</td>
</tr>
<tr>
<td>Barreira et al 2018</td>
<td>CNN vs neurologists to detect LVOs from CTA</td>
<td>Acute ischemic stroke with or without LVO</td>
<td>Three-center retrospective review</td>
<td>CNN, Viz.ai</td>
<td>Several experienced stroke neurologists</td>
<td>Sensitivity 90%, specificity 83%, accuracy 86%, PPV 82%, NPV 91%, AUC 90%</td>
<td>875</td>
<td>CNN AI detects more than 82% of anterior circulation LVOs</td>
</tr>
<tr>
<td>Boldsen et al 2018</td>
<td>AI vs radiologist for an ATLAS core stroke</td>
<td>Acute anterior circulation stroke, MRI DWI scan</td>
<td>Multicenter, retrospective case training, validation</td>
<td>RFL</td>
<td>One imaging expert</td>
<td>Dice coefficient 0.61</td>
<td>109</td>
<td>AI provides reliable penumbra and core size</td>
</tr>
<tr>
<td>Kellner et al 2018</td>
<td>AI vs clinical scales to identify severe stroke types, including LVO</td>
<td>Any brain neurological presentation with neurological imaging</td>
<td>Multicenter, combined pilot and validation study</td>
<td>ML subtype not defined</td>
<td>Multiple trial coordinators and clinicians</td>
<td>Sensitivity 93%, specificity 87%, AUC 95%</td>
<td>252</td>
<td>Volumetric impedance phase shift spectroscopy run by ML reliably identifies severe strokes</td>
</tr>
<tr>
<td>Lucas et al 2018</td>
<td>CNN vs radiologist to predict CTP stroke growth</td>
<td>Patients with a stroke, each with a full set of CTP data</td>
<td>SCRV AI development, validation</td>
<td>CNN</td>
<td>One expert manual rater</td>
<td>Dice coefficient 0.46–0.53</td>
<td>29</td>
<td>CNN predicts infarct growth well</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Study</th>
<th>Aim</th>
<th>Inclusion and exclusion criteria</th>
<th>Study type</th>
<th>AI method and Software</th>
<th>Secondary validation,</th>
<th>Accuracy metrics</th>
<th>n</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goebel et al 2018&lt;sup&gt;22&lt;/sup&gt;</td>
<td>ASPECTS from: AI softwares vs each other vs radiologist</td>
<td>Acute MCA ischemia, no hemorrhage or artifacts</td>
<td>SCRV, head-to-head scoring</td>
<td>Frontier vs Brainomix vs radiologist</td>
<td>Two senior radiologists</td>
<td>Agreement interclass correlation: human and Brainomix 0.71–0.84 vs human and frontier 0.47–0.54</td>
<td>150</td>
<td>High ASPECTS agreement between consensus radiologists and Brainomix</td>
</tr>
<tr>
<td>Guberina et al 2018&lt;sup&gt;23&lt;/sup&gt;</td>
<td>AI vs radiologist ASPECTS with acute and chronic strokes</td>
<td>Acute MCA ischemia within 6 hours from symptom onset who were eligible for reperfusion therapy</td>
<td>SCRV, definite infarct core on head CT at 24 hours</td>
<td>RFL, Brainomix e-ASPECTS</td>
<td>Three neuroradiologists</td>
<td>AI sensitivity 83% vs human 73%, specificity 57% vs 84%</td>
<td>119</td>
<td>RFL has slightly higher sensitivity but worse specificity for ASPECTS</td>
</tr>
<tr>
<td>Nagel et al 2017&lt;sup&gt;19&lt;/sup&gt;</td>
<td>AI vs radiologist ASPECTS</td>
<td>Acute anterior circulation ischemic stroke syndrome, baseline and 24-36-hour follow-up head CT scans</td>
<td>Retrospective, multicenter comparison study</td>
<td>RFL, Brainomix e-ASPECTS</td>
<td>Three radiologists</td>
<td>AI sensitivity ~45%, specificity ~92%, accuracy ~86%</td>
<td>132</td>
<td>RFL was non-inferior to neuroradiologists for ASPECTS</td>
</tr>
<tr>
<td>Lisowksa et al 2017&lt;sup&gt;28&lt;/sup&gt;</td>
<td>CNN vs radiologist acute stroke hypodensity and dense vessel sign detection</td>
<td>Acute ischemic stroke within 6 hours from symptom onset with non-contrast CT</td>
<td>Retrospective, pooled patients from three studies</td>
<td>CNN, Keras on Theano (Python)</td>
<td>One researcher, one radiologist</td>
<td>AI AUC 92–96%</td>
<td>51</td>
<td>CNN detection of asymmetrical hypodensity and dense vessel can detect ischemia well</td>
</tr>
<tr>
<td>Herweh et al 2016&lt;sup&gt;6&lt;/sup&gt;</td>
<td>AI vs neurologist ASPECTS</td>
<td>Acute stroke, CT and MRI with DWI images &lt;2 hours apart</td>
<td>Retrospective single center study</td>
<td>RFL, Brainomix e-ASPECTS</td>
<td>Three stroke experts and three neurology trainees, DWI as gold standard</td>
<td>AI sensitivity 46%, specificity 95%, AI correlation coefficient 0.44 vs humans 0.19–0.38</td>
<td>34</td>
<td>RFL accuracy for ASPECTS was similar to neurologists</td>
</tr>
<tr>
<td>Bentley et al 2014&lt;sup&gt;37&lt;/sup&gt;</td>
<td>AI vs radiologist prediction of sICH, mRS</td>
<td>Acute ischemic stroke treated with IV tPA</td>
<td>SCRV, moderate-supervision of AI</td>
<td>SVM</td>
<td>Radiologist SEDAN and HAT scoring</td>
<td>Radiologist 67% vs AI 74% AUC</td>
<td>112</td>
<td>AI is superior to radiologist scales for sICH</td>
</tr>
<tr>
<td>Takahashi et al 2014&lt;sup&gt;81&lt;/sup&gt;</td>
<td>AI vs neuroradiologist detection of MCA dot sign</td>
<td>Anterior ischemic stroke, first head CT within 24 hours</td>
<td>Two-center, retrospective study</td>
<td>SVM</td>
<td>Two neuroradiologists</td>
<td>AI sensitivity 98%, false positive rate 1.3% per image</td>
<td>7</td>
<td>AI SVM is accurate to detect the MCA dot sign</td>
</tr>
<tr>
<td>Thornhill et al 2014&lt;sup&gt;93&lt;/sup&gt;</td>
<td>AI vs neuroradiologist intraluminal thrombus from plaque</td>
<td>Patients with acute TIA or stroke with suspected free-floating luminal thrombus on CTA</td>
<td>SCRV, training and ML validation</td>
<td>ANN, SVM</td>
<td>Two neuroradiologists</td>
<td>SVM sensitivity 88%, specificity 71%, AUC 85%</td>
<td>23</td>
<td>AI is able to distinguish shape signatures of intraluminal thrombus from plaques</td>
</tr>
</tbody>
</table>

AI, artificial intelligence; ANN, artificial neural network; ASPECTS, Alberta Stroke Programme Early CT Score; AUC, area under the curve; CBV, cerebral blood volume; CNN, convolutional neural network; CTA, CT angiography; CTP, CT perfusion; CT, computed tomography; CT, computed tomography; ET, endovascular thrombectomy; ICA, internal carotid artery; LINDA, lesion identification with neighborhood data analysis; LVO, large vessel occlusion; MCA, middle cerebral artery; ML, machine learning; mRS, modified Rankin Scale; NPV, negative predictive value; PPV, positive predictive value; RFL, random forest learning; SCRV, single-center, retrospective review; sICH, symptomatic intracerebral hemorrhage; SVM, support vector machine (form of artificial neural network); TIA, transient ischemic attack; tPA, tissue plasminogen activator.
These LVO detection studies variably report algorithm performance compared with individual humans, with broad sensitivities of 67–98% and AUC ranging from 83% to 93%. Many use human consensus scoring of the CTA or ASPECTS as the gold standard for comparison. Of these studies reviewed for AI-mediated LVO detection, four clearly describe use of CNNs, and this ML type is the most common AI used. These studies using a CNN to directly detect LVOs from CTA images report a sensitivity range from 67% to 94%, mean 83%, and specificity from 52% to 94%, mean 91%. 12 18 20

AI algorithms are also used for acute prognosis prediction. This may facilitate immediate treatment planning—for example, whether or not to offer IV tissue plasminogen activator or ET due to possible postintervention complications, or near-term outcome prediction, such as risk of intracranial hemorrhage. In general, AI has benefited these prognoses. 21 34–36 For example, one study showed that AI led to an absolute increase of 7% accuracy in predicting a symptomatic intracerebral hemorrhage. 17

**Figure 2** Brainomix (A) non-contrast head CT, (B) colorimetric analysis and overlayed e-ASPECTS, (C) e-CTA with LVO marked by red circle, and (D) Olea sphere thresholded core in red (<40% CBF threshold) and penumbra in yellow (>6s Tmax threshold) for a patient with an acute stroke with a left MCA occlusion. Images courtesy of Brainomix. ASPECTS, Alberta Stroke Program Early CT Score; CBF, cerebral blood flow; CTA, CT angiography; LVO, large vessel occlusion; MCA, middle cerebral artery.
Figure 3 RAPID image processing outputs. (A) CT perfusion (CTP) map of a right middle cerebral artery (MCA) syndrome. The CTP stroke core of 1 mL is represented in pink at the top left and to the right is the penumbra of 129 mL. MR perfusion status is shown 24 hours after thrombectomy below for validation. (B) CT angiography image of a patient with a left MCA occlusion (left) and the RAPID CT angiography detection of a >55% reduction in blood vessel density asymmetry, an indirect suggestion of a relevant flow limiting the underlying large vessel occlusion. Images courtesy of iSchemaView. CBF, cerebral blood flow.

no published validation are data available for this CTA software. In December 2018, iSchemaView received 510(k) approval for a thrombectomy selection guide. This incorporates inclusion criteria from major clinical trials that used RAPID for patient selection and applies this to the individual scan, providing a simple binary output of whether or not the patient is a candidate for thrombectomy. Finally, if ET is selected for the individual patient, then the interpreting provider manually activates all stroke treatment systems.

GE Healthcare also offers a platform for perfusion imaging in patients with suspected LVO. The software, FastStroke and CT perfusion 4D Neuro, is, similarly, automated software for analyzing CT angiographic and perfusion images related to stroke and tumor angiogenesis. It offers color-coded arterial and venous displays termed ColorViz for collateral circulation assessment and unique volumetric visualization of functional core and penumbra perfusion maps. Further software details and AI algorithm metrics were not provided by GE.

The newest AI stroke diagnostic software is from Viz.ai. It received de novo regulatory clearance from the FDA in February 2018 for the first computer-aided triage and notification platform to identify LVO strokes in CTA imaging (Viz LVO). This was followed in April 2018 with 510(k) clearance for Viz.ai’s CTP software. The Viz LVO and Viz CTP software platforms encompass similar perfusion features to the CTP outputs by iSchemaView, Brainomix/Olea Sphere, and GE 4D Neuro, but combine additional components for automating expedited stroke care—namely, (1) delivery of dynamic CTP and CTA images directly to the users’ mobile device for real-time 3D image manipulation, (2) automatic direct LVO detection from CTA data (Viz LVO), (3) in-built motion correction capability within Viz CTP, and (4) automatic activation of emergent treatment systems when a LVO is detected (figure 4). In comparison, for European users only, Brainomix is also beginning to offer a mobile device PACs feature platform and direct LVO detection.

Viz.ai uses a CNN algorithm to automatically detect stroke patterns of LVOs. If a LVO is detected, the output of the CNN automatically alerts the stroke treatment team, without any action required by the provider who ordered the radiologic scan. The alert is sent via a mobile phone application compliant with the Health Insurance Portability and Accountability Act to the emergency room provider, neurologist, and the neurointerventional
surgeon. This immediate communication demonstrated faster notifications, saving on average 32 min before LVO treatment. Within the app, the provider can activate emergency transport systems for ET (figure 4).

**DISCUSSION AND CONCLUSIONS**

AI in acute stroke diagnostic imaging uses different ML algorithms for a variety of tasks, including identification of hypodensities on non-contrast CT scans and LVOs on CTA scans. The ML systems encompass mostly unmodified input abstraction input into a RFL or CNN function approximation. Reported diverse ML performance and accuracy measurements, however, limit comparison between algorithms. Many systems are evaluated against different standards on different quality input data, leading to inconsistent and potentially biased differences in accuracy metrics. Of the available studies, comparisons were made using the reported data. ASPECTS’ comparative accuracy with humans when the gold standard is different, either non-contrast CT or MRI with DWI, may be similar or better, depending on the AI algorithm. Specifically, RFL algorithms for ASPECTS were used in over 10 studies for validation, with nearly 70% sensitivity and over 80% specificity as compared with humans (table 2). Analysis of CNN algorithms for LVO detection were limited as data were extracted primarily from study abstracts. Systems using CNNs for LVO detection report performance metrics on average 8–10% greater than those of ML employing RFL, up to 85% mean sensitivity for automatic LVO detection. Notably, core and perfusion studies from RAPID CT and MR have the highest metrics for AI accuracy, >80%, with some datasets showing 100% sensitivity to predict favorable perfusion mismatch.

However, diagnosis of acute stroke by AI has not yet been perfected and errors still exist (table 2). AI algorithms in the peer-reviewed literature report a broad sensitivity metrics, with a mean of 68%, suggesting that some AI algorithms may miss up to one in three findings on imaging output (table 2). The reported reasons for failure of the algorithms are often similar. These include radiologic scan abnormalities from pre-existing central nervous system injury, inadequate contrast boluses, inability to correct for patient motion, and tortuous vessels that preclude evaluation of the contrast column. 12 14 24 Despite this, some versions of AI have good predictive power to increase standardized accurate stroke diagnosis and LVO triage. In such cases, AI used to assist in interpretation of traditional stroke imaging outputs may reduce false-negative human errors in image interpretation, increase efficiency of stroke triage, and therefore minimize neuronal death and long-term morbidity and mortality.

Comprehensive platforms by iSchemaView, Viz.ai, and Brainomix use different AI methods yet can automatically detect the presence of LVOs and, with the Viz.ai software, also instantly activate emergency LVO stroke treatment teams. The benefit of these AI algorithms is seen by improved sensitive stroke detection, as an aid to physician decision-making, and by acting as a catalyst for timely LVO detection. Yet lower specificities reported by some algorithms, in the 50% range, indicate that critical review of the radiological images by a physician remains important. As the current LVO detection methods aim for higher positive detection rates, the burden on the physician to navigate an increased number of scans containing false positives will also be noticed.

Additional validation studies for LVO diagnostics are rapidly developing as ML algorithms improve. It is important to recognize that FDA and CE mark approval of these algorithms allows them to be used by trained professionals, physicians, or medical technicians, and not as a substitute for human decision-making. Therefore, in clinical practice we predict that use of the AI software may assist clinicians in their detection accuracy of ischemic stroke.
Finally, among the AI software platforms commonly used in healthcare institutions, iSchemaView RAPID, Viz LVO and Viz CTP, General Electric, and Brainomix, each provide similar features (figure 5, online supplementary table II). Among these, AI for LVO detection is in different stages of validation. Some features differ significantly based on the software platform. iSchemaView and Brainomix AI (indirect form) do not directly detect LVOs, but infer LVO presence based on asymmetry in collateral blood vessel density. Brainomix (direct form) and Viz.ai offer direct LVO detection, though only Viz.ai has reported validation metrics. AI for both direct LVO detection and automatic emergency LVO treatment system activation is available only through the Viz.ai platform.

This study and systematic review have some limitations. In contrast to AI in ASPECTS and perfusion imaging, the emerging AI technologies in the past 1–2 years applied to direct and indirect LVO detection have only been described in abstracts and posters. No rigorous comparison or systematic studies have been published. Reporting of AI in LVO detection thus required evaluation of non-peer reviewed abstracts, posters, FDA, CE material, which permits exposure and reporting bias. The selected statistical comparisons and metrics reported in each study reviewed also vary. ML systems are evaluated against different standards of different qualities on different data, and thus metrics of sensitivity and specificity are not generalizable between ML systems.

There is a critical need for a clear definition of ‘ground truth’ against which algorithms are evaluated consistently. Having a consistent set of metrics is essential to improve ML in acute stroke care.

Overall, AI applied to stroke diagnosis and LVO detection has the potential to standardize and improve care. There remains a paucity of randomized controlled trials comparing AI software. Systematic and standardized methods for validation and comparison of this class of tools are needed.

Contributors NM, FH: protocol and project development, data collection and management, data analysis, manuscript writing and editing. MU, GDH: Data analysis, manuscript writing.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID IDs Nick M Murray http://orcid.org/0000-0003-3861-0958 Ferdinand K Hui http://orcid.org/0000-0003-7597-8866

REFERENCES
New Devices and Techniques


