

**NATIONAL INSTITUTE FOR HEALTH AND CARE  
EXCELLENCE**

**DIAGNOSTICS ASSESSMENT PROGRAMME**

**Diagnostics consultation document**

**Artificial intelligence (AI)-derived software to  
help clinical decision making in stroke**

The National Institute for Health and Care Excellence (NICE) is producing guidance on using software with artificial intelligence derived algorithms for analysing CT brain scans in the NHS in England. The diagnostics advisory committee has considered the evidence and the views of clinical and patient experts.

**This document has been prepared for public consultation.** It summarises the evidence and views that have been considered, and sets out the recommendations made by the committee. NICE invites comments from registered stakeholders, healthcare professionals and the public. This document should be read along with the [evidence](#) (the diagnostics assessment report and the diagnostics assessment report addendum).

The advisory committee is interested in receiving comments on the following:

- Has all of the relevant evidence been taken into account?
- Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?
- Are the recommendations sound, and a suitable basis for guidance to the NHS?

**Equality issues**

NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Please let us know if you think that the recommendations may need changing to meet these aims. In particular, please tell us if the recommendations:

- could have a different effect on people protected by the equality legislation than on the wider population, for example by making it more difficult in practice for a specific group to access the technology
- could have any adverse effect on people with a particular disability or disabilities.

Please provide any relevant information or data you have about such effects and how they could be avoided or reduced.

**Note that this document is not NICE's final guidance on software with artificial intelligence derived algorithms for analysing CT brain scans. The recommendations in section 1 may change after consultation.**

After consultation, the committee will meet again to consider the evidence, this document and comments from the consultation. After considering the comments, the committee will prepare its final recommendations, which will be the basis for NICE's guidance on the use of the technology in the NHS in England.

For further details, see the [diagnostics assessment programme manual](#).

**Key dates:**

Closing date for comments: 19 October 2023

## 1 Recommendations

NICE is aware that companies are reviewing their CE marking in response to changing regulations and advances in digital health technologies.

### Can be used in the NHS with evidence generation

1.1 The following artificial intelligence (AI)-derived software can be used in the NHS, to support review and reporting of CT brain scans for people who have had a stroke, while [more evidence is generated](#):

- e-Stroke
- RapidAI.

These technologies can be used once they have Digital Technology Assessment Criteria (DTAC) approval from NHS England.

1.2 The software should only be used with healthcare professional review and centres should maintain existing scan reporting protocols to reduce the risk of incorrect results. Centres should ensure that images shared between different stroke centres can be remotely reviewed to help with decision making by healthcare professionals at a different site.

## Can only be used in research

1.3 More research is needed on the following AI-derived software to support review and reporting of CT brain scans for people who have had a stroke:

- Accipio
- Aidoc
- BioMind
- BrainScan CT
- Cercare
- CINA Head
- CT Perfusion 4D
- icobrain ct
- Neuro Solution
- qER.

See [1.5 for key outcomes for further research.](#)

1.4 Access to the technologies listed in section 1.3 should be through company or research funding (non-core NHS funding).

## Evidence generation and further research

1.5 Further evidence is needed on:

- the impact of the addition of AI-derived software on a healthcare professional's ability to identify people for whom thrombolysis and thrombectomy is suitable (see section 3.7)
- test failure rate with causes of failures (see section 3.12)
- the impact of using the software on time to thrombolysis or thrombectomy (see section 3.9)
- the impact of using the software on how many people have thrombolysis or thrombectomy (see section 3.10).

## Why the committee made these recommendations

Stroke adversely affects quality of life for many people who survive it. Faster and greater access to treatment could improve clinical outcomes and so quality of life after stroke. AI-derived software used alongside healthcare professional interpretation of CT brain scan images could guide and speed up decision making in stroke, for example, decisions on thrombolysis and thrombectomy treatment. The AI-derived software uses fixed (or static) algorithms in clinical practice, with AI used to derive new versions of the algorithm.

Clinical evidence on the software is limited in quality. There is no evidence on their diagnostic accuracy when used alongside healthcare professional review that met review inclusion criteria. Some studies on 2 technologies (e-Stroke and RapidAI) in clinical practice suggest that people had faster or greater access to treatment after using the software, but it is unclear to what extent this is an effect of the software. A small increase in people having thrombectomies because of AI-derived software in the economic model would likely make the software cost effective.

The software is already widely used in the NHS, and should always be used with healthcare professional review. Because an important potential benefit of the technologies is improving image sharing between centres to help with time-sensitive decision making, centres should ensure that shared images are of sufficient quality to allow remote image review. Healthcare professionals should be cautious when changing their findings based on software results. Existing reporting protocols (or those used before the AI-derived software was adopted in a centre) should be maintained in centres using these technologies.

e-Stroke and RapidAI can be used in the NHS while further evidence is generated to help better determine their cost effectiveness. Other technologies that have no evidence on how they impact time or access to treatment should only be used in research.

## 2 The diagnostic tests

### Clinical need and practice

- 2.1 Stroke is a serious life-threatening medical condition that happens when blood supply to a part of the brain is severely compromised. More than 100,000 strokes happen in the UK every year. The average age for stroke across the UK varies, with a median age of 77 (interquartile range 67 to 85 years). A quarter of strokes happen in people of working age. Stroke is a leading cause of disability and one of the most common causes of early death in the UK.
- 2.2 Treatment of stroke depends on the cause, length of time the blood supply to brain has been compromised and the severity of the damage caused by the stroke. In ischaemic stroke (the most common type), when blood supply to the brain gets blocked by a clot, treatments aim to restore blood flow by dispersing the clot with an intravenous injection of a clot-busting drug (thrombolysis) and mechanically removing the clot (thrombectomy). Thrombolysis needs to be started within 4.5 hours of onset of stroke symptoms and thrombectomy should be offered as soon as possible to people who were last known to be well within the last 6 hours or considered as soon as possible for people who were last known to be well between the last 6 and 24 hours. In the less common type of stroke, intracerebral haemorrhage, when a weakened blood vessel in the brain bursts and blood leaks into soft brain tissues, these treatments would be harmful and should not be offered.
- 2.3 CT brain scans are used to help guide treatment choice. In people with a suspected acute stroke, a non-enhanced CT scan is used first to determine if the stroke is ischaemic so that thrombolysis can start. In people with confirmed ischaemic stroke, a CT angiography is then used to confirm the presence of a clot and to assess if it is a large vessel occlusion (a clot in a location where it could be removed by thrombectomy). When the impact of stroke is likely to be more severe

because the blood supply to the brain has been reduced for a longer time (between 6 and 24 hours), CT perfusion is used to assess if there is the potential to salvage brain tissue by doing a thrombectomy.

- 2.4 Software with artificial intelligence (AI)-derived algorithms can be used to analyse CT brain scan images from people with suspected acute stroke to detect and report irregularities. The result of this analysis is intended to support the scan review and reporting by a trained healthcare professional. By identifying, quantifying and highlighting stroke-related changes in the brain, the AI-derived algorithms may support clinical decisions about suitability of an appropriate time-sensitive treatment. Using the software in the radiology pathway may lead to quicker review of scans by a multi-site clinical team, improved decisions about treatment, expedited patient transfer, faster access to the correct treatment and improved patient outcomes. Some software has features that can prioritise the review of stroke CT scans.

## **The interventions**

### **Accipio (MaxQ AI)**

- 2.5 Accipio is a software with AI-derived algorithms. It detects and analyses intracranial haemorrhage on non-enhanced CT brain scans and large vessel occlusions on CT angiography brain scans. The cost of Accipio is not known. The company says that the Accipio suite is CE marked as a class IIa medical device. The technology has been discontinued since January 2022.

### **Aidoc (Aidoc)**

- 2.6 Aidoc is a software platform that includes AI-derived stroke-related algorithms Aidoc ICH and Aidoc LVO. Aidoc ICH detects intracranial haemorrhage on non-enhanced brain CT scans. Aidoc LVO detects large vessel occlusions on CT angiography brain scans. The platform also includes Aidoc Mobile, a communication component to help

communication between healthcare professionals in the stroke pathway. The manufacturer estimates that the licence fee for Aidoc for centres doing up to 45,000 CT brain scans per year is around £24,800 per year. For centres doing more than 45,000 CT brain scans per year, the licence fee is estimated as to be around £32,900 per year.

As well as the individual software package, Aidoc has partnered with icometrix (see section 2.13) for a 'stroke solution' (AIDOC ICH for detecting intracranial haemorrhage on non-enhanced CT brain scans, AIDOC LVO for detecting large vessel occlusion on CT angiography brain scan and icobrain ctp is used for CT perfusion brain scan analysis in ischaemic stroke). The company says that Aidoc software is a CE-marked (class I medical device) AI triage and notification platform.

### **BioMind (BioMind.ai)**

2.7 BioMind is a software that includes AI-derived algorithms for detecting, locating and assessing the severity of intracerebral haemorrhage on non-enhanced CT brain scans. The cost of BioMind is not known. The company says that BioMind is a CE-marked AI-derived platform (class unknown).

### **BrainScan CT (Brainscan.ai)**

2.8 BrainScan CT is a software with AI-derived algorithms. It detects and locates intracerebral haemorrhage and acute ischaemic stroke on non-enhanced CT brain scans. The cost of BrainScan CT not known. The company says that BrainScan CT is a CE-marked (class IIa medical device) AI-derived platform.

### **Cercare Perfusion (Cercare Medical)**

2.9 Cercare Perfusion is a software with AI-derived algorithms for CT and MRI brain scans. It uses data on tissue oxygenation to provide identification of suspected stroke lesions. It provides an overview of brain tissue status on

CT perfusion brain scans. The cost of Cercare Perfusion is not known. The company says that Cercare Perfusion is a CE-marked AI-derived platform (class unknown).

### **CINA head (Avicenna)**

2.10 CINA head is a software platform that includes AI-derived algorithms CINA-ICH, CINA-LVO and CINA-ASPECTS. CINA-ICH detects intracranial haemorrhage on non-enhanced CT brain scans and prioritises them on the radiologist's worklist. CINA-ASPECTS analyses non-enhanced CT brain scans to help characterise early ischaemic brain tissue injury. CINA-LVO detects and prioritises the review of large vessel occlusions on CT angiography brain scans. Assuming a minimum of 1,000 scans a year, cost ranges from around EUR 7.08 per scan for centres doing up to 5,000 CT brain scans a year to around EUR 5.27 per scan for centres doing over 20,000 CT brain scans a year. The company says that the technology is CE marked (class I medical device).

### **CT Perfusion 4D (GE Healthcare)**

2.11 CT Perfusion 4D is a software with AI-derived algorithms. It provides an overview of brain tissue status on CT perfusion brain scans. The cost of CT Perfusion 4D is not known. The company says that CT Perfusion 4D is a CE-marked medical device (class unknown).

### **e-Stroke (Brainomix)**

2.12 e-Stroke is a software platform with AI-derived algorithms e-ASPECTS, e-CTA and e-CTP. e-ASPECTS detects acute ischaemic stroke on non-enhanced CT brain scans. e-CTA detects and locates large vessel occlusions on CT angiography brain scans. e-CTP analyses CT perfusion brain scans to provide information about brain tissue status. The cost of the software licence for a comprehensive stroke centre is around £30,000 per year and for an acute stroke centre around £15,000 per year. The company says that e-Stroke is CE marked (class IIa medical device).



### **icobrain ct (icometrix)**

2.13 icobrain ct is a software platform that includes AI-derived algorithm icobrain ctp for analysing CT perfusion brain scans to determine the presence of potentially salvageable brain tissue in ischaemic stroke. The cost of icobrain ct ranges from around £20,000 to around £45,000 per year depending on volume. As well as the individual software package, icometrix has partnered with Aidoc on a 'stroke solution' (see section 2.6). The company says that the icobrain ct is CE marked (class I medical device).

### **Neuro Solution (Nanox.AI)**

2.14 Neuro Solution is a software with AI-derived algorithms. It detects intracranial haemorrhage on non-enhanced CT brain scans. Until November 2021, this technology was known as Zebra-Med (Zebra Medical Vision). The cost of Neuro Solution is not known. The company says that Neuro Solution is CE marked (class I medical device).

### **qER (Qure.ai)**

2.15 qER is a software with AI-derived algorithms for detecting intracerebral haemorrhage and areas of brain tissue death (infarct) on non-enhanced CT brain scans. The cost of qER is not known. The company says that qER is a CE-marked triage and notification medical device (class unknown).

### **RapidAI (Ischemaview)**

2.16 RapidAI is a software platform that includes AI-derived stroke-related algorithms Rapid ICH, Rapid ASPECTS, Rapid CTA, Rapid LVO and Rapid CTP for analysing CT brain scans. Rapid ICH detects intracerebral haemorrhage on non-enhanced CT brain scans. RAPID ASPECTS assists in assessing the extent of disease on non-enhanced CT scans from people who have ischaemic stroke caused by a large vessel

occlusion. Rapid CTA and RAPID LVO detect and locate large vessel occlusions on CT angiography brain scans. Rapid CTP analyses CT perfusion brain scans to give information about salvageable brain tissue. RapidAI is provided on an annual subscription fee basis. An average cost per centre is around £20,000 per year. The company says that RapidAI is CE marked (class IIa medical device).

### **Viz (Viz.ai)**

2.17 Viz is a software platform that includes AI-derived stroke-related algorithms Viz ICH, Viz LVO and Viz CTP. Viz ICH detects intracranial haemorrhage on non-enhanced CT brain scans. Viz LVO detects large vessel occlusions on CT angiography brain scans. Viz CTP analyses CT perfusion brain scans to provide information about salvageable brain tissue. The cost of the software for a comprehensive stroke centre is in the region of around £30,000 per year and for a primary stroke centre around £15,000 per year. The company says that Viz is CE-marked as a class I medical device.

## **The comparator**

### **CT brain scan review by a healthcare professional without assistance from AI software**

2.18 Non-enhanced CT brain scans may be reviewed by a radiologist, specialist radiologist, radiographer, stroke physician or emergency medicine physician, depending on availability of staff. CT angiography and CT perfusion brain scans are more likely to be reviewed by a radiologist, neuroradiologist or an interventional neuroradiologist, who specialise in interpretation of these scans.

## **3 Committee discussion**

The [diagnostics advisory committee](#) looked at evidence for AI software across 3 indications. Evidence was considered from several sources, including a diagnostics

assessment report and an overview of that report. Full details are in the [project documents for this guidance](#).

### **Quality of life is important to people who survive stroke**

3.1 The patient expert explained that stroke adversely affects quality of life for many people who survive it. In addition to physical disability, long-term effects include fatigue, cognitive impairment, difficulty with language or speech (aphasia), poor mental health and emotional lability (exaggerated emotions that can be difficult to control). Around 50% of people who survive a stroke at a working age never return to work. Stroke often also substantially affects the lives of relatives and friends. The patient expert advised that it is important to understand the effect of AI-derived software when used alongside clinician interpretation of CT brain images on clinical outcomes and the related quality of life after stroke. The committee recognised that quality of life is important to people who survive stroke. Experts further highlighted the benefits of greater access to treatment for people who present with stroke after a longer time from symptom onset, such as people with wake-up strokes (which happen while a person is sleeping so it is not clear when the stroke occurred).

### **The AI-derived software do not automatically adapt and improve if the software is employed in the NHS**

3.2 The committee discussed the nature of the algorithms in AI software and if the software could learn from the CT scan data in the setting it was used in. The manufacturers said that data from scans the software is used on in clinical practice is not used to further develop algorithms in the software. Instead, the algorithms in the software are developed using CT scans held by the company or accessed through research studies. Then regulatory approval is sought before an updated static algorithm is released for use in clinical practice. The committee recognised that all AI software in clinical settings use fixed algorithms and cannot adapt and improve in real time using data from the clinical practice setting in which they are used.

## **The AI-derived software are already widely used in the NHS and further changes to the stroke care pathway are ongoing**

3.3 Since the initial committee meetings, the [Getting It Right First Time \(GIRFT\) national report for stroke](#) has published. This includes a recommendation to increase regional availability of AI support tools and training. AI decision support has been implemented at 99 of 107 stroke units in England. All other identified centres are actively working on plans to go live before the end of 2023. Experts also highlighted that stroke care is a complex care pathway and changes to practice are ongoing, which makes it difficult to measure the impact of any single change on the pathway.

## **The image sharing function of the technologies could be a key driver of potential benefit**

3.4 Several stakeholders and experts highlighted that the rapid image transfer function of the technologies is an important feature because it allows stroke physicians and interventional neuroradiologists working at comprehensive stroke centres to quickly review CT scans taken at other sites. This can help support decisions about whether a transfer is needed for a person presenting at a centre that cannot do thrombectomies and has less experience in interpreting CT images. Several stakeholders highlighted issues with current image sharing systems in England, and that rapid image transfer is currently being provided by the AI-derived software. They further stated that the NHS imaging system would not be able to replicate this function, or that it would need substantial cost to upgrade existing systems.

## **The software should only be used alongside healthcare professional decision making**

3.5 Experts highlighted the importance of AI-derived software only being used to assist CT brain scan review by a healthcare professional. Also, they

advised that healthcare professionals should be cautious when changing their findings based on software results.

## Clinical effectiveness

### **No published evidence was found for 9 of the 13 technologies in the assessment**

3.6 The committee considered the available evidence for each technology and indication. It noted that the external assessment group's (EAG's) review found no published evidence that met review inclusion criteria for Accipio, Aidoc, Biomind, Brainscan CT, Cercare (Perfusion), CT Perfusion 4D, icobrain ct, Neuro Solution or qER for the indications in the assessment. An updated review done by the EAG for the third committee meeting in August 2023 also found no further studies that met the review inclusion criteria for these technologies. So, the committee was unable to consider these technologies further as part of its discussions and recommended more research. Only 1 study was identified that assessed the CINA head software, and this only reported accuracy data for the technology used as a standalone intervention and not alongside clinician interpretation (as it is intended to be used). There were more studies, including some reporting outcomes such as the impact of software on treatment (see sections 3.9 and 3.10) for the remaining 3 technologies (e-Stroke, RapidAI and Viz).

### **Impact of the AI-derived software's analytical functions on a healthcare professional's ability to identify people for treatment is uncertain**

3.7 The EAG's initial review found 15 diagnostic accuracy studies but these all evaluated the performance of the AI-derived software as a standalone intervention and not alongside clinician interpretation (as it is intended to be used). Also, the risk of bias because of patient selection in many studies was high, particularly when they used a case-control study design, or unclear because of inadequate reporting. The reference standard used

in the studies ranged from review by a single clinician to a panel of clinicians with or without access to clinical data alongside images. But, it was often unclear whether these clinicians were blinded to the output from the AI software and difficult to determine if they were likely to correctly classify the target condition because their experience was not clearly reported. So, because the studies were not generalisable to how the technologies would be used in practice, limited conclusions could be made about the accuracy of the technologies. Also, the committee noted that none of the studies separately reported accuracy for people aged over 80 with cerebrovascular disease, when the interpretation of scans is often more challenging. An updated review done by the EAG for the third committee meeting in August 2023 also found no studies that compared the performance of a healthcare professional reviewing scans, with and without use of the software, that met review inclusion criteria. The committee concluded that the impact of the analytical functions of the AI-derived software on a healthcare professional's ability to identify people for whom treatments like thrombolysis and thrombectomy are suitable is uncertain.

### **It is difficult to draw conclusions on the reported comparative accuracy data**

3.8 The committee recognised that 1 study (Seker et al. 2020), relevant to guiding mechanical thrombectomy decisions for people with an ischaemic stroke using CT angiography, reported some comparative accuracy data. It reported data both on the accuracy of e-CTA software (Brainomix) alone and for scan reviews done by clinicians of varying experience alone compared with a common reference standard. This was an experienced neuroradiologist who had access to both imaging and clinical data. This is important because the usefulness of AI-software-assisted scan review may vary between centres with differing levels of stroke specialism, and between different types of clinicians (for example, doctors in hospital emergency departments, stroke specialists, radiologists and

neuroradiologists). But the committee noted that it is difficult to draw conclusions from the study on how the software would perform when used alongside clinician review because it did not provide information on whether clinicians and the software missed the same or different cases.

### **It is uncertain whether using AI-derived software to help guide treatment decisions in stroke leads to faster access to treatment**

3.9 In the EAG's initial review there were 7 observational studies that compared time to treatment before and after implementing AI-derived software in clinical practice, assessing e-Stroke, RapidAI and Viz. Most of the studies suggested that time to treatment for people who had thrombectomy or thrombolysis had reduced after implementing the software. The EAG reported that there was a high risk of bias in these studies because of the limited information they included. The studies were all retrospective, study populations and stroke care settings were not clearly described, the point in the care pathway when the software was used and by whom was often unclear. Also, it was unclear if the before and after populations had similar characteristics, and whether adding the software was the only change to the care pathway. Because only people with a positive scan result were included in the studies, it is unclear whether people with a false negative result would experience a delay in treatment. An updated review done by the EAG for the third committee meeting in August 2023 identified further studies (using e-Stroke and Viz) that assessed the impact of test use on time to treatment. Studies done in centres that needed to transfer people to have a thrombectomy reported reductions in time to thrombectomy, although the EAG highlighted concern about study quality, including the same issues highlighted in the original report. The EAG noted that the UK implementation report (an interim report of an ongoing study by an Academic Health Science Network) did not collect data on the effects on time to treatment for centres that needed to transfer people for thrombectomy. The committee recalled that the image sharing function of the technologies may be a

driving factor in improving time to treatment (see section 3.4) but that other changes to stroke care may also be factors (see section 3.3). It concluded that it is uncertain whether using AI-derived software to help guide treatment decisions in stroke leads to faster access to thrombolysis or thrombectomy.

### **It is uncertain whether using AI-derived software leads to increases in the number of people having a thrombectomy**

3.10 An updated review done by the EAG for the third committee meeting in August 2023 identified 2 studies that assessed impact of e-Stroke on the numbers of thrombectomies done. Both showed increases in the proportion of people who had a stroke having a thrombectomy after implementation of e-Stroke. The EAG stated that neither study provided sufficient information to establish that populations were comparable before and after the implementation of the software, and that this was the only change to the care pathway. It further highlighted that 1 study (an interim report provided by an Academic Health Science Network of an ongoing UK implementation study) was done during the COVID-19 pandemic and it was unclear how rates of CTA scanning and thrombectomy may have been affected by this. The AHSN work reported that the proportion of people having thrombectomy who had presented more than 6 hours after the onset of symptoms increased significantly after implementation of e-Stroke. The committee noted that study authors suggested that this may be because e-Stroke's rapid image sharing functionality allows quicker access to thrombectomy specialists, and helps their decision making. This could make it less likely they will decline a transfer for a patient who was last well more than 6 hours ago if they can review their CT scans. Experts highlighted that a NICE recommendation about offering thrombectomy for people between 6 hours and 24 hours of symptom onset (including wake-up strokes) included the need to establish the potential to salvage brain tissue, as shown by imaging such as CT perfusion (CTP) showing limited infarct core volume. They further



highlighted that RapidAI software had been used to analyse CTP scans in identifying people eligible for studies underpinning this recommendation. Experts disagreed about the extent that analysis from the AI-derived software of CTP scans adds to information routinely available from CT scanners and would be needed in practice for thrombectomy to be available for this later presenting group. Experts further highlighted that the National Optimal Stroke Imaging Pathway includes doing CTP scans at the same sitting as plain CT head and CTA, and that CTP scans are increasingly available at acute stroke centres. But such centres have less expertise to interpret these scans and would benefit from rapid support from comprehensive stroke centre teams. The EAG highlighted that there is little data on technology performance when assessing CTP scans. In the AHSN work, this function of the e-Stroke package was only available in comprehensive stroke centres and the extent it was used was uncertain.

### **It is unclear whether using AI-derived software to help guide treatment decisions in stroke leads to better clinical outcomes**

3.11 The committee noted that the studies that compared time to treatment before and after implementing AI-derived software provided limited information on how it affected clinical outcomes. In particular, there was no information on clinical outcomes when the software was used for guiding thrombolysis treatment decisions for people with suspected acute stroke using a non-enhanced CT scan. Six studies, in which the software was used for guiding mechanical thrombectomy using CT angiography or CT perfusion brain scans, reported on the proportion of people who were functionally more independent (with modified Rankin Scale [mRS] score 2 or less), length of hospital stay, mean 90-day mRS score and rate of complications and death during hospital stay after software implementation. The committee noted that the results from these studies were conflicting, with some reporting a positive and others a negative impact. The EAG advised that the studies were unlikely to have been

appropriately set up to adequately capture any differences in clinical outcomes. So, the reported data is unlikely to show the true effects of implementing the technologies. The EAG also highlighted that the evidence described outcomes only for people who had a thrombectomy. Clinical experts explained that while using AI-derived software could help improve outcomes for people who are offered treatment if they have it sooner, it could also worsen outcomes for people who were not offered treatment or who had incorrect treatment if their diagnosis was missed because of the influence of the software on clinical decision making. An updated review done by the EAG for the third committee meeting in August 2023 identified further studies that assessed the software's impact on mRS and mortality. Results were mixed, in terms of indicating possible benefit or detriment, and the 1 study that had longer follow up (6 months) had very high levels of missing data.

### **More information about the reliability of AI-derived software to help guide treatment decisions in stroke is needed**

3.12 Only 1 published study (Kauw et al. 2020) reported on the technical failure rate of AI-derived software. This study reported that the software failed to process CT perfusion brain scan data and return results to assist the review of 20 of the 176 scans (11%) included in the analysis. Causes for failures were severe motion, streak artifact and poor arrival of contrast. The clinical experts advised that it is possible that the failure in clinical practice may be even higher. The patient expert raised concerns that technical failures could result delays in diagnosis and access to time-sensitive treatments. The committee concluded that the reliability of AI-derived software to help guide treatment decisions in stroke in clinical practice is not clear. It recommended further research to measure technical failure rates of AI-derived software used to help guide treatment decisions in stroke. Information about the reasons for test failures should be recorded.

## Cost effectiveness

### **There was not enough clinical evidence to evaluate the cost effectiveness of AI-derived software in 2 of the 3 assessed indications**

3.13 Evidence on using AI-derived software for guiding thrombolysis treatment decisions for people with suspected acute stroke using a non-enhanced CT scan, and mechanical thrombectomy treatment decisions for people with ischaemic stroke using CT perfusion after a CT angiography brain scan, was very limited. In particular, there was no evidence on the diagnostic accuracy of the technologies when used alongside clinician interpretation or how they might perform relative to clinician alone in either indication (see section 3.7). No clinical outcomes were reported for using AI-derived software to guide thrombolysis treatment decisions for people with suspected acute stroke (see section 3.11). So, the EAG did not build health economic models to evaluate the cost effectiveness of the software in these 2 indications. The committee concluded that it would be useful to understand the cost effectiveness of the AI-derived software in these indications but accepted that there is currently not enough data available to inform modelling.

### **Accuracy estimates in the model for using AI-derived software in thrombectomy decisions may not reflect the accuracy seen in clinical practice**

3.14 The EAG explained that because there was more data related to using AI-derived software for guiding mechanical thrombectomy decisions for people with ischaemic stroke using CT angiography than for the other 2 indications, it could build an exploratory economic model for this. Because no diagnostic accuracy data was available for using the technologies as intended (see section 3.7), the EAG elicited accuracy estimates for the model from clinical experts. These estimates were sought for a hypothetical average AI-derived software when used alongside clinician interpretation and for the comparator in the model, clinician interpretation

alone. The committee noted that it is challenging for people to estimate something like accuracy that they cannot directly see. The committee concluded that while expert elicitation is an appropriate method to obtain model inputs when data is scarce, it is uncertain if the accuracy estimates in the model reflect the accuracy of the AI-derived software that would be seen in clinical practice.

### **Health-related quality of life, given the exploratory nature of the model, is adequately captured**

3.15 The committee considered whether the model captured the effect that having a stroke has on people's quality of life. The EAG explained that the utility values for health-related quality of life used in the base case were linked to the mRS health states, from Rivero-Arias et al. (2010). This study used mRS and EQ-5D-3L information that was collected from people with stroke or transient ischaemic attack who took part in the Oxford Vascular Study (OXVASC) in the UK. The committee concluded that the health-related quality of life, given the exploratory nature of the model, was adequately captured. But, it recalled how important quality of life is to people who survive stroke (see section 3.1) and that further research on the impact of the AI-derived software on quality of life after stroke using EQ-5D would be beneficial (see section 3.22). It further considered that better understanding of health-related quality of life after stroke, for example aspects such as cognitive or emotional difficulties and fatigue that are important to people who survive stroke, would be helpful. It noted that the [research priorities from the Stroke Priority Setting Partnership](#) (led by the Stroke Association with the James Lind Alliance) and [research needs from the National Stroke Programme](#) (NHS Accelerated Access Collaborative) include understanding and managing emotional and psychological effects of stroke that may be less visible.

### **Small increases in the number of thrombectomies done in the EAG's model are enough for the test to be cost effective**

3.16 In the EAG's model, the increase in the number of people having thrombectomies if the software is used does not need to be large for it to be cost effective. This includes if specificity is worse with the addition of AI-derived software to healthcare professional review (that is, there are more false positive results), based on a study identified by the EAG (Andralojc et al. 2023). The committee noted at the third committee meeting there was now some evidence that the technologies increase the number of people having thrombectomy but recalled that this was uncertain (see section 3.10). The EAG commented that if using the technologies means different people are identified for thrombectomy this could impact on how effective it is. That is, the relative effectiveness of thrombectomy in the additional people that get it, would be equal to the effectiveness of thrombectomy for people that already get it (without the addition of the software). Experts acknowledged this but commented that studies routinely show benefits from increased thrombectomy use.

### **Further evidence on test performance is important alongside data on impact on time to, or occurrence of, treatment**

3.17 The EAG's model included additional cost impact for false positive results (for ambulance transfer and time on a stroke unit). Experts agreed that such cases would be detected before having an unnecessary thrombectomy, but also highlighted that time spent reviewing these scans increases the work pressures on healthcare professionals. The EAG highlighted uncertainty in how software use will affect numbers of false positive results, and that this will impact cost effectiveness. The impact of false negative cases (people with eligible large vessel occlusions [LVOs] who do not have thrombectomy) was included in the model. Experts commented that such cases would likely eventually be detected, for example if ongoing symptoms suggested an issue. But this could mean delays to treatment (increased time to thrombectomy), which may impact

eligibility for treatment (numbers having thrombectomy). Data on the impact of AI-derived software on the performance of healthcare professionals to identify people for whom thrombolysis or thrombectomy is suitable would help assess this. The EAG stated that data on time to treatment alone is insufficient to assess cost effectiveness, because the technologies can affect who has treatment, as well as how fast this is done. The committee agreed that further data on the impact of the addition of AI-derived software on the performance of healthcare professionals to identify people eligible for thrombolysis or thrombectomy is important to generate, alongside data on time to treatment and how many people have treatment.

**Cost effectiveness of AI-derived software is uncertain, but it is plausible they are cost effective**

3.18 The committee considered whether it was possible to determine the cost effectiveness of AI-derived software for guiding mechanical thrombectomy decisions for people with an ischaemic stroke using CT angiography from the EAG's model. It recalled that the model was built using diagnostic accuracy estimates elicited from experts (see section 3.14). This meant that the model did not reflect any of the individual software but modelled a hypothetical average AI-derived software. The committee noted that in reality, the different technologies may perform differently from this modelled average technology but acknowledged that there was no evidence on their performance when used as intended (see section 3.7). The committee concluded that cost effectiveness of AI-derived software for guiding mechanical thrombectomy decisions for people with an ischaemic stroke using CT angiography is uncertain. But it recalled that small increases in the number of people having thrombectomy caused by AI-derived software use in the EAG's model probably would be enough for the technologies to be cost effective, and there was at least some data to suggest this may occur (see sections 3.10 and 3.16). The committee concluded that while cost effectiveness is uncertain, the technologies

could potentially be cost effective. But further data is needed to confirm this.

### **Some technologies can be used in the NHS while further evidence is generated**

3.19 Further evidence is needed to better estimate the cost effectiveness of the technologies, so the committee did not recommend routine use in the NHS. It recalled that the technologies are already widely in use (see section 3.3). Experts at the third committee meeting highlighted benefits they are seeing from the technologies, and the committee considered that it was not appropriate to recommend stopping use. There was more data for 3 of the technologies (e-Stroke, RapidAI and Viz), including impact on time to treatment and how many people get treatment. The committee recalled there was uncertainty about how using the technologies impacted a healthcare professional's ability to identify people for thrombolysis or thrombectomy. But it concluded that, as long as these technologies were only used alongside clinician decision making (see section 3.5), they could be used in the NHS while further evidence is generated. The remaining technologies should only be in research. Because the image sharing function of the technologies is likely to be a large part of their value (see section 3.4), it is important that centres using the technologies ensure that the images that are shared can be remotely reviewed to help with decision making by healthcare professionals at a different site.

## **Research considerations**

### **Ongoing use in the NHS and data collection**

3.20 The EAG commented that the widespread use of these technologies in the NHS in England limits the potential for commissioning some types of high-quality primary research. It suggested that retrospective re-analyses of stored radiology reports, collected during the implementation of these technologies, may have potential to inform estimates of the accuracy of AI-derived software technologies in combination with clinical judgement.

Committee members also highlighted the potential use of retrospective data to help answer outstanding evidence issues, potentially making use of data routinely collected in the NHS, such as the Sentinel Stroke National Audit Programme (SSNAP). The committee were aware that data collection as part of the AHSN's work is ongoing and full analyses are yet to be completed, which may provide useful further evidence.

**Data should be collected from a representative population and reflect how the software may add benefit to the NHS**

3.21 The EAG highlighted that it was not clear from available data whether the populations assessed were representative of the UK stroke population in terms of age, comorbidities and family background. The committee emphasised that further data collection should be done in populations that represent people having treatment in the NHS, including assessing factors that could impact on technology effectiveness such as age, sex, ethnicity, socioeconomic status and comorbidities. The committee also recalled that people presenting longer after onset of symptoms may be in particular need of greater access to treatment (see section 3.1) and encouraged this group to be included in further evidence generation. The committee recalled that the most benefit of the technologies may be for people presenting to sites that cannot do thrombectomies (and so a transfer is needed for this). Data collection should reflect this, for example by assessing the impact of software on the performance of clinicians working in acute rather than comprehensive stroke centres. Ideally studies should describe the level of experience of the clinicians interpreting the CT brain scans and include both people who had and did not have a particular type of treatment. Studies should also consider reporting data separately in subgroups when using the technologies may be particularly useful or less effective (for example in older people, particularly those aged over 80, with small vessel disease and calcification of the cerebrovasculature, people with an unknown time of stroke onset or wake-up stroke and people who have had a previous stroke).



## Data on clinical outcomes would be beneficial

3.22 Both the EAG's model and previous assessments of tests related to stroke done for NICE diagnostic assessment programme guidance have used a linked evidence approach to estimate the impact of test use on clinical outcomes like mRS, rather than needing direct evidence of this. The committee also noted that in the EAG's model only very small improvements in mRS are needed for the technologies to be cost effective. Direct data on the impact of AI-derived technology use on clinical outcomes like mRS or mortality would be beneficial for this assessment but not essential.

## 4 Implementation

NICE intends to develop tools, in association with relevant stakeholders, to help organisations put this guidance into practice.

In addition, NICE will support this guidance through a range of activities to promote the recommendations for further research. The research proposed will be considered by the NICE Medical Technologies Evaluation Programme research facilitation team for developing specific research study protocols as appropriate. NICE will also incorporate the research recommendations in section 1.5 into its [guidance research recommendations database](#) and highlight these recommendations to public research bodies.

## 5 Review

NICE reviews the evidence 3 years after publication to ensure that any relevant new evidence is identified. However, NICE may review and update the guidance at any time if significant new evidence becomes available.

Brian Shine

Chair, diagnostics advisory committee

September 2023

## 6        **Diagnostics advisory committee members and NICE project team**

### **Committee members**

This topic was considered by the [diagnostics advisory committee](#), which is a standing advisory committee of NICE.

Committee members are asked to declare any interests in the test to be assessed. If it is considered there is a conflict of interest, the member is excluded from participating further in that assessment.

The [minutes of each committee meeting](#), which include the names of the members who attended and their declarations of interests, are posted on the NICE website.

Additional specialist committee members took part in the discussions for this topic:

### **Specialist committee members**

#### **Sotirios Bisdas**

Consultant neuroradiologist, University College London Hospitals NHS Trust, and associate professor of neuroradiology, Queen Square Institute of Neurology, University College London

#### **Margaret Cheng**

Lay member

#### **Tim England**

Professor of stroke medicine, University of Nottingham, and honorary consultant stroke physician, University Hospitals of Derby and Burton

#### **Carole Gavin**

Consultant in emergency medicine, Salford Royal NHS Foundation Trust

#### **Nigel Hoggard**

Professor of neuroradiology and honorary consultant neuroradiologist, University of Sheffield and Sheffield Teaching Hospitals NHS Foundation Trust

Diagnostics consultation document – Artificial intelligence (AI) software to help clinical decision making in stroke  
Page 26 of 28

Issue date: September 2023

© NICE 2023. All rights reserved. Subject to [Notice of rights](#).

**Grant Mair**

Senior clinical lecturer and honorary consultant neuroradiologist, University of Edinburgh and NHS Lothian

**Kiruba Nagaratnam (until July 2023)**

Consultant stroke physician, Royal Berkshire NHS Foundation Trust

**Jonathan Shapey**

Senior clinical lecturer and honorary consultant neurosurgeon, King's College London and King's College Hospital

**Li Su**

Chair of neuroimaging, University of Sheffield, and senior research fellow, University of Cambridge

**David Werring**

Professor of clinical neurology and consultant neurologist, University College London and University College London Hospitals NHS Foundation Trust

**Clinical experts**

**Kausik Chatterjee**

Consultant Stroke Physician, Countess of Chester Hospital

**Martin James**

Consultant Stroke Physician and Honorary Professor, Royal Devon and Exeter NHS Foundation Trust and University of Exeter Medical School

**Robert Simister**

Consultant Neurologist and Stroke Physician, University College London Hospitals

**NICE project team**

Each diagnostics assessment is assigned to a team consisting of a technical analyst (who acts as the topic lead), a technical adviser and a project manager.

**Anam Khan**

Topic lead

**Thomas Walker**

Technical adviser

**Donna Barnes**

Project manager

ISBN: