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

Section A: External Assessment Report - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
British and Irish Association of Stroke Physicians (BIASP)	1			We recognise the poor evidence base for the accuracy of AI algorithms for acute stroke CT analysis which is well assessed by the updated systematic review.	No response required.
British and Irish Association of Stroke Physicians (BIASP)	2			We would like to particularly highlight that the current use of AI in stroke is not primarily diagnostic but also essential to facilitate remote imaging review, early rapid clinician notification and expedite decision making and image transfer between centres.	As noted in the addendum, whilst this application of the technology is acknowledged, assessment of image sharing technologies was not in the scope for this assessment. Remote image sharing is a different application to AI- derived software image interpretation and an assessment of image sharing technologies may have needed to include additional non-AI comparators.
British and Irish Association of Stroke Physicians (BIASP)	3			Unfortunately, the inevitable conclusions of the report on the basis of the current evidence conflict with the roll-out of software across a very large proportion of UK/Ireland stroke services, which was accelerated by the results of the reperfusion trials using the technology, the effect of the COVID-19 pandemic on stroke services, and the desire by GIRFT and others for rapid adoption. We request that NICE does not recommend stopping use of the technology in stroke services. The practical use of AI technology in stroke is facilitating huge advances in the delivery of extended window thrombolysis and maximising the access to timely thrombectomy which is currently only delivered in few specialist centres.	Statement of opinion, not supported by evidence, no response required.
British and Irish	4			We would support NICE describing the uncertainties and caution that is required in using the technology (based on	No response required.

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
Association of Stroke Physicians (BIASP)				the existing diagnostic accuracy studies) and that expert radiological support for imaging interpretation should be available in addition to AI algorithms for acute stroke imaging analysis for clinicians.	
British and Irish Association of Stroke Physicians (BIASP)	5			As this is such a rapidly emerging field - we would like to request that NICE repeats this appraisal as soon as the results of the cluster randomised trial(s) of AI algorithms for acute stroke CT analysis are known.	Decision for NICE.
British and Irish Association of Stroke Physicians (BIASP)	6			We feel an important research recommendation of the NICE guidance would be to address the remaining uncertainties about this in a UK context. We appreciate the limitations in this as AI is being used routinely already.	Decision for NICE.
British and Irish Association of Stroke Physicians (BIASP)	7			Please consider, the cost effectiveness analysis and impact of AI in thrombolysis and thrombectomy decisions can be dependent on differing time points. E.g. using AI CT perfusion to make lysis decisions within a standard treatment window is unnecessary so would not be cost effective. Use after the standard time windows for thrombolysis and thrombectomy should have an entirely different assessment as the treatment can not be offered without them.	The differing time windows for CTA and CTP were reflected in the scope and inclusion criteria, specified for the original DAR and carried forward to the current addendum. Unfortunately the lack of evidence about the effects of AI-derived software interpretation of CTP prevented its inclusion in the cost-effectiveness modelling undertaken for the original DAR and no new evidence was identified during the preparation of the addendum.
British and Irish Association of Stroke Physicians (BIASP)	8			We would like to recognise that the evidence base for AI reporting would need a much stronger basis if it were used solely without parallel radiology support (which is generally the current practice). AI is used as a supportive tool for clinicians and radiologists along with clinical judgement as per many other routinely available diagnostic methods.	It is unclear what the stakeholder means by ‘stronger basis’ in this context. There is a substantial body of evidence to inform the accuracy of these technologies alone. However, as noted in both the original DAR and the addendum, this cannot provide any information about their effects on clinicians’ judgement (combined accuracy).

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
British and Irish Association of Stroke Physicians (BIASP)	9			For a comprehensive evaluation we would recommend using additional parameters such as reduction in door to needle time (DTN)/door in door out (DIDO) time plus facilitation of out of hours thrombectomy and image transfer times.	The outcomes included in both the addendum and the original DAR were those specified, *a priori*, in the scope and in the agreed protocol for this topic. Door to needle time data were included, where reported. Out of hours thrombectomy rates were not explicitly listed as an outcome, but would have been extracted (with thrombectomy rates) had data been reported. DIDO time and image transfer time were not among the outcome measures specified for this assessment and the reasons for this are discussed in the addendum. Inclusion of these outcome measures would require a new assessment with a new scope/protocol.
British and Irish Association of Stroke Physicians (BIASP)	10			We would like to highlight that as the technologies are being widely used across the UK, that a more pragmatic approach to assessment would be analysing which platforms are best/most cost effective and ways of implementing into care pathways to maximise benefit to patients and systems.	The discussion section of the addendum includes a detailed discussion of the deficiencies of the current evidence base with respect to a range of possible approaches to considering cost-effectiveness. We would like to note that even a 'pragmatic' approach requires evidence/data.
British and Irish Association of Stroke Physicians (BIASP)	11			The cost-effective analysis does not take into account the substantial cost needed to upgrade current NHS image transfer systems to facilitate the emerging and geographically challenging landscape of stroke medicine. This essential rapid image transfer is currently being provided by this AI technology.	As noted in previous response, remote image sharing is a different application to AI- derived software image interpretation and an assessment of image sharing technologies may have needed to include additional non-AI comparators.
Oxford Academic Health Science Network	12			The number of people using AI decision support software for suspected acute stroke and using it regularly is increasing, despite there being questions about the accuracy of the technology. This consistent use confirms that the level of accuracy is acceptable amongst clinicians for rapid triage and decision support.	Widespread use and clinical acceptance are not evidence that an intervention is safe and effective/cost-effective. Also, if the technology was already known to be accurate enough then why was there a need to examine "Effectiveness", "Accuracy" and "Safety", which

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
				The purpose of the unpublished evaluation was not to validate the AI algorithm, but to determine how the technology is used and valued in a real-world situation. The report, shared with NICE and included in this review, is one in a series of evaluation reports, covering the 3 year evaluation period starting in March 2021. Subsequent reports will include the publication of a value assessment (September 2023) and a final evaluation report (March 2024). Copies of previous reports, including details of the evaluation approach and methodology are available on request.	are explicit themes stated in the report of the evaluation. The stakeholder's statement that their report was not intended as a research study was acknowledged in the addendum. The addendum includes any potentially relevant evidence that meets the *a priori* specified inclusion criteria for this assessment. The stakeholder was asked, at the start of this up-date, to provide details of their methods, in particular, full details of survey methods and all outcomes and methods of analysis/analysis plan for quantitative analyse; these were not provided.
Oxford Academic Health Science Network	13			Clarification questions on the unpublished evaluation report. Given that the authors of the report were approached to provide validation of several points within the report, it is not clear why some questions remained unanswered and unclear within the review. It would be helpful if the responses were provided in full alongside the additional questions.	The response provided did not address all questions raised and some responses did not provide the information/clarification required. A copy of the full set of questions submitted to the stakeholder was included in the addendum (Appendix 4). Whether or not to provide a copy of the full response is a decision for NICE.
Oxford Academic Health Science Network	14	93	Appendix 4	Clarification questions on the unpublished evaluation report. Given that the authors of the report were approached to provide validation of several points within the report, it is not clear why some questions remained unanswered and unclear within the review. It would be helpful if the responses were provided in full alongside the additional questions.	Please see previous response.

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
Oxford Academic Health Science Network	15	7	1	Research Q1: Does AI-derived software assisted review guide thrombolysis treatment? (abridged) Most strokes are diagnosed Ischaemic or Haemorrhagic on a plain CT scan and are identifiable from one another without decision support software. The impact of AI in this diagnosis and treatment is limited. This reflects a lack of understanding in the delivery of thrombolysis at a service level. The decision to administer thrombolysis is largely down to clinical practice, which considers multiple factors in the patient's history such as co-morbidities, medication, last known well. It is also common practice for patients to start IV thrombolysis whilst still on the CT scanner. CTP scan and interpretation with AI would be a better marker for the impact of AI as this helps to determine how much brain tissue is salvageable when time of stroke or onset of symptoms is unknown or outside of the traditional treatment time window. However, CTP scanning is not available at all stroke units in the NHS and we therefore have limited data at this point to analyse the impact of this.	The research questions are those that were defined by the scoping process for the original DAR, which included input from experienced clinicians across relevant disciplines. Data about the effects of AI-derived software technologies on the interpretation of CTP and subsequent time-to-treatment, treatment rate sand clinical outcomes were very limited (original DAR) and the addendum did not identify any new data (including in the report submitted by the stakeholder).
Oxford Academic Health Science Network	16	17	Table 2, and associated text.	Your review has chosen to exclude studies as they do not provide information about how the 'addition of AI-derived software technology might affect the performance of human readers and the clinical decision points specified in the three research questions.' It seems amiss to exclude these sources when a number have been carried out in NHS stroke units and give valuable insight into the impact of AI.	The inclusion or exclusion of studies is not an arbitrary choice made by the EAG, it is based on inclusion criteria which were specified, *a priori*, for this assessment and which were arrived at following a full scoping process which included input from clinical experts and stakeholders. It is not clear how the stakeholder considers that studies, which do not address consider how AI effects image interpretation at a clinical decision

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
					point can 'give valuable insight into the impact of AI.'
Oxford Academic Health Science Network	17	18	3.2.1	2nd paragraph. Glasgow has been included as a site that has CTP in the subsequent submission. Southend Hospital, although having access to e-CTA, was not reporting numbers of scans due to an error (this site was using a research version of the software). This accounts for the differences in numbers reported in the review.	We are not clear what the stakeholder is clarifying with this comment; 'Glasgow' is not mentioned in the report submitted by the stakeholder or in their subsequent response to questions from the EAG. The report does include data for Southend Hospital, but does not mention that this site was using a research version of the software.
Oxford Academic Health Science Network	18	22	3.2.2	1st paragraph. The authors of the unpublished report stated that it is not intended to be an academic, peer reviewed or scientific publication. However, the reviewers have assessed as such whilst also not requested further information on the evaluation methodological approach.	Please see previous response (comment 12) and copy of the submitted questions provided in Appendix 4 of the addendum.
Oxford Academic Health Science Network	19	23	1st para	A t-test approach had been used to identify the unknown variances and remove outliers, which is stated by yourselves earlier in your review.	The rule by which outcomes were deemed different enough from others (so-called 'outliers') to be excluded is not equivalent to a justification for excluding outcomes simply because they are different. Removing such observations without any justification deprives us of some of the evidence and what we might infer from that evidence.
Oxford Academic Health Science Network	20	23	2nd para	"The unpublished report…. provided time to treat data however, these data were only for patients directly admitted to the CSC." This point is made several times through the document, that no equivalent data (scan to MT) was reported for patients who received MT following transfer from an ASC. Scan to MT at an ASC is not a good measure of the impact of AI in stroke care. This measure of time starts at the first scan and finishes at the time of arterial puncture. The first scan is normally a CT scan, not CTA or CTP. It then	This issue has been extensively discussed, both in the original DAR and the addendum. 1) There is an inherent deficiency in consideration of effectiveness/cost-effectiveness based on time to decision/intervention outcomes only, in that the addition of AI-derived software technologies also has the potential to

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
				includes the time it takes to arrange for transfer to the CSC, which is part of the pathway that is outside of the control of the stroke unit and therefore not linked to the use of AI. The unpublished evaluation report, therefore, rightly uses the DIDO or door in door out time, to account for the impact of AI on stroke care at an ASC. Subsequently, in the same paragraph, the reviewer is stating that 'It should be noted that this Diagnostic Assessment seeks to evaluate whether the introduction of AI-derived software technologies will result in changes that translate to changes in outcome for the patients. The scope for this topic did not include DIDO or clinical decision times as outcomes'. We disagree with this as your research questions investigate how AI is assisting treatment decisions, surely the time to treatment is a key factor in ascertaining this. Furthermore, this critique is inconsistent with the remainder of your review which has a focus on time to treatment for thrombolysis and scan to MT. Finally, within this same paragraph, the reviewer states that 'it is, therefore, not possible to make any link between the data provided on changes in time to treatment and data provided on clinical outcomes.' This point has been well studied and evidenced regardless of the use of AI decision support (Holodinsky et al, 2018) and a reduction in DIDO at ASC is found to improve patient outcomes. Whilst the data that is presented in our report does not provide that link, it would be remiss not to include DIDO as a measure of impact of AI on clinical decision making.	change which patients are referred for MT 2) Any consideration of time-to-treatment requires evidence that time savings can be achieved, which are sufficient to impact clinical outcomes. This evidence is not available in relation to reductions in DIDO or time to treatment decision. Also, as noted on the addendum, the stakeholder's report indicated that the median DIDO times provided by the report, for 16/20 ASCs, also did not indicate a consistent reduction in time following implementation (median DIDO times increased, following implementation, at 9/16 sites, were unchanged at one site and decreased at 6/16 sites) 3) Holodinsky et al 2018 has been repeatedly cited, by some stakeholders, as providing evidence that a reduction in DIDO times improves patient outcomes. However, the significant variable in the Holodinsky model was actually much more specific (reduction in DIDO time to 60 minutes overall, comprising 30 minutes to IVT and 30 minutes from IVT to transfer). This is not the same as 'reduction in DIDO at the ASC' and is not evidence of clinical benefit from any reduction of DIDO time, irrespective of the magnitude of any such reduction or the administration/timing of administration of IVT.

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
					It may also be worth noting that the shortest median post-implementation DIDO time, for any site, listed in the stakeholder's report was 90 minutes.
Oxford Academic Health Science Network	21	26	Bottom of page	The reviewer refers to the communication aids included in the e-Stroke technology (mobile app and web access to brain images). The RWE is a whole system review, we are not isolating one aspect of functionality from another - we are assessing impact on its' entirety.	Please see previous response (comment 11). In addition, please note that (as discussed in the addendum) any effect of these technologies on communication/image sharing cannot be separated from their effects on image interpretation/decision making. As the stakeholder notes, this is a 'whole system'.
Oxford Academic Health Science Network	22	27	2nd Para	Sentence beginning 'Of particular note….' This statement is incorrect. The data presented in the graphs is the number of CTA scans processed by e-Stroke, not the number undertaken. We have clarified this directly with colleagues at NICE via e-mail on Monday 7th August. The purpose of this analysis is to determine whether use of e-Stroke overtime has changed and if that has an impact on patients and clinical outcomes i.e., increased and more consistent use leads to better outcomes. It is important to note that some sites have automatic processing of scans and equally as important to understand how stroke services differ from site to site. It is also worth noting that during COVID, clinical audit activities were suspended so data during this period may not considered as robust, i.e., activities may still have been happening but not recorded.	This point was not clear from the plots, as originally presented. An amended version of the addendum has been provided following additional information provided by the stakeholder, post-submission. No data were provided on the numbers of CTA scans undertaken during the pre-implementation, comparator period. We note the stakeholder's comment in relation to data collection during the COVID period.
Oxford Academic Health Science Network	23	27	Table 3	OAHSN and Brainomix have been given a U for Q2 – Did the study population include an appropriate spectrum of patients. Our report goes into some detail on the variation of patients included in the study, those that present to ASC or	The question concerns the comparability of presenting patients in the pre- and post-implementation periods. The stakeholder's report includes no such comparison, nor does it

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
				CSC, those that present greater or less than 6 hours after onset of symptoms etc. This is not unknown. In reference to comment 2 above, it is uncertain why this information wasn't requested at the same time as other clarifications were sought. In response to Q5, the authors of the report informed the reviewer that pathway maps had been obtained from all sites involved in the evaluation and could be provided yet this shows as an N in the table.	report information about patients in a format that could facilitate should a comparison. This information was requested (please see Appendix 4 of the addendum: 'Please provide any information that you have (patient characteristics) to inform the question of whether patient populations were comparable before and after implementation of e-Stroke' The questions to the stakeholder included: 'Please provide details of how e-Stroke was incorporated in the care pathway (at each site if there were differences between sites) – please include details of type/seniority of clinician and the number and order of clinicians if more than one involved' The stakeholder's response stated that: '*Clinician will vary depending on site and we would need to seek further information on this. We do have pathways mapped for almost all sites which indicates when e-Stroke is used in the pathway. A summary of this can be provided later if deemed useful by the panel'* 'Noting that the implementation period appears to have overlapped the COVID-19 pandemic, please provide any information that you have about any changes in the care pathway that occurred (other than the implementation of e-Stroke) during the study period' The stakeholder's response stated:' *We don't have a full picture from all sites at present, this will be reported in our final evaluation report'.*

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
					As noted in the addendum, the stakeholder's report includes the following text , in relation to changes (other than the implementation of the AI-derived software technology) occurring during the study period: *'Despite this being a multi-year real-world evaluation, which relies on an iterative approach to adapt to a fast-changing stroke landscape which continuously sees new challenges and practices evolving, our hypothesis and value proposition have remained unvaried: "e-Stroke aids the evaluation of imaging in patients with suspected acute stroke and decisions for reperfusion therapies. This leads to a reduction in disability and enhanced quality of life with associated cost savings for the Health and Social Care System". Our hypothesis relies on the assumption that the benefits of e-Stroke will be maximised through quality improvement initiatives. This is because clinical outcomes are likely to improve because of faster diagnosis and treatment which are being facilitated by the technology, but also because of improvements across the acute stroke pathway. We have now entered phase three of the evaluation. The closer we get to the end of the project, the more our efforts are being directed towards the identification of promising areas for quality improvement as our goal is to support the Integrated Stroke Delivery Networks (ISDNs) to optimize the benefits of e-Stroke, driving change and maximising impact on operational and clinical outcomes.'* This text describes other changes that were expected to occur during the evaluation period, and arbitrarily ascribes any changes and any

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
					benefits that may have accrued from such changes to the AI-derived software technology.
Oxford Academic Health Science Network	24	28	2nd Para	Potential sampling bias. 'In addition, the majority (27/34) respondents were from clinical disciplines who do not have expertise in interpreting brain scans or conducting MT.' This statement implies that stroke consultants do not have expertise in interpreting brain scans, which is incorrect. In fact, this presents the question of impact. The introduction of AI in stroke care is to support those clinicians that are not experts in their field, but those that would use decision support to inform their next steps. This does not suggest that they are relying on the AI, but when having more information available to them would refer to the CSC, for a specialist opinion and subsequent referral for MT. AI should not just be compared against experienced radiologists, but with other experienced disciplines within the stroke unit.	The stakeholder does not comment on the primary point made under this heading, which was that no information was provided about survey response rates and it was not clear who/how many people had been asked to participate in the survey or how potential participants had been selected. The stakeholder's point about the potential use of AI to support clinicians in referring to CSC for specialist opinion and subsequent referral for MT is a valid one; unfortunately, their report does not include any information to support the suggestion that the software is being used in this way, and the copy of the survey provided did not include any questions about this specific use.
Oxford Academic Health Science Network	25	29	1st Para	Potential reporting bias. The reviewer suggests that the definitive statement about use of e-Stroke in the Royal Sussex is invalid. However, there is only 1 wte stroke consultant in post at the Royal Sussex, so this statement is valid. Furthermore, numbers of responders were made available to the reviewer, by local and clinical discipline through prior requests.	The point being made was that this is a very definitive statement, which is being made on the basis of opinion from one person. It is not clear, from the stakeholder's comment, whether the single stroke consultant at the Royal Sussex was the only person at this site to be interpreting CTA scans/using e-Stroke. The stakeholder did provide, in response to questions from the EAG, overall absolute numbers of survey respondents, and numbers by clinical discipline and site. However, these data were not provided no information was provided about how survey participants were selected or about how many people were invited to participate (response rate).

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
Oxford Academic Health Science Network	26	29	Final para	Collection of information about usage, workflow, and accuracy from a survey of clinicians. The reviewer suggests that 'a number of survey questions were of questionable value.' The evaluation has utilised a formative, mixed methods approach, converging qualitative and quantitative data in a cross-case analysis. We used Theory of Change to determine what was important to understand from a clinical perspective through stakeholder engagement with experienced clinicians. The qualitative data, and the questions which are deemed to be of questionable value by the reviewer, are used to gain perceptions of the teams using the technology, which is then converged with quantitative data to determine impact. Furthermore, how could data on the proportion of cases that e-Stroke is used for, or, how frequently you agree with e-Stroke, be collected at implementation (as suggested by the reviewer), when the technology has only just started to be utilised?	The stakeholder's point appears to be that they have combined qualitative and quantitative information in their report. Whilst qualitative data may be useful to understand some aspects of users' perceptions and e.g. how these might affect willingness to use the technology, it is difficult to see what collecting data on users opinions about e.g. whether *'Brainomix has reduced the time taken to start thrombolysis'* could add to measuring the time take to start thrombolysis before and after implementation of Brainomix. It should also be noted that the stakeholder's report does not include any quantitative data about accuracy. All statements about accuracy are reliant upon qualitative/survey derived information. The EAG did NOT suggest that *'data on the proportion of cases that e-Stroke is used for, or, how frequently you agree with e-Stroke, be collected at implementation*'. We noted that it is important to use direct measurement to establish that a change in, e.g. rates of intervention or time to intervention, has occurred between the periods before and after the implementation of the technology, and that qualitative data may then be useful to explore how the technology may have contributed to any such change.
Oxford Academic Health Science Network	27	32		In relation to our comment regarding research question 1 above (3), we would like to reiterate that it is unlikely that AI-derived software will have an impact on supporting earlier treatment with thrombolysis.	Please see previous response.

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
				This is also in reference to our comment 8 where the reviewer has stated that clinical decision times are not in the scope of this review.	
Oxford Academic Health Science Network	28	40		Hypothesis. 'This statement appears to assume that any beneficial changes, across the acute care pathway, that may occur during or subsequent to the implementation period are attributable to the effects of e-Stroke.' This is not what the statement means. It suggests that if our hypothesis is correct, clinical outcomes are likely to improve in combination with quality improvement activities across the stroke pathway ensuring the optimal use of e-Stroke. In the same paragraph the reviewer suggests that the report indicates 'small increases in the rates of MT between 0.57% and 3.46% at CSCs.... and between 1.42% and 6.51% at ASCs' These increases require context and are not considered small increases when the target for MT treatment is 10%. This confirms that the reviewers are not well informed on stroke care and the complexity of delivery in the NHS.	We cannot see any text in the stakeholder's report that acknowledges the possibility that their hypothesis, *"e-Stroke aids the evaluation of imaging in patients with suspected acute stroke and decisions for reperfusion therapies. This leads to a reduction in disability and enhanced quality of life with associated cost savings for the Health and Social Care System"*, may not be correct and evidence was not presented to support *'reduction in disability and enhanced quality of life with associated cost savings for the Health and Social Care System'* We acknowledge the stakeholder's point regarding the size of the reported effect, but not that there are some concerns around the reliability of these estimates and that it remains important to establish that any additional thrombectomies result in beneficial clinical outcomes.
Oxford Academic Health Science Network	29	41		Observation – why is the reviewer carrying out their own statistical analysis on data they are quoting to be incomplete and inconsistent rather than confirming the dataset with the authors of the report? There was nothing incorrect however, it would seem that the issue is that we have run additional request for on the updated dataset which included an additional 3 months of data.	The stakeholder was asked a series of questions (see Appendix 4 of the addendum) in an attempt to provide usable data for this assessment. Both the stakeholder's initial report their response to these questions lacked clarity and consistency. Given the limited timeframe allowed, and the incomplete response received, the EAG

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
					attempted some analysis based on what data were available; it is for the committee to decide whether or not these analyses were useful.
Oxford Academic Health Science Network	30	43		Suggestion that Stroke Consultants have no expertise in interpreting brain scans – see comment above. There are further pedantic comments on this page, one saying that radiographers have not been included, when they have, which is clear from the associated text. There also appears to be a misunderstanding by the reviewer that radiographers have experience in interpreting brain scans.	We acknowledge the stakeholder's comment regarding stroke consultants and apologise for any perceived implication that stroke consultants lack expertise. The key points about the survey remain its small size (maximum 34 in total and maximum 5 per site), poor reporting (e.g. not information about participant selection or response rates), and use to collect qualitative opinion about quantitative outcomes.
Oxford Academic Health Science Network	31	49	4.2	Further information on the approach to the health economics are available and could have been requested at any time but were not. As a full economic assessment is to be reported in September of this year, this section is intentionally light touch, providing, as stated in the document, an overall approach. Not sure why a critique has been included in this scope, which we understood to be a review of the literature. The critique goes into some detail when no further clarification was ever sought from the authors; Health economists and statisticians at the University of Oxford have both helped to develop and ratify this approach.	The EAG have clearly indicated that the economic assessment is reported only as a proposal. It was critiqued in order provide as much information about a part of the evaluation that at least appeared to be related to cost effectiveness. The EAG did not consider that it was a good use of the limited time available for this assessment to probe any further.
Oxford Academic Health Science Network	32	62	2nd Para	The reviewer refers to confounders. It should be noted that the AI is not linked to ambulance transfer times and availability. The reviewer says that RWE is not considered as valuable as other research methods because it includes confounding factors (p25); However, DIDO is considered a	Please see previous response (comment 20)

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
				better more fined measure than Scan to MT because it excludes the variation seen in patient transfer times to Thrombectomy Centres. Contradiction on whether DIDO is a good measure when considering confounding factors. In this statement the reviewers suggests that the reasoning for including DIDO times at acute centres as they 'may be confounded by the variation seen in patient transfer time to the thrombectomy centre caused by pressures on ambulance availability' is not an example of confounding factors but rather a measure of the true, real-world effects of the intervention. Whilst this is RWE we have taken measures to ensure it is as robust as possible. Through regression analysis we have identified that ambulance times have a significant impact on the effectiveness of e-Stroke; this further supports our decision to include DIDO.	
Oxford Academic Health Science Network	33	63	1st Para	Agree with original comments from your review regarding rapid image transfer. PACs systems are never going to be able to transfer images with the same speed and quality as AI-derived software solutions can, allowing for instant communication with the thrombectomy centres. We agree that a wider scope for this assessment should be considered.	Please see previous responses (11 and 21)
Viz.ai	34	17	3.2 Addendum 2	We would like to highlight two relevant, independent studies for potential inclusion. One was an FDA-sponsored study published prior to the May 23 cut-off date for the literature review, and the other one was published a month afterward. 1. Kunst et al. "Real-World Performance of Large Vessel Occlusion Artificial Intelligence–Based	We would like to note that Viz.ai did not provide an evidence submission at the start of this up-date. Had they done so, all cited references would have been checked against the inclusion criteria for this assessment (as is usual procedure for the EAG) and included in the addendum as appropriate.

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
				Computer-Aided Triage and Notification Algorithms—What the Stroke Team Needs to Know." **J Am Coll Radiol.** 2023 May 16:S1546-1440(23)00335-6. doi: 10.1016/j.jacr.2023.04.003. Epub ahead of print. PMID: 37196818. 2. Delora et al. "Viz LVO versus Rapid LVO in detection of large vessel occlusion on CT angiography for acute stroke." **J NeuroIntervent Surg** 2023;0:1–5. doi:10.1136/jnis-2023-020445.	In addition, the very short timelines specified for this work did not allow for the conduct of any additional up-date searches during the production of the addendum (as would usually be undertaken for a DAR or EVA). None of the six references listed were identified by our up-date searches for this addendum. We have looked into the reasons for this (details below). In addition, we have obtained copies of all the listed references. The very small number of working days between the provision of this list and the committee date means that we cannot provide a full written inclusion assessment/summary of results or an amended report. However, we have provided (below) a brief summary of the key points of each study. **Kunst et al. 2023** This study would be retreived by our search strategy, however, it was not in the original search output file (i.e. it was not available/indexed at the time of searching) The study provides a comparison of the stand-alone accuracy of RAPID and Viz, but no data on accuracy of either device combined with a human reader and no data for any other specified outcome. **– no data that would have been included in the current addendum, study would have been listed in Table 2** **Delora et al 2023** This study did not appear on Medline until June 2023 (after our search date).
Viz.ai	35	36	3.2.4 Addendum2	We would like to highlight three additional, relevant, independent studies that are not currently included in Table 5 or the preceding discussion. In February 2023, multiple conference abstracts were published, and while none of the studies reported clinical outcomes, the time to treatment outcomes are some of the strongest to date because they are based on more rigorous study designs and multi-centre data. In a multi-centre, cluster randomized clinical trial, implementation of Viz LVO significantly improved door-to-groin puncture time in both a univariate analysis and a multivariable adjusted mixed-effects model. - Martinez-Gutierrez et al. "Machine Learning-Enabled Automated Large Vessel Occlusion Detection Improves Treatment Times: A Multi-Center Cluster Randomized Clinical Trial." **Stroke**. 2023;54:ATMP41 In a multi-centre, prospective cohort study, implementation of Viz LVO significantly improved the median door-in-door-out time of LVO acute ischemic stroke patient transfers from spoke to hub. - Chaudhry et al. "Effect Of Automated Large Vessel Occlusion Detection On Door-in-door-out	

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
				Times At Primary Stroke Centers: A Multi-center Prospective Cohort Study." **Stroke**. 2023;54:A27. A multi-centre study involving 14,116 patients across 166 facilities and 17 U.S. states compared time to treatment outcomes between two contemporary cohorts consisting of hospitals implemented with Viz LVO and hospitals without AI-based stroke software. The median door to neurointerventionalist notification time was 39.5 minutes faster in the Viz LVO hospitals, and the door to thrombolytic needle time was 4 minutes faster. - Sevilis et al. "Validation Of Artificial Intelligence To Limit Delays In Acute Stroke Treatment And Endovascular Therapy (VALIDATE)." **Stroke**. 2023;54:AWP81	The study provides a comparison of the stand-alone accuracy of RAPID LVO and Viz LVO, but no data on accuracy of either device combined with a human reader and no data for any other specified outcome. **– no data that would have been included in the current addendum, study would have been listed in Table 2** **Martinez-Gutierrez et al 2023** This study was in Northern Light Life Sciences Conference Abstracts (searched resource), but as a title only (no abstract) which did not include sufficient information to be retrieved by our search strategy. This reference is a conference abstract providing a brief description of some results from a cluster randomised controlled trial of the implementation of Viz LVO. A shorter time from door to groin puncture (DTG) was reported for centres with implementation (n=50 patients, median DTG 73 min [IQR 32 to 102 min]) than in centres without implementation (n=199, median DTG 91 min [IQR 63 to 108 min]). The abstract reports that, adjusting for age, sex and baseline NIHSS resulted in an improvement in DTG of 15 min (95% CI 4 to 26 min). The abstract does not provide any detail about how Viz LVO was implemented or about any other changes in the care pathway. No clinical outcomes data were reported. **– the results described would have met inclusion criteria for the current addendum** **Chaudhry et al 2023**
Viz.ai	36	36	3.2.4 Addendum 2	We would also like to highlight for consideration an additional, relevant, independent study that compares both time to treatment and clinical outcomes in a cohort of patients following implementation of Viz LVO to a cohort of historical controls. - Hassan et al. "Artificial Intelligence–Parallel Stroke Workflow Tool Improves Reperfusion Rates and Door-In to Puncture Interval." **Stroke Vasc Interv Neurol.** 2022;2:e000224. DOI: 10.1161/SVIN.121.000224	
Viz.ai	37	66	6.2 Addendum 2	Referencing comment 2, we would like to note the recent results for the time to treatment outcome (door-to-groin puncture) in a multi-centre, cluster randomized clinical trial studying the effects of Viz LVO implementation.	

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
					This study was in Northern Light Life Sciences Conference Abstracts (searched resource), but did not include sufficient information to be retrieved by our search strategy. This reference is a conference abstract which does not report data for any specified outcome (DIDO times only**– no data that would have been included in the current addendum** **Sevilis et al 2023** This study was in Northern Light Life Sciences Conference Abstracts (searched resource), but did not include sufficient information to be retrieved by our search strategy. This reference is a conference abstract comparing centres that did and did not use Viz LVO. The abstract reported a shorter median door to needle (DTN) time for IVT for patients in centres with Viz LVO, 40 min (30 to 525 min, not clear if range or IQR) than for patients in centres without Viz LVO, 44 min (32 to 57.5 min). The abstract reported that a total of 8557 AI patients and 5559 non-AI patients were included, but it was not clear how many patients were included in each analysis. No clinical outcomes data were reported and there was no information about patient characteristics, or about how Viz LVO was implemented or about any other changes in the care pathway. **– the results described would have met inclusion criteria for the current addendum** **Hassan et al 2022**

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
					This reference is not in Embase or Northern Light Life Sciences Conference Abstracts. Our addendum includes data from two publications (references 9 and 10 in the addendum) which appear to relate to the same study. These included publications report data for patients transferred from a PSC to the study authors' CSC. The additional publication, listed here, reports data for patients who presented directly to the CSC. The publication reports a pre- post-implementation comparison for Viz LVO, used as an automated alert, triage system in a CSC. Baseline participant characteristics indicated no significant differences between the presenting populations pre- and post-implementation. The study reported a reduction in mean door to groin puncture time for thrombectomy patients who presented directly to the CSC, from 206±169.1 min (n=86) pre-implementation to 119±83 min (n=102) post-implementation. The study also reported no significant difference in the rate of symptomatic haemorrhage pre- to post-implementation (7/86 [8.1%] vs 6/102 [5.9%]), no significant difference in mortality at discharge pre- to post-implementation (18/86 [20.9% vs 23/102 [22.5%]) and an increase in the proportion of patients with good TICI (2b-3) (73/86 [84.9%] to 96/102 [94.1%]). **– the results described would have met inclusion criteria for the current addendum** In summary, three of the six listed studies reported some data that would have been included in our addendum. These studies

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
					provide some additional data for Viz LVO, but would not substantively change the overall conclusions of our addendum.
NHS England	38	General		To: Gail Allsopp Chief Medical Officer NICE 22 July 2022 Dear Gail, Re: Artificial intelligence (AI) software to help clinical decision making in stroke (diagnostics consultation document) Thank you for sharing the draft consultation document offering guidance on using artificial intelligence (AI) software for analysing CT brain scans in the NHS, and for the opportunity to hold a discussion regarding your findings. AI software to augment clinical decision-making in stroke is currently being used in approximately 75% of acute trusts in the NHS, with strong clinical support and over 50,000 scans having been interpreted in the last year alone. The findings within the report therefore sit at odds with current practice and NICE (NG128) 2019 guidance on access to mechanical thrombectomy at 6-24 hours. Evidence reviews for the update to the National Clinical Guideline for Stroke (due for publication in March 2023) will be supporting thrombolysis and thrombectomy beyond 4.5-6 hours with evidence from randomised controlled trials that used AI-supported interpretation of acute brain imaging, particularly CT perfusion. The recommendations are also counter to a number of national policy directives regarding the value of AI in stroke care, which have cascaded throughout clinical teams and the wider NHS over the past three years:	

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
				• NHS Long Term Plan (January 2019) • Diagnostics: Recovery and Renewal (October 2020) • National Stroke Service Model (May 2021) • Artificial Intelligence (AI) Software in Neuroscience for Stroke Decision Making Support Procurement Framework (February 2022) • GIRFT Programme National Specialty Report (April 2022) • Stroke RightCare Toolkit (May 2022) We therefore find the guidance not only at odds with current practice, but also narrowly focused solely on the diagnostic performance characteristics of AI decision support, without paying appropriate heed to the existing randomised trial evidence demonstrating the clinical utility of the technology as a component part of hyperacute stroke care. We would question whether the scope and methodology of the appraisal appropriately reflected this current state of evidence and practice, including whether the evidence review was sufficiently broad. We have serious concerns regarding your recommendations, with a potentially significant impact on patient safety should the report be published at this juncture. There is a risk that existing hyperacute stroke pathways will be disrupted as a result of publication, introducing avoidable delays to time dependant interventions, resulting in harm to patients. We recommend that publication of this document is paused until these issues are addressed and the apparent contradiction to the NICE 128 guidance is resolved. We would strongly encourage consideration of the value of other evidence sources available, such as the 50,000	

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
				routine AI cases acquired from the recent GIRFT / specialised commissioning mechanical thrombectomy quality reviews, and from the ongoing evaluation by the Oxford Academic Health Science Network of the Brainomix e-Stroke Suite software. NHS England and GIRFT are very keen to work with NICE, alongside the ongoing endeavours of Oxford AHSN, to establish a systematic collaboration project to gather and review data from both users and suppliers of AI software in order to derive additional evidence on the efficacy and benefits of this technology. It is feasible that this may be incorporated into the next evaluation cycle, with the aim of producing a more detailed summary by April 2023. [redacted], and [redacted], [redacted], have also offered their support to identify the best way forward for this important programme. We look forward to working with you further to ensure that this technology continues to contribute to delivery of the National Optimal Stroke Imaging Pathway, ensuring rapid diagnosis and access to disability-saving interventions for patients suffering a stroke. Yours sincerely, [redacted] NHS England	

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
				[redacted] NHS England [redacted] NHS England [redacted] [redacted] NHS England CC. Mark Chapman Nick Baillie Rebecca Albrow [redacted]	
NHS England	39	General		14 August 2023 Dear Donna, Re: NICE assessment: stakeholder consultation on addendum for 'Artificial intelligence (AI) software to help clinical decision making in stroke' Thank you for sharing the addendum document to the original consultation offering guidance on using artificial intelligence software for analysing CT brain scans in the NHS. We wrote to you in July 2022 (attached) outlining our serious concerns regarding the original recommendations, with a potential significant impact on patient safety should the report be published at this juncture.	

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
				We can find no evidence in the addendum of attention paid to our primary concern of the disparity from the NICE (NG128) guidance, updated National Clinical Guideline for Stroke National Clinical Guideline for Stroke (strokeguideline.org), NHS and international clinical practice with this new appraisal. The former all acknowledge the pivotal use of AI in decision making to deliver recanalisation therapy in the time window 4.5-24hrs. The wealth of real-world data now available and acknowledged in this addendum appears to have been dismissed. We remain committed to work with you to explore how we may reach an appropriate consensus of narrative that will enable the significant improvements in care seen since the introduction of AI into over 90% of stroke services in England to continue. Yours sincerely [redacted] NHS England [redacted] NHS England	