Canada's Drug and Health Technology Agency

CDA-AMC Health Technology Review

RapidAI for Stroke Detection: Main Report

August 2024

Rapid Review with Expert Input

Abbreviations

AI	artificial intelligence
CDA-AMC	Canada's Drug Agency
CI	confidence interval
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HTA	heath technology assessment
ICH	intracranial hemorrhage
LVO	large vessel occlusion
mRS	modified Rankin Scale
NICE	National Institute for Health and Care Excellence
NIHSS	National Institutes of Health Stroke Scale
QUADAS-2	Quality Assessment of Diagnostic Accuracy Studies 2
ROBINS-E	Risk Of Bias In Nonrandomized Studies – of Exposures
ROBINS-I	Risk Of Bias In Nonrandomized Studies – of Interventions

Key Messages

What is the Issue?

- Stroke is a sudden loss of neurological function caused by poor or interrupted blood flow within the brain. It is 1 of the leading causes of death and disability in Canada. Symptoms include numbness or weakness, difficulty speaking, dizziness, loss of vision, sudden trouble walking, and loss of balance or coordination. Patients with suspected stroke require prompt evaluation to determine the type of stroke, to assess the severity of damage, and to guide treatment decisions.
- RapidAI is an artificial intelligence-enabled software platform that facilitates the viewing, processing, and analysis of CT images to aid clinicians in assessing patients with suspected stroke. Understanding the potential benefits and harms of using RapidAI is important to clarify its role in stroke detection.

What Did We Do?

- We sought to identify, synthesize, and critically appraise literature evaluating the effectiveness, accuracy, and cost-effectiveness of RapidAI for detecting large vessel occlusion (i.e., ischemic stroke) and intracranial hemorrhage (i.e., hemorrhagic stroke).
- We searched key resources, including journal citation databases, and conducted a focused internet search for relevant evidence published up to July 22, 2024. We screened citations for inclusion based on predefined criteria, critically appraised the included studies, narratively summarized the findings, and assessed the certainty of evidence.
- We highlighted and reflected on the ethical and equity implications of using RapidAI for stroke detection found in the clinical literature, integrating these considerations throughout the review.
- We engaged a patient contributor who had experienced a hemorrhagic stroke, to learn about her experience, perspectives, and priorities. Additionally, we incorporated feedback from clinical and ethics experts, the manufacturer, and other interested parties.

What Did We Find?

- We found 2 cohort studies and 11 diagnostic accuracy studies that assessed the effectiveness and accuracy of RapidAI for detecting stroke. Among these, 3 studies evaluated RapidAI as it is intended to be used in clinical practice (i.e., to complement clinician interpretation of CT images), while the remaining 10 studies assessed RapidAI as a standalone intervention.
- The patient contributor identified important outcomes for stroke care, including speed and accuracy of diagnosis, minimizing the damaging effects of stroke, and reducing mortality rates. She also highlighted ethical considerations regarding the use of artificial intelligence in health care, such as data privacy, equitable access, and informing patients about the use of AI technologies in the care pathway.

- Low certainty evidence suggests that evaluation of CT angiography images by Rapid LVO combined with clinician interpretation, compared to clinician interpretation alone, may result in clinically important reductions in radiology report turnaround time in patients with suspected stroke. For detecting intracranial hemorrhage, low certainty evidence suggests that Rapid ICH combined with clinician interpretation, using clinician interpretation as a reference standard, has a sensitivity of 92% (95% CI, 78 to 98%) and a specificity of 100% (95% CI, 98 to 100%). However, estimates of sensitivity and specificity for detecting large vessel occlusion varied, based on studies using different modules of RapidAI as a standalone intervention, providing only indirect accuracy data.
- The effects of RapidAI on other time to intervention metrics, measures of physical and cognitive function, and response to therapy (e.g., reperfusion rates) were very uncertain. We did not identify any evidence on the effects of RapidAI on many important clinical outcomes, including patient harms, mortality, health-related quality of life, length of hospital stay, and health care resource implications.
- We did not find any studies on the cost-effectiveness of RapidAI for detecting stroke that met our selection criteria for this review.
- Ethical and equity considerations related to patient autonomy, privacy, transparency, access, and algorithmic bias have implications across the technology lifecycle when using RapidAI for detecting stroke.

What Does This Mean?

- RapidAI has potential to improve the diagnostic process by assisting clinicians interpreting CT images for patients with suspected acute stroke; however, the impact on patient-important outcomes is uncertain.
- To improve the certainty of findings, there is a need for evidence from robustly conducted studies at lower risk of bias that enroll diverse patient populations and measure outcomes that are important to patients, with improved reporting.
- The cost-effectiveness of RapidAI for stroke detection is currently unknown.
- In addition to the evidence on the effectiveness and accuracy of RapidAI for detecting stroke, decision-makers may wish to reflect on the ethical and equity considerations that arise during the deployment of AI-enabled technologies, such as those related to autonomy, privacy, transparency, and explainability of machine learning models, and the need for considerations related to equity and access in their design, development, and deployment.

Context and Policy Issues

What Is Stroke?

Stroke, also known as cerebrovascular accident, is a life-threatening medical condition characterized by loss of neurological function. Symptoms include numbness or weakness, difficulty speaking, dizziness, loss of vision, sudden trouble walking, and loss of balance or coordination.[1] In Canada, stroke is 1 of the leading causes of death and disability,[2,3] with more than 100,000 stroke events resulting in hospital or emergency department presentations each year.[4]

Health care inequities exist in stroke incidence, prevalence, symptoms, quality of care, and outcomes across factors such as race, ethnicity, gender, sex, disability status, age, geographic location, and socioeconomic status.[5-12] For instance, about 59% of stroke-related deaths in Canada occur in women, and women also tend to have worse outcomes after stroke than men.[10,11] Regarding geographic location, data from 2007 to 2011 indicate that rural hospitals in Canada have significantly higher 30-day in-hospital mortality rates following stroke compared to urban academic hospitals and the national average.[13]

Strokes are classified as either ischemic or hemorrhagic. Ischemic strokes, the most common type, occur when a blood vessel supplying the brain becomes blocked or clogged, impairing blood flow.[14] Hemorrhagic strokes occur when a blood vessel ruptures, causing bleeding into the brain.[14]

What Is the Current Practice?

Acute stroke assessment involves prompt and comprehensive evaluation of individuals suspected of having a stroke to determine the type of stroke, the severity of symptoms, and their eligibility for specific treatments. It typically includes an assessment of the

individual's medical history, physical examination (e.g., heart rate and rhythm, blood pressure, temperature, oxygen saturation), neurological examination, laboratory tests (e.g., blood work including electrolytes, random glucose, complete blood count, coagulation status, and creatinine), and cerebrovascular imaging with CT scans or magnetic resonance imaging.[15]

Stroke diagnosis and intervention is time sensitive. The 2022 Canadian Stroke Best Practice Recommendations[15] on acute stroke management emphasize the importance of performing neurovascular imaging and determining treatment eligibility without delay, as the potential benefits of stroke interventions and patient outcomes decline over time.[16-18] Following ischemic stroke, an estimated 1.9 million neurons die every minute in which it is untreated.[16] Accurately determining whether a stroke is ischemic or hemorrhagic with neuroimaging studies, often using CT scans, is crucial for selecting appropriate treatment options.[14] Some of the main interventions for ischemic stroke include intravenous thrombolysis and mechanical thrombectomy, whereas hemorrhagic stroke treatments aim to control bleeding, reduce pressure on the brain, and prevent further complications.[15,19] Misidentifying the type of stroke can lead to inappropriate treatments that may exacerbate the condition; for example, administering thrombolysis for a hemorrhagic stroke can worsen the bleeding.

What Is RapidAI and What Are Its Potential Benefits?

RapidAI (iSchemaView, Inc., Menlo Park, California) is an artificial intelligence (AI)-enabled software platform that provides tools for various indications and uses including stroke and cerebrovascular imaging. It includes a range of products that facilitate the viewing, processing, and analysis of CT images (including non-contrast CT, CT angiography, and CT perfusion) that can be used by clinicians to assist in assessing patients with suspected stroke and deciding appropriate treatment.[20] RapidAI is intended to complement, rather than replace clinician interpretation of CT images and should be used as a supportive tool rather than a standalone diagnostic intervention.[21] The software includes several static AI-derived algorithms for evaluating the brain's physiological status, including Rapid ICH, Rapid ASPECTS, Rapid CTA, Rapid LVO, Rapid CTP, and Rapid HVS.[20,22] While RapidAI has numerous features and functionalities, what is relevant for this report is its capability to detect suspected intracranial hemorrhage (ICH) and large vessel occlusion (LVO) that is used to inform stroke diagnosis. Potential benefits may include reducing crucial time intervals in stroke care, improving workflow efficiencies in emergency settings, and facilitating faster triage, decision-making, and treatment initiation, which could result in better patient outcomes.

As of March 2024 (i.e., the time of our checking of its regulatory status), RapidAI (version 4.9.2.1) is licensed for sale in Canada as a Class III medical device. The application runs on standard computers or virtual platforms integrated into existing radiology workflows and uses existing forms of cerebrovascular imaging and computer information systems (e.g., radiology information systems and picture archiving and communication systems).

Why Is It Important to Do This Review?

Globally, we are seeing an increase in medical devices relying upon software incorporating AI.[23] With the inherent nature of AI being a disruptive technology in health care, its comprehensive assessment within health technology assessment (HTA) is essential to ensure that digital health technologies are adequately equipped to balance benefits and harms, while being interoperable and equitably accessible for people living in Canada. To this end, Canada's Drug Agency (CDA-AMC) recognized the importance and timeliness of assessing AI-enabled medical devices and included an activity in its 2024-25 Annual Business Plan to support the implementation of AI by several means, including the assessment of an AI technology in radiology using a digital health assessment framework.[24] This review addressed that objective, using RapidAI for stroke detection as the topic, alongside an implementation review (found on the project website).

CDA-AMC has recently established information-sharing and collaborative relationships with various organizations, including an international HTA partnership of HTA bodies.[25] From this partnership, CDA-AMC learned of the Scottish Health Technology Group's evidence framework, which outlines an approach to digital health technology assessment. This includes an HTA framework[26] and Digital Technology Assessment Criteria as an add-on component.[27] This review on RapidAI applied the Scottish Health Technology Group's HTA framework.[26] The accompanying implementation review (found on the project website) applied Digital Technology Assessment Criteria,[27] leveraging existing work to ensure alignment and harmonization across organizations and to gain efficiency and sustainability and planning to share our experiences with the framework and the criteria with the international partners.

Objective

To support decision-making about the use of RapidAI for the detection of stroke, this rapid review summarizes and critically appraises the available studies on the effectiveness, accuracy, and cost-effectiveness of RapidAI to detect ICH and LVO.

Research Questions

1. What is the effectiveness and accuracy of RapidAI in stroke detection?
2. What is the cost-effectiveness of RapidAI in stroke detection?

Methods

Study Design

To inform the design of this review, we conducted an informal scoping search of existing literature. We identified several systematic reviews[28-31] that assessed the use of imaging software with AI features for informing clinical decision-making in stroke; however, there was a lack of up-to-date systematic reviews that matched the scope of our research questions (e.g., no reviews were specific to RapidAI for stroke detection). Considering the project's timelines (April to September 2024), designed to share our experiences assessing an AI technology with the international partners in a timely manner, we did not anticipate being able to undertake a full systematic review. Therefore, we conducted a rapid review of primary studies and economic evaluations to address research questions 1 and 2. Our methods were documented in an a priori project plan and were guided by methodologic standards for Cochrane rapid reviews.[32]

To address health equity in our review process, we followed guidance outlined in the *Cochrane Handbook for Systematic Reviews of Interventions*.[33] Specifically, we:

- Searched for and included nonrandomized studies, which may be more likely than randomized controlled trials to include diverse groups of patients, including members of equity-deserving groups.
- Included outcomes that were prioritized by the patient contributor and clinical and ethics experts who were consulted during this project. When data for prioritized outcomes were missing, we noted this as a limitation of the evidence.
- Investigated subgroup effects as informed by PROGRESS-Plus criteria[34,35] that were identified by the experts consulted during this project as important in the context of stroke care or AI.
- Used PROGRESS-Plus[34,35] to guide data extraction and examine who may or may not have been represented in the included studies.

Our methods for this review were guided by the Scottish Health Technologies Group's HTA framework.[26] When possible, we provided information addressing the items for each of the domains set out in the framework.[26]

This rapid review was conducted alongside an implementation review for AI-enabled medical devices (found on the project website).

Literature Review

Literature Search Methods

An information specialist conducted a literature search on key resources including MEDLINE, Embase, Scopus, the Cochrane Database of Systematic Reviews, the International HTA Database, the websites of Canadian and major international health technology agencies, as well as a focused internet search. The search approach was customized to retrieve a limited set of results, balancing comprehensiveness with relevancy. The search strategy was comprised of both controlled vocabulary, such as the

National Library of Medicine's MeSH (Medical Subject Headings), and keywords. Search concept was developed based on the elements of the research questions and selection criteria. The main search concept was the technology, RapidAI. No search filters were applied to limit the retrieval by study type. The search was completed on April 23, 2024, with no date or language limits. Monthly search alerts updated the database literature searches until June 24, 2024. These searches were supplemented by reviewing the bibliographies of key papers (e.g., related systematic reviews) and through contacts with the manufacturer.

Screening and Study Selection

The literature review management software DistillerSR (Evidence Partners, Ottawa, Ontario) facilitated screening and study selection. Pilot exercises were used on a selected sample of records to familiarize the reviewers with the selection criteria (Table 1) and to determine the appropriateness of single reviewer screening. In the first level of screening, 2 reviewers independently screened 20% of the titles and abstracts retrieved from the literature searches. Agreement between the reviewers was 93.5% (58 of 62 piloted citations); because agreement exceeded the 80% a priori threshold, 1 reviewer completed the remainder of the title and abstract screening.

Full texts of titles and abstracts that were judged to be potentially relevant were retrieved. During the pilot exercise, 2 reviewers independently assessed 20% of the full texts (11 articles) for inclusion based on the selection criteria. The selection decisions of the 2 reviewers during the full-text pilot were in full agreement. Therefore, 1 reviewer completed the selection of the remaining potentially eligible studies at the full-text level.

We applied a tiered approach when selecting studies for inclusion. For both the detection of ICH and LVO, we aimed to include studies assessing the effectiveness (including harms) and accuracy of RapidAI to support the review of CT scans by health care providers, as it would be used in clinical practice (i.e., RapidAI as an adjunct or aid to clinician interpretation versus clinician interpretation without the use of RapidAI). However, if we did not identify any relevant evidence that made this comparison, or if the identified studies only provided outcome data for a subset of patients who underwent imaging for suspected stroke, we planned to include studies that examined RapidAI as a standalone intervention and not alongside clinician interpretation. We identified evidence evaluating our initial comparison of interest for the detection of ICH. We also identified evidence evaluating this comparison for the detection of LVO, but these studies only provided outcome data for patients who underwent tissue plasminogen activator therapy or mechanical thrombectomy following diagnosis, and not for the entire population of patients who underwent imaging for suspected LVO (i.e., outcome data for false negative and true negative cases were not reported). Therefore, we also included studies comparing RapidAI as a standalone intervention for the detection of LVO.

Table 1: Selection Criteria

Criteria	Description
Population	People (any age) with a suspected acute stroke
Intervention	RapidAI to support the review of CT scans by health care providers to detect intracranial hemorrhage and large vessel occlusion; RapidAI used as a standalone intervention (i.e., not alongside clinician interpretation) to detect intracranial hemorrhage and large vessel occlusion[a]
Comparator or reference standard	CT scan review by health care providers alone (e.g., assessment by a single radiologist, consensus obtained from a panel of neuroradiologists)
Outcomes	Q1: Clinical effectiveness, including benefits and harms (e.g., functional status [e.g., mRS score], time to intervention,[b] mortality, recanalization rate, length of hospital stay, health-related quality of life, patient harms); health care resource implications (e.g., time usage or savings, cases diagnosed per unit time); diagnostic test accuracy (e.g., sensitivity, specificity, positive predictive value, negative predictive value, area under the ROC curve) Q2: Cost-effectiveness (e.g., cost per QALY gained)
Study designs	Q1: Randomized controlled trials and nonrandomized studies Q2: Economic evaluations
Publication date	No date limits

Language	English

CT = computed tomography; mRS = modified Rankin Scale; QALY = quality-adjusted life-year; ROC = receiver operating characteristic.
[a] RapidAI as a standalone intervention was only considered relevant when evidence on RapidAI to support clinician interpretation was unavailable.
[b] Includes measures such as time from door to imaging, time from door to needle (i.e., tissue plasminogen activator therapy), and time from door to groin puncture (i.e., thrombectomy).

Exclusion Criteria

We excluded studies not meeting the inclusion criteria outlined in Table 1, duplicate publications, and studies of any design published only as abstracts, conference proceedings, presentations, thesis documents, or pre-prints. For feasibility reasons, we also excluded studies published in non-English languages.

Additionally, we excluded studies that evaluated RapidAI for estimating ischemic core volumes, automatically calculating Alberta Stroke Program Early CT scores,[36] or predicting favourable or poor outcomes following intervention, as these are intended to inform treatment selection (e.g., selecting patients for reperfusion therapy) rather than determine whether stroke has occurred. We also excluded studies that evaluated the diagnostic accuracy of RapidAI in patient populations that included those who were imaged for non-stroke related indications (e.g., populations who were not suspected of having acute stroke or were described as healthy controls).

Data Extraction

We extracted data directly into tables created in Microsoft Word, which were developed, piloted, and modified, as necessary. In the pilot round, 2 reviewers independently extracted data from a sample of included studies (i.e., 1 cohort study and 1 diagnostic accuracy study), then met to resolve disagreements through discussion. Once both reviewers were satisfied with the content and usability of the tables, formal data extraction was performed by 1 reviewer, and a second reviewer independently verified the study characteristics and outcomes data for accuracy and completeness. Disagreements were resolved through discussion.

Relevant information that was extracted included study characteristics, methodology (e.g., study design), population, intervention, comparator, and results regarding the outcomes of interest. We used PROGRESS-Plus[34,35] to guide data extraction. Each included publication was checked to determine if relevant PROGRESS-Plus criteria were reported by study authors to describe the participants, and the reviewer made note of where these were not included. When available, detailed participant characteristics were extracted and reported in tables. The main PROGRESS-Plus criteria include place of residence, race, ethnicity, culture, language, occupation, gender, sex, religion, education, socioeconomic status, social capital, personal characteristics associated with discrimination (e.g., age, ability, disability), features of relationships (e.g. smoking spouse or partner), and time-dependent relationships (e.g., leaving the hospital, respite care, other instances where a person may be temporarily at a disadvantage).[34,35] Relevant PROGRESS-Plus criteria and outcomes of interest were selected because they were important for those who might be affected by the intervention, including the patient contributor and experts who were engaged or consulted during this project. We classified the designs of included diagnostic accuracy studies based on guidance from the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy[37] and Mathes and Pieper.[38]

When relevant data were conflicting in the included studies (e.g., there were discrepancies between values reported in the abstract and the main text of a publication), we reported all values and described the inconsistency. Due to the project's timelines, we did not attempt to contact the corresponding authors of included studies to obtain missing information or to clarify conflicting information. We did not extract data presented only as figures or graphs that would require manual estimation or extraction using image processing software.

Critical Appraisal of Individual Studies

The risk of bias for included studies was evaluated using the Risk Of Bias In Nonrandomized Studies – of Interventions (ROBINS-I) tool[39] for nonrandomized studies and the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) checklist[40] for diagnostic accuracy studies. Published guidance for each tool was followed.[39-41] For nonrandomized studies, we assessed the risk of bias in the effect of assignment to the intervention at the level of the reported effect (i.e., outcome-level). Where possible, we attempted to predict the direction of the potential bias. A rationale was provided for decisions about the risk of bias for both the

domain-level and overall assessments. Based on guidance from Risk Of Bias In Nonrandomized Studies – of Exposures (ROBINS-E),[42] we did not proceed with assessing other domains of bias for studies deemed at critical risk of bias due to baseline confounding or in the measurement of the outcome. To capture other considerations not considered in risk of bias assessments, we consulted additional tools and checklists for assessing methodological and reporting quality in studies of AI-enabled software for clinical decision support (e.g., APPRAISE-AI tool)[43] and medical imaging (e.g., CLAIM guideline).[44] Studies were not excluded from this review based on the critical appraisal results.

Two reviewers independently piloted the selected tools across at least 1 study of each included design and met to resolve any disagreements, to ensure a mutual understanding of the tools. After piloting, 1 reviewer completed the risk of bias assessments for the remaining studies. The second reviewer verified all judgements and justifications. Any disagreements were resolved through discussion

Data Analysis and Synthesis

Detailed descriptions of study characteristics, as well as the results of the risk of bias assessments, from eligible studies were provided in tables, together with a narrative summary in the main text. The study and patient characteristics were considered in the analysis of the effectiveness and accuracy within and across the studies. Throughout our analysis and synthesis, we attended to relevant ethics and equity considerations (e.g., algorithmic bias, lack of representation, data ownership, explainability) alongside evaluations of effectiveness and accuracy (further described in the Ethics and Equity Considerations section of this report within Methods).

A narrative synthesis was conducted as per existing guidance by Popay et al.[45] We first grouped studies that were similar in their design and population, intervention, comparator, and outcomes. Next, we developed a preliminary synthesis by organizing the results and identifying patterns in the size and direction of effects. We evaluated within- and between-study relationships and described our findings about the direction and magnitude of the observed effects. Outcomes were reported in the measurement units used by the study authors, and results were interpreted with due consideration for the differences in the instruments of assessment across studies. In some instances, we combined summary statistics (i.e., means and standard deviations) across 2 groups using standard formulae[46] to streamline the presentation of population characteristics across included studies. When not reported in publications, we calculated diagnostic accuracy metrics (e.g., sensitivity, specificity, negative predictive value, positive predictive value) with Clopper-Pearson exact 95% confidence intervals (CIs) from the available data (e.g., true positives, false positives, true negatives, false negatives) via the EpiR package in R.[47] Additionally, we calculated risk differences with 95% CIs for some binary outcomes when they were not reported in the study but were considered useful for informing assessments of certainty via the PropCIs package in R.[48]

Certainty of the Evidence

One reviewer assessed the overall certainty of the evidence for all outcome-comparisons using the methods of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group, following a minimally contextualized approach.[49-54] A second reviewer verified all assessments and rationales. Reviewers discussed discrepancies and reached consensus on the assessments.

Evidence from nonrandomized studies reporting on clinical benefits, harms, or resource use started at low certainty evidence acknowledging the risk of residual confounding and selection bias. Evidence for studies of diagnostic accuracy started at high certainty evidence. The certainty of evidence could be rated down for concerns due to risk of bias, inconsistency in effects, imprecision, indirectness, or publication bias. Evidence from nonrandomized studies could be rated up in the case of a large effect or if all plausible confounding is judged to reduce the size of the effect. For effectiveness results, where possible, we judged our certainty in a clinically important effect using thresholds informed by a clinical expert consulted during this project. When no reasonable thresholds could be determined, we judged our certainty in a non-null effect (i.e., using the null as the threshold). For diagnostic accuracy results, we judged our certainty in the point estimates for sensitivity and specificity. We rated down for imprecision if the boundaries of the 95% CIs included values that could lead to different conclusions about RapidAI's value based on input from a clinical expert consulted during this project.

When there were outcome-comparisons with data from multiple studies, the certainty of evidence was assessed using published guidance on the use of GRADE in the absence of a single estimate of effect,[55] as no quantitative syntheses were conducted. The results of GRADE assessments were reported in summary of findings tables. We used GRADE informative statements when describing our certainty in the results (i.e., "results in" for high certainty evidence, "likely results in" for moderate certainty evidence, "may result in" for low certainty evidence, and "very uncertain" for very low certainty evidence).[56]

Ethics and Equity Considerations

Ethics and equity considerations were considered as core elements of this review, and findings are integrated throughout the presentation of results. Alongside the application of the PROGRESS-Plus tool to investigate elements of health equity arising in this review, an analysis of additional ethics and equity considerations relevant to RapidAI was informed by key items arising in the Scottish Health Technology Group's HTA framework.[26] These included items related to the use and impact of the technology within health care systems, the safety, acceptability, and credibility of the technology, and expectations around the performance of the technology. The analysis also drew from and was framed by other foundational Ethics of AI tools and frameworks (e.g., UNESCO Recommendation on the Ethics of Artificial Intelligence,[57] WHO Guidance: Ethics and Governance of Artificial Intelligence for Health),[58] the ethical considerations proposed by the AI Task Force of the Society of Nuclear Medicine and Molecular Imaging,[59] as well as the EUnetHTA Core Model 3.0 Ethics Domain[60] and the Equity Checklist for HTA.[61]

Prompts and guiding principles from these sources were used to identify and reflect on ethics and equity considerations of RapidAI throughout the technology lifecycle relevant to patients, providers, and health care systems.

With the assistance of a reviewer with ethics expertise, the findings of these prompts and considerations of elements of PROGRESS-Plus were synthesized into analytic categories representing the key ethical and equity considerations related to RapidAI and digital health technologies more broadly.

Patient and Clinician Engagement

Invitation to Participate and Consent

Patient and clinician engagement is an important component of our projects, as it allows us to consider their experiences when writing our report. We disseminated a patient engagement request for individuals with lived experience of a hemorrhagic stroke through several large patient advocacy groups. We also sent an engagement request to several clinics that specialize in AI and use RapidAI, seeking a clinician to participate in a 1:1 engagement. Interested individuals — 1 patient and 1 clinician — responded to our outreach requests, and a Patient Engagement Officer conducted introductory discussions by email or Zoom. During these initial discussions, the Patient Engagement Officer described CDA-AMC and gave an overview of the purpose and scope of the project and the nature of the engagement. Both interested parties were invited to participate, and the interested patient agreed. The clinician declined due to time constraints.

The Patient Engagement Officer obtained informed consent from the patient contributor to participate in a discussion with CDA-AMC project team members and for a recording and summary of the discussion to be shared with the broader project team for their review. The patient contributor was offered a gift card as a gesture of thanks for her time and expertise and was offered the opportunity to be thanked by name in the report or to remain anonymous.

Engagement Activities

We invited the patient contributor to participate in an interview facilitated by the Patient Engagement Officer. Three members of the project team also attended. The purpose of attending the dialogue is for the project team members to hear directly from an individual with lived experience of a hemorrhagic stroke and ask questions relating to what they are reading in the literature, including AI use in health care. This offered insights to the project team members and allowed for a more nuanced understanding of the literature.

With the patient contributor's consent, we recorded the dialogue for note-taking purposes, and so that other members of the project team could review and learn from the conversation. We structured the interview into 2 parts: (i) the patient contributor's lived

experience with a stroke, and (ii) AI in the interpretation of imaging results. Patient involvement was guided by the Guidance for Reporting Involvement of Patients and the Public (version 2) Short Form reporting checklist,[62] which is outlined the Supporting Information document (found on the project website; refer to *Patient Engagement*).

The Patient Engagement Officer subsequently drafted a summary of the conversation and sent it to the patient contributor for review and approval. The summary was used as a prompt for the authors of this report as they were drafting the report and was not published. To inform this review, we used the summary to extract any discussion points or themes related to RapidAI for stroke detection considerations.

External Review

Peer Review

Before the review phase began, 1 clinical expert with expertise in stroke assessment, 1 clinical expert with expertise in AI radiology, and 1 ethics expert with expertise in AI reviewed the project plan. The same experts will review this draft version of the report, and their feedback will be incorporated into the final version of the report.

Manufacturer Review

The manufacturer reviewed a preliminary list of included studies. The manufacturer will review this draft version of the report, and their feedback will be incorporated into the final version of the report.

Feedback Opportunity

This draft version of the report will be posted on the CDA-AMC's website. This allows for interested parties the opportunity to provide feedback on the draft report. All feedback will be considered, and, where appropriate, revisions will be made to the draft report and reflected in the final version.

Summary of Evidence

Quantity of Research Available

We identified a total of 309 citations via the electronic literature search and excluded 242 records following screening of titles and abstracts. We retrieved 20 additional potentially relevant publications from the grey literature search, from the search alerts, by reviewing bibliographies of key papers, and through contacts with the manufacturer. From full-text review of the 87 potentially relevant articles, we excluded 76 for various reasons, and 11 publications met the inclusion criteria and were included in this report. These comprised 2 cohort studies and 11 diagnostic accuracy studies described in 11 publications (2 publications assessed clinical outcomes before and after implementing RapidAI using a non-concurrent cohort study design and diagnostic accuracy outcomes using a cross-sectional study design). A Preferred Reporting Items for Systematic Reviews and Meta Analyses (PRISMA)[63] flow chart that shows the study selection process, and lists of included and excluded studies, with details describing the rationale for those excluded, are presented in the Supporting Information document (found on the project website; refer to *Selection of Included Studies*, *List of Included Publications*, and *List of Excluded Publications and Reasons for Exclusion*).

Summary of Study Characteristics

Additional details regarding the characteristics of included studies are provided the Supporting Information document (found on the project website; refer to *Characteristics of Included Publications*).

Included Studies for Question 1: Effectiveness and Accuracy

We identified 2 retrospective non-concurrent cohort studies[64,65] and 11 cross-sectional diagnostic accuracy studies[64-74] to address this research question. The 2 cohort studies[64,65] and 1 diagnostic accuracy study[70] examined the effectiveness or accuracy of

RapidAI for supporting health care providers' review of imaging results, while 10 diagnostic accuracy studies[64-69,71-74] evaluated RapidAI as a standalone intervention (i.e., without clinician interpretation).

The cohort studies were conducted in the US, at a single comprehensive stroke centre[64] or including patients from a large multi-hospital network with several comprehensive stroke centres.[65] Seven diagnostic accuracy studies were single-centre and conducted in Australia,[67,74] the UK,[68,71] or the US.[64,70,72] The remaining diagnostic accuracy studies were multi-centre and included patients from several hospitals in Australia[66] or the US[65] or included patient data from a variety of sources (such as recent cerebrovascular trials and hospitals in multiple countries).[69,73]

Patient Population

Across all studies, populations consisted of patients presenting with neurological deficit or suspected acute stroke, all of whom underwent diagnostic imaging such as non-contrast CT, CT angiography, CT perfusion, or a combination of imaging techniques (i.e., multimodal imaging). The mean or median ages of participants ranged from 61[70] to 75[64] years. One cohort study[69] and 1 diagnostic accuracy study[70] did not report the age of participants.

The reporting of sex and gender of study participants varied across the included studies (described using the original terms provided by the study authors):

- Eight studies (described in 7 publications)[64,65,68,71-74] reported the proportion of female participants, which ranged from 41.3%[68] to 55.2%.[64]
- Five studies[65,66,71,73,74] reported the proportion of male participants, which ranged from 44.4%[66] to 56.8%.[74]
- Two studies[68,74] reported the proportion of participants who were women, which ranged from 41.3%[68] to 43.2%.[74]
- Two studies[67,74] reported the proportion of participants who were men, which ranged from 49.6%[67] to 56.8%.[74]
- Three studies[65,69,70] did not report the sex or gender of study participants.

None of the included studies provided information on how sex and gender were defined or measured, nor did they report on sex or gender categories beyond female, male, women, or men.

Two studies (described in 1 publication)[64] reported participant race but did not indicate how it was recorded. No other PROGRESS-Plus criteria[34,35] (including those identified by the clinical experts consulted during this project as important in the context of stroke care or AI) were reported, such as place of residence, ethnicity, culture, language, occupation, religion, faith, spirituality, education, socioeconomic status, social capital, or disability status. One study[73] presented subgroup analyses for some PROGRESS-Plus criteria,[34,35] including for age (20 to 29 years versus 40 to 59 years versus ≥ 60 years) and sex (female versus male).

RapidAI

The included studies examined various versions and modules of RapidAI, including Rapid ICH (version not reported), Rapid LVO (versions 1.0 or 5.2.2, or as part of Rapid 4.9), Rapid CTA (as part of Rapid 4.9 or RapidAI 5.1, or version not reported), RapidAI (version not reported), and the Rapid NCCT Storke platform (using Rapid HVS to detect LVO). Some of the included studies[64,66,67,74] provided high-level overviews of RapidAI's process for analyzing CT images to detect ICH or LVO. For example, Delora and colleagues[66] explained that Rapid CTA detects LVO by aligning CT angiography with a known reference image and removing bone tissue based on voxel location and brightness. An algorithm then calculates the vessel density using brightness values, compares them to the opposite side of the brain, and applies a threshold to classify images as LVO-positive or LVO-negative. However, there was limited descriptions of the types of AI models used by RapidAI and their structures (e.g., inputs, outputs, intermediate layers, and connections). Consequently, we cannot assess the appropriateness of the methods used to train the machine learning models, and based on the literature we reviewed, there is limited information to help users understand how RapidAI models were developed and what factors influence their performance.

Outcomes

Measures of effectiveness included time to intervention metrics (e.g., time from door to groin puncture for thrombectomy), functional status (e.g., modified Rankin Scale [mRS] scores), and response to therapy (i.e., using the Thrombolysis in Cerebral Infarction [TICI] scale).

The diagnostic accuracy studies[64-74] reported various parameters of diagnostic performance, including sensitivity, specificity, positive predictive value, and negative predictive value, along with the numbers of true positives, false negatives, false positives, and true negatives. One study[74] also reported the area under the receiver operating characteristic curve for detecting different types of LVO.

None of the included studies reported on measures of mortality, length of hospital stay, health-related quality of life, or patient harms (e.g., administration of harmful therapies or undertreatment due to inaccurate diagnosis), which were identified as important by the clinical experts and patient contributor who were consulted or engaged during this project. None of the included studies reported on health care resource implications.

Included Studies for Question 2: Cost-Effectiveness

We did not identify any studies that evaluated the cost-effectiveness of RapidAI (with or without clinician interpretation) to detect ICH or LVO.

Summary of Critical Appraisal

Full details are provided in the Supporting Information document (found on the project website; refer to *Critical Appraisal of Included Studies*). Overall, the cohort studies exhibited critical risk of bias across all outcome domains. Ten diagnostic accuracy studies[64-73] were assessed as having a high or unclear risk of bias in at least 1 domain, with most studies having high or unclear risk of bias for multiple domains. One diagnostic accuracy study[74] exhibited low risk of bias across all 4 domains. For applicability concerns, 10 studies[64-69,71-74] were labelled as high risk as they evaluated CT images by RapidAI alone, which differs from our primary review question and how RapidAI is used in clinical practice (i.e., alongside clinician interpretation).

Risk of Bias in Nonrandomized Studies

In both cohort studies,[64,65] there were no adjustments for confounding variables (e.g., age, sex or gender, race, ethnicity, disease severity, comorbidities, neurologist experience) across most or all reported outcomes, including time to intervention and measures of response to therapy. Soun and colleagues[64] applied adjustments for some measures of functional status, including for the effects of high cholesterol, heart disease, atrial fibrillation, therapies received, and NIHSS score on admission. However, the authors did not provide a rationale for how these factors were selected and did not describe how these confounding variables were measured. It was unlikely that these adjustments adequately controlled for all potential sources of baseline confounding. Consequently, both cohort studies[64,65] were judged to be at critical risk of bias across all outcome domains, and we did not proceed with detailed risk of bias assessments.

Risk of Bias in Diagnostic Accuracy Studies

Seven studies[64,67,68,70-72,74] exhibited low risk of bias for patient selection, as they enrolled consecutive samples of patients suspected of having stroke while minimizing inappropriate exclusions, creating representative samples. Four studies were at high or unclear risk of bias for patient selection, as they enrolled patients based on the results of the reference standard,[73] applied inappropriate exclusion criteria (e.g., only included patients who tested positive on the index test),[65] or their methods for selecting patients were unclear.[66,69] None of the included studies exhibited applicability concerns regarding patient selection, as the patient populations from each study were directly relevant to the review question (i.e., people with suspected stroke).

Four studies[65,67,69,74] demonstrated a low risk of bias for the conduct and interpretation of the index test. In these studies, clinicians interpreted the index test results without knowledge of the reference standard results, and they likely pre-specified the relative vessel density thresholds for detecting LVO. Seven studies[64,66,68,70-73] exhibited an unclear risk of bias for the index test because there was insufficient information to determine whether they had pre-specified positivity thresholds for detecting LVO or ICH. For applicability

concerns, the index test domain for 1 study[70] was labelled as low risk because the execution and interpretation of the index test reflected its use in practice (i.e., RapidAI assisting clinician interpretation of CT images). However, 10 studies[64-69,71-74] were labeled as high risk because they evaluated CT images by RapidAI alone, which differs from our primary review question and how RapidAI is used in clinical practice (i.e., alongside clinician interpretation).

The reference standard domains of 5 studies[68,69,72-74] were judged to be at low risk of bias as the reference standards were likely to correctly classify the target condition and were interpreted without knowledge of the results of the index tests. Three studies[64,66,70] exhibited high risk of bias. In 2 of these studies,[64,66] those interpreting the reference standard were not blinded to the results of the index test, which may bias estimates of agreement between the index test and the reference standard. The index test served as the reference standard for some patients in the study by Eldaya and colleagues,[70] which is likely to inflate the estimates of diagnostic accuracy for the index test. Three studies[65,67,71] were at unclear risk of bias as the authors did not report whether interpretation of the reference standard was independent of the index test. Across all 11 studies,[64-74] there were no applicability concerns for the reference standard domain, as the target conditions were likely to be correctly classified by the reference standards and were directly relevant to the review. However, there was variability in the composition of the reference standards used. Although the rates of misclassification by the reference standards were excepted to be low, they could have been influenced by several factors, including the number of radiologists involved (e.g., a single radiologist versus a panel of radiologists), their level of training (e.g., years of experience), the types of imaging results available (e.g., CT angiography versus CT angiography and CT perfusion), and whether reference standard assessors had access to clinical history and other diagnostic tests results (e.g., blood tests). Because the studies used imperfect reference standards, it is possible that the diagnostic performance of RapidAI may not accurately reflect situations where RapidAI correctly classified patients, but the reference standard did not.

Seven studies[65-67,71-74] exhibited low risk of bias for the flow and timing domain as all patients received the same reference standard and all or nearly all patients were included in the analyses. Two studies were at high risk of bias related to flow and timing as not all patients received the same reference standard[70] or they excluded a considerable number of patients from the analysis.[68] The risk of bias for the flow and timing domain was unclear in 2 studies[64,69] because there was insufficient information to determine whether all patients were included in the analyses.

Table 2 presents a summary of the risk of bias assessments for the 11 diagnostic accuracy studies.[64-74]

Table 2: Summary of Risk of Bias in the Included Diagnostic Accuracy Studies Using QUADAS-2[40]

Study Citation	Risk of bias				Applicability concerns		
	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
LVO							
Delora et al. (2024)[66]	Unclear	Unclear	High	Low	Low	High	Low
Slater et al. (2024)[67]	Low	Low	Unclear	Low	Low	High	Low
Chan et al. (2023)[68]	Low	Unclear	Low	High	Low	High	Low
Soun et al. (2023)[64]	Low	Unclear	High	Unclear	Low	High	Low
Yedavalli et al. (2023)[69]	Unclear	Low	Low	Unclear	Low	High	Low
Mallon et al. (2022)[71]	Low	Unclear	Unclear	Low	Low	High	Low
Schlossman et al. (2022)[72]	Low	Unclear	Low	Low	Low	High	Low
Adhya et al. (2021)[65]	High	Low	Unclear	Low	Low	High	Low
Dehkharghani et al. (2021)[73]	High	Unclear	Low	Low	Low	High	Low
Amukotuwa et al. (2019)[74]	Low	Low	Low	Low	Low	High	Low
ICH							
Eldaya et al. (2022)[70]	Low	Unclear	High	High	Low	Low	Low

ICH = intracerebral hemorrhage; LVO = large vessel occlusion; QUADAS-2 = Quality Assessment of Diagnostic Accuracy Studies 2.

Additional Methodological Considerations

The included studies generally adhered to the reporting guidelines outlined in the APPRAISE-AI tool[43] and the CLAIM guideline.[44] For example, the authors often indicated in the title that their study evaluated an AI-enabled tool for stroke. They provided background information on the clinical issue, specified whether the study was prospective or retrospective, clearly outlined the objectives of the study, described their sources of data, patient eligibility criteria, interventions, comparators (or reference standards), and main outcomes. The authors also contextualized their results, acknowledged limitations of their approach, provided conclusions with appropriate caveats, and disclosed relevant financial relationships and potential conflicts of interest.

However, several items from the APPRAISE-AI tool[43] and the CLAIM guideline[44] were underreported across all included studies, particularly items that are specific to AI-enabled digital health technologies during the model development phase, including:

- Methods used for developing final datasets, including data abstraction, cleaning, imputation, and preparation.
- Descriptions of the models and the software libraries used.
- Details of the training approach, including data augmentation and hyperparameter tuning.
- The process for splitting data into training, testing, and validation partitions.
- Approaches for model evaluation, including the metrics used to assess performance and calibration, bias assessments, sensitivity analyses, and error analyses.

According to the APPRAISE-AI tool[43] and the CLAIM guideline,[44] reporting these aspects helps to promote clear and transparent scientific communication about AI applications in health care, improving investigator accountability and increasing the overall quality of AI research. The principle of transparency as it relates to AI algorithms is further described in the Ethics and Equity Considerations section of this report within Summary of Evidence.

Summary of Findings

A detailed overview of the main study findings is presented in the Supporting Information document (found on the project website; refer to *Main Study Findings*).

Clinical Effectiveness of RapidAI for Stroke Detection

RapidAI to support the review of CT scans by health care providers for stroke detection

Table 3 presents the findings for time to intervention outcomes, functional status, and response to therapy based on evidence from 2 cohort studies[64,65] (1 evaluating clinician interpretation of imaging results with Rapid LVO and the other evaluating clinician interpretation of imaging results with Rapid CTA). Outcome measures were heterogeneous across the studies, and conclusions were often limited due to critical risk of bias and imprecision.

Among patients presenting with acute ischemic stroke, CT angiography with Rapid LVO may reduce radiology report turnaround time (1 study;[64] ≤ 760 patients; low certainty) by 8.6 (SD = 32.8) minutes, which was considered clinically important by the clinical expert we consulted. It is very uncertain whether Rapid LVO has any effect on door to intervention times (1 study;[64] ≤ 105 patients; very low certainty) and whether Rapid CTA has any effect on the time interval between CT angiography and groin puncture for thrombectomy (1 study;[65] 146 patients; very low certainty).

The evidence is very uncertain about the effects of both Rapid CTA (1 study;[65] 141 patients) and Rapid LVO (1 study;[65] ≤ 105 patients) on measures of functional status. It is also very uncertain whether Rapid LVO has any effect on response to therapy (1 study;[64] 80 patients).

Outcomes related to patient harms (e.g., rate of symptomatic intracranial hemorrhage) were not measured in the included studies but were identified as critically important by the clinical expert we consulted.

Table 3: Summary of Findings for Clinician Interpretation of CTA Imaging With RapidAI Versus Clinician Interpretation of CTA Imaging Without RapidAI for People With Suspected Acute Stroke

Outcome and follow-up	Intervention	Participants (studies), N	Absolute effects			Certainty[a]	What happens
			Without RapidAI	With RapidAI	Difference		
Time to intervention							
Radiology report turnaround (minutes), mean (SD)[b]	Rapid LVO (Rapid v4.9)	≤760[c] (1 NRS)[64]	30.6 (29.9)	22 (35.1)	−8.6 (32.8)[d,e]	Low *(due to risk of bias)*	CTA with Rapid LVO may result in a clinically important reduction in radiology report turnaround time.
Door to intervention (minutes), median (IQR)	Rapid LVO (Rapid v4.9)	≤105[c] (1 NRS)[64]	Door to intervention (without vs. with Rapid LVO): • Door to image: 11 (8 to 20) vs. 13 (7 to 20) • Door to intubation: 65 (46 to 73) vs. 70 (58 to 86) • Door to needle (tPA): 37 (26 to 44) vs. 42 (30 to 53) • Door to puncture: 97 (80 to 107) vs. 101 (90 to 113) • Door to revascularization: 155 (123 to 197) vs. 158 (131 to 192) Between-group differences were NR for any outcome.			Very low *(due to risk of bias and imprecision)*	The evidence is very uncertain about the effect of CTA with Rapid LVO on door to intervention times.
CTA to groin puncture (minutes), mean (SD)	Rapid CTA (version NR)	146 (1 NRS)[65]	92 (NR)	68 (NR)	−24 (NE)[d,f]	Very low *(due to risk of bias and imprecision)*	The evidence is very uncertain about the effect of CTA with Rapid CTA on CTA to groin puncture times.
Functional status							
Neurological deficit (per NIHSS score), median (IQR)[g,h]	Rapid LVO (Rapid v4.9)	NR[c] (1 NRS)[64]	NIHSS scores (without vs. with Rapid LVO): • 36-hours post-treatment: 10 (5 to 18) vs. 11 (2 to 20) • At discharge: 5 (1 to 100) vs. 8 (2 to 20) • Change from admission to discharge: −7 (−2 to −13) vs. −3 (0 to −7) Between-group differences were NR for any outcome.			Very low *(due to risk of bias and imprecision)*	The evidence is very uncertain about the effect of CTA with Rapid LVO on post-treatment neurological deficit.
Proportion of patients with significant morbidity or	Rapid LVO (Rapid v4.9)	105 (1 NRS)[64]	177 per 1,000 (NR)	233 per 1,000 (NR)	55 more per 1,000 (103 less to 213 more)[i]	Very low *(due to risk of bias and imprecision)*	The evidence is very uncertain about the effect of CTA with Rapid LVO on the proportion of patients

Outcome and follow-up	Intervention	Participants (studies), N	Absolute effects			Certainty[a]	What happens
			Without RapidAI	With RapidAI	Difference		
mortality (defined as mRS score ≥ 5) at discharge (95% CI)[h,i]							with significant morbidity or mortality.
Disability and dependence in daily activities (per 90-day mRS score), mean (SD)[h,i]	Rapid CTA (version NR)	141 (1 NRS)[65]	4.47 (NR)	3.90 (NR)	−5.7 (NE)[d]	Very low *(due to risk of bias and imprecision)*	The evidence is very uncertain about the effect of CTA with Rapid LVO on disability and dependence in daily activities.
Proportion of patients considered to be functionally independent (defined as mRS score ≤ 2) at 90 days (95% CI)[h,i]	Rapid CTA (version NR)	141 (1 NRS)[65]	230 per 1,000 (NR)	343 per 1,000 (NR)	114 more per 1,000 (35 less to 262 more)[j]	Very low *(due to risk of bias and imprecision)*	The evidence is very uncertain about the effect of CTA with Rapid LVO on the proportion of patients considered to be functionally independent.
Response to therapy							
Proportion of patients with TICI scores of 0, 1, 2A, 2B/C, or 3[k]	Rapid LVO (Rapid v4.9)	80 (1 NRS)[64]	TICI scores (without versus with Rapid CTA): • Score of 0: 43 vs. 88 per 1,000 (difference, 45 more per 1,000) • Score of 1: 22 vs. 0 per 1,000 (difference, 22 less per 1,000) • Score of 2A: 86 vs. 118 per 1,000 (difference, 32 more per 1,000) • Score of 2B/C: 391 vs. 500 per 1,000 (difference, 109 more per 1,000) • Score of 3: 457 vs. 294 per 1,000 (difference, 163 less per 1,000)			Very low *(due to risk of bias and imprecision)*	The evidence is very uncertain about the effect of CTA with Rapid LVO on the proportion of patients with TICI scores of 0, 1, 2A, 2B/C, or 3.
Patient Harms							
NR	—	—	No data available.			NA	There is no evidence for the effect of CTA with RapidAI on patient harms.

CI = confidence interval; CTA = computed tomography angiography; IQR = interquartile range; LVO = large vessel occlusion; mRS = modified Rankin Scale; NA = not applicable; NE = not estimable; NIHSS = National Institutes of Health Stroke Scale; NR = not reported; NRS = nonrandomized study; SD = standard deviation; TICI = Thrombolysis in Cerebral Infarction; tPA = tissue plasminogen activator.

[a] Detailed reasons for certainty of evidence ratings are provided in the Supporting Information document (found on the project website; refer to *Reasons for Certainty of Evidence Ratings*).

[b] Defined as the time from when the CTA images are available for the radiologist to the earlier time of either the report being available or read-back verification was provided for the clinicians.

[c] The sample size for the analysis was not explicitly reported; as such, the amount of missing data is unknown.

[d] NR in the study. Imputed by the review team using standard calculations per the Cochrane Handbook for Systematic Reviews of Interventions,[46] assuming a correlation coefficient of 0.5.

[e] A difference of 5 to 10 minutes between groups was identified by the clinical expert we consulted as a threshold of clinical importance for this outcome.

[f] A difference of 10 to 20 minutes between groups was identified by the clinical expert we consulted as a threshold of clinical importance for this outcome.

[g] The National Institutes of Health Stroke Scale is a 15-item neurological examination stroke scale used for evaluating stroke-related neurological deficit. Total scores range from 0 to 42, with higher scores indicating more severe neurological deficit.

[h] Based on the information provided in the publication, this analysis appears to only include patients who received acute therapies for stroke (e.g., tPA therapy, thrombectomy). Therefore, outcomes for all patients who were assessed by RapidAI are not captured, including those who may have been misdiagnosed.

[i] The modified Rankin Scale is a clinician-reported tool for measuring the degree of disability and dependence in daily activities in people who have experienced stroke. Scores range from 0 (no symptoms at all) to 6 (death). A higher score indicates higher disability.

[j] The risk difference (95% CI) was NR in the study; it was imputed by the review team from the available data via the PropCIs package in R.[48]

[k] The Thrombolysis in Cerebral Infarction (TICI) scale is a grading system used to evaluate the degree of perfusion obtained following recanalization of an arterial occlusion. The TICI scale ranges from 0 (no reperfusion) to 3 (complete reperfusion).

Source: Soun et al. (2023) and Adhya et al. (2021).[65]

Diagnostic Accuracy of RapidAI for Stroke Detection

This section provides a summary of measures of diagnostic performance reported in the included diagnostic accuracy studies. Additional results for other outcomes (e.g., positive predictive value, negative predictive value, area under the receiver operating characteristic curve) are presented in the Supporting Information document (found on the project website; refer to *Main Study Findings*).

RapidAI to support the review of CT scans using clinician interpretation (or clinician consensus) as the reference standard

Table 4 presents the findings for the diagnostic accuracy of Rapid ICH with clinician interpretation for detecting ICH using non-contrast CT in people with suspected acute stroke. The sensitivity and specificity were 92% (95% CI, 78% to 98%) and 100% (95% CI, 98% to 100%), respectively (1 study;[70] 307 patients). The certainty of evidence was low, owing to risk of bias and imprecision.

The study by Eldaya et al. (2022)[70] was the only study that examined the diagnostic accuracy of RapidAI as it would be used in clinical practice (i.e., as a tool to support the review of CT scans). The included cohort studies[64,65] did not provide any direct measures of patient harms; however, the rates of false positives and false negatives can be used to infer when patient harms may occur, as they can lead to inappropriate treatment decisions resulting in unnecessary interventions or missed opportunities for timely care. Out of 307 test results included in the analysis by Eldaya and colleagues,[70] Rapid ICH with neuroradiologist interpretation resulted in 1 false positive (0.3% of total tests) and 3 false negative results (1% of total tests) when detecting ICH on non-contrast CT against a reference standard of neuroradiologist interpretation in cases of concordant results or consensus diagnosis by a panel of 3 neuroradiologists in cases of discordant results. The ethical implications of false positives and false negatives are discussed further in the Ethics and Equity Considerations section of this report within Summary of Evidence.

Table 4: Summary of Findings for the Diagnostic Accuracy of RapidAI with Clinician Interpretation Relative to Clinician Interpretation (or Clinician Consensus) for Suspected Acute Stroke

Index Test	Participants (studies), N	% positive per reference standard (range)	Sensitivity range (95% CI range)	Specificity range (95% CI range)	Certainty[a]
ICH					
Rapid ICH (version NR) with interpretation by a neuroradiologist	307 (1 study)[70]	12%	92% (78% to 98%)	100% (98% to 100%)	Low *(due to risk of bias and imprecision)*

CI = confidence interval; ICH = intracranial hemorrhage; NR = not reported.

[a] Detailed reasons for certainty of evidence ratings are provided in the Supporting Information document (found on the project website; refer to *Reasons for Certainty of Evidence Ratings*).

Source: Eldaya et al. (2022).[70]

RapidAI as a standalone intervention using clinician interpretation (or clinician consensus) as the reference standard

Table 5 summarizes the diagnostic accuracy of RapidAI as a standalone intervention compared to clinician interpretation (or clinician consensus) for detecting LVO on CT images (i.e., non-contrast CT or CT angiography). There was heterogeneity across the included studies for the types of LVO assessed, the versions of RapidAI examined, and reference standards used.

For each of these assessments, we judged our certainty in the comparison without rating down for indirectness or applicability concerns (based on QUADAS-2 assessments). We acknowledge that RapidAI is not used as a standalone diagnostic tool in practice, but as a supportive tool for clinicians. Consequently, these findings inform the accuracy of RapidAI when used independently and are unlikely to be directly applicable to clinical practice (where clinicians would have the opportunity to either refute the RapidAI results or change their own interpretation based on RapidAI results).

For detecting LVO of the M1 segment of the middle cerebral artery and the internal carotid artery using CT angiography, Rapid LVO as part of RapidAI v4.9 (2 studies;[64,72] 574 patients), Rapid LVO v1.0 (1 study;[73] 217 patients), Rapid CTA as part of RapidAI v4.9 (1 study;[74] 477 patients), and Rapid CTA (version not reported) (1 study;[68] 88 patients) had sensitivity values ranging from 90% to 96%, and specificity values ranging from 76% to 98%. Across these 5 studies,[64,68,72-74] the proportion of false positives and false negatives ranged from 1% to 13% and 1% to 20% of total tests, respectively (1 study[64] did not report the number of false positives and false negatives). The certainty of evidence was moderate or low, due to risk of bias or imprecision.

For the remaining comparisons, estimates of sensitivity ranged between 62% and 92%, while specificity ranged from 65% to 93%. When reported, the proportion of false positives ranged from 10 to 48% of total tests, and the proportion of false negatives ranged from 0% to 10% of total tests. Differences in the types of LVOs included, versions of RapidAI tested, study populations, and reference standards used may have contributed to the variability in results. The certainty of the evidence for these findings was low or very low, or there was insufficient information to judge. The certainty of the evidence was rated down due to imprecision or risk of bias and imprecision.

In addition to the primary analyses reported above, one study[73] included subgroup analyses to evaluate the diagnostic accuracy of Rapid LVO v1.0 for detecting LVO of the M1 segment of the middle cerebral artery and the internal carotid artery using CT angiography. Their findings indicated that the accuracy of Rapid LVO appeared similar across subgroups for age (20 to 29 years versus 40 to 59 years versus ≥ 60 years) and sex (female versus male). However, the study may not have been powered to detect subgroup differences, there were no tests for subgroup differences, and many other patient characteristics may impact the performance of RapidAI but were not investigated. The overall lack of similar and more robust investigations in the available evidence has potential implications for ethics and equity considerations related to inclusiveness and algorithmic bias, as it is unclear whether the accuracy of RapidAI is robust across diverse patient populations.

Table 5: Summary of Findings for the Diagnostic Accuracy of RapidAI Alone Relative to Clinician Interpretation (or Clinician Consensus) for Suspected Acute Stroke

Index Test	Participants (studies), N	% positive per reference standard (range)	Sensitivity range (95% CI range)	Specificity range (95% CI range)	Certainty[a]
M1 MCA and ICA LVO					
Rapid LVO (RapidAI v4.9) alone	574 (2 studies)[64,72]	11% (NR in 1 study)	90% to 96% (73% to 98%)[b]	85% to 86% (80% to 90%)[b]	Low *(due to risk of bias and imprecision)*
Rapid LVO (v1.0) alone	217 (1 study)[73]	50%[c]	96% (91% to 99%)	98% (94% to 100%)	Moderate *(due to risk of bias)*
Rapid CTA (RapidAI v4.9) alone	477 (1 study)[74]	16%	94% (86% to 98%)	76% (72% to 80%)	Moderate *(due to imprecision)*
Rapid CTA (version NR) alone	88 (1 study)[68]	15%	92% (64% to 100%)	85% (75% to 92%)	Low *(due to risk of bias and imprecision)*
Rapid NCCT Stroke platform (version NR) alone[d]	244 (1 study)[69]	47%	63% (54% to 72%)	95% (88% to 98%)	Low *(due to risk of bias and imprecision)*
M1 and M2 MCA and ICA LVO					
RapidAI alone (version NR)	84 (1 study)[71]	73%	74% (61% to 84%)	65% (43% to 84%)	Low *(due to risk of bias and imprecision)*
Rapid CTA alone (RapidAI v4.9)	477 (1 study)[74]	22%	92% (85% to 96%)	81% (77% to 85%)	Low *(due to risk of bias and imprecision)*

Index Test	Participants (studies), N	% positive per reference standard (range)	Sensitivity range (95% CI range)	Specificity range (95% CI range)	Certainty[a]
M1 MCA LVO					
RapidAI alone (version NR)	73 (1 study)[71]	51%	89% (75% to 97%)	77% (60% to 90%)	Low *(due to risk of bias and imprecision)*
Rapid LVO alone (as part of RapidAI v4.9)	247 (1 study)[74]	9%	91% (72% to 99%)	86% (80% to 90%)	Low *(due to risk of bias and imprecision)*
M2 MCA LVO					
Rapid LVO (RapidAI v4.9) alone	NR (1 study)[72]	NR	80% (NR)	NR	Insufficient information to judge
Rapid CTA (RapidAI v4.9) alone	477 (1 study)[74]	6%	86% (67% to 96%)	68% (63% to 72%)	Low *(due to risk of bias and imprecision)*
ICA LVO					
Rapid LVO (RapidAI v4.9) alone	235 (1 study)[72]	5%	82% (48% to 98%)	86% (80% to 90%)	Low *(due to risk of bias and imprecision)*
LVO of the ICA, M1 or M2 MCA, basilar artery, or intracranial vertebral artery					
Rapid LVO alone (RapidAI v5.1)	500 (1 study)[67]	13%	62% (48% to 75%)	93% (90% to 95%)	Low *(due to risk of bias and imprecision)*
Undefined LVO[e]					
Rapid LVO (v5.2.2) alone	360 (1 study)[66]	13%	87% (74% to 95%)	85% (80% to 88%)	Low *(due to risk of bias and imprecision)*
Rapid CTA (version NR) alone at < 45% relative vessel density threshold	310 (1 study)[65]	52%	80% (73% to 86%)	71% (63% to 78%)	Very low *(due to risk of bias and imprecision)*

CI = confidence interval; CTA = computed tomography angiography ICA = internal carotid artery; LVO = large vessel occlusion; MCA = middle cerebral artery; NCCT = non-contrast computed tomography; NR = not reported.

[a] Detailed reasons for certainty of evidence ratings are provided in the Supporting Information document (found on the project website; refer to *Reasons for Certainty of Evidence Ratings*).

[b] The 95% CI was not reported and not calculable in 1 study.

[c] The study by Dehkharghani et al. (2021) used case-control selection; as such, 50% of patients had LVO.

[d] Unlike other index tests for detecting LVO, the Rapid NCCT Stroke platform uses NCCT images (rather than CTA, which is typically required for confirming LVO).[69]

[e] Unclear which types of LVO were considered eligible. In Delora et al. (2024), the study population included patients with occlusions of the ICA, M1 MCA segment, and M2 MCA segment. The study population from Adhya et al. (2021) included patients with occlusions of the ICA, carotid terminus, M1 MCA segment, and M2 MCA segment.

Source: Delora et al. (2024),[66] Slater et al. (2024),[67] Chan et al. (2023),[68] Soun et al. (2023),[64] Yedavalli et al. (2023),[69] Mallon et al. (2022),[71] Schlossman et al. (2022),[72] Adhya et al. (2021),[65] Dehkharghani et al. (2021),[73] and Amukotuwa et al. (2019).[74]

Cost-Effectiveness of RapidAI for Stroke Detection

No relevant evidence was identified regarding the cost-effectiveness of RapidAI for the detection of ICH or LVO in people with suspected stroke; therefore, no summary can be provided.

Ethics and Equity Considerations

Several ethics and equity considerations related to the use of AI for detecting stroke can be drawn from our summary and analysis of the clinical evidence. We primarily leveraged and adapted the WHO Guidance on Ethics and Governance of Artificial Intelligence for Health,[58] to organize and reflect on these considerations and their implications. We focused our discussion on the 4 most relevant and applicable of the 6 consensus principles to ensure AI benefits the public identified in this Guidance (i.e., protect autonomy, promote human well-being, human safety and the public interest, ensure transparency, explainability and intelligibility, and ensure inclusiveness and equity). The 2 remaining consensus principles focus on fostering responsibility and accountability and promoting artificial intelligence that is responsive and sustainable, which are described in the accompanying implementation review (found on the project website). While these considerations are relevant to the implementation of RapidAI, we found insufficient information to address them adequately in this review. Specifically, we are unable to report on how patients and clinicians were involved in the development of RapidAI, who holds accountability if issues arise during its use, and how the developers are continuously, systematically, and transparently monitoring RapidAI to determine whether it is working according to its expectations. Additionally, we drew from the ethical considerations proposed by the AI Task Force of the Society of Nuclear Medicine and Molecular Imaging,[59] whose recommendations on the major ethical considerations during the deployment of AI-enabled medical devices are directly relevant to the use of AI in diagnostic workups.

Many ethical considerations related to AI in health care, as outlined by the Scottish Health Technology Group's HTA framework[26] and other foundational ethics of AI tools and frameworks, are often inadequately addressed in studies evaluating the effectiveness or accuracy of commercialized AI-enabled medical devices (i.e., the types of evidence included in this review). Instead, these ethical questions often arise at both earlier and later stages of the technology lifecycle. For example, during the initial design phase, issues such as algorithmic bias and lack of representation can occur. Later in the technology lifecycle, during widespread implementation into health systems, concerns about accessibility (i.e., who has access to the technology) and accountability (i.e., ensuring the technology is performing according to communicated expectations) become prominent.

Autonomy and Privacy

The principle of protecting human autonomy in the health care context indicates that humans should retain full control of health care decisions and raises considerations for both clinicians and patients.[57,58] The adoption of AI into health care settings could lead to situations where clinical decision-making is shifted from clinicians to AI-enabled tools, thus potentially limiting the autonomy of clinicians in making care decisions and patients in being a part of these decisions. RapidAI can be considered as a supportive tool that assists clinicians by providing 1 piece of information for diagnosing patients undergoing assessment for acute stroke. By not acting as the sole source of clinical decision-making for stroke diagnosis and treatment, clinicians and patients can still exercise their autonomy in interpreting and making decisions about how to translate the results of this tool into clinical practice. This perspective aligns with RapidAI's indications for use.[21]

Patients have the right to be informed of the role, potential risks, benefits, and alternatives of medical procedures or interventions to be able to make informed decisions of whether to undergo treatment or testing.[59] This informed consent process is more complex for clinical decision support tools like RapidAI, as they do not directly interact with patients. During our discussion with the patient contributor, she asked whether patients are typically informed that AI will be or was involved in making their diagnosis, as it was unclear to her what standard practice is. She did not know whether AI had been used to inform a diagnosis in her experience. Further engagement with clinicians during the course of this review may have been able to further explore how the results of tools like RapidAI are communicated with patients and used in the context of shared clinical decision-making.

Implementers of RapidAI could consider whether and how to notify patients that AI is being used and when to engage in the informed consent process. For individual clinicians, this may involve engaging in shared decision-making with patients or substitute decision-makers before using RapidAI. For institutions, it could involve creating policies that specify how and when to disclose and discuss the use of AI with patients. While some individuals may view consulting an AI tool similarly to consulting a colleague or a medical textbook, which would not always be shared with the patient, findings from a recent survey-based study[75] indicate that people perceive AI tools differently and support the disclosure of when and how AI tools are used. According to Herington and colleagues,[59] patients ought to be informed of the use of an AI-enabled medical device during diagnostic or therapeutic interventions, the benefits and harms of using the technology, and any known limitations of the AI-enabled technology. Importantly,

these conversations should convey the benefits and risks of an AI-enabled medical device using language that is clinically relevant and understandable by the patient, rather than with abstract measures of performance.[59] Within the context of using RapidAI for stroke detection, this could mean informing patients that RapidAI may speed up the diagnostic process by reducing the time needed to review CT imaging results, but that the evidence exploring its effect on direct patient outcomes (e.g., functional status) is very uncertain. Additionally, these conversations could include a discussion of the potential risks of using RapidAI, such as the rates of false positive and false negative results and their potential consequences, or patient data privacy considerations. Sharing this type of information can promote patient autonomy by helping patients become fully informed about the implications of the technology.

Data privacy and security are important components of protecting human autonomy that must be safeguarded when implementing AI technologies.[57,58] In Canada, there are a number of data privacy laws and regulations at both the federal and provincial or territorial levels (e.g., the *Personal Health Information Protection Act*, the *Privacy Act*, the *Personal Information Protection and Electronic Documents Act*, and the *Freedom of Information and Protection of Privacy Act*) that establish rules for how personal information (including personal health information) is collected, used, and disclosed, which are described in the accompanying implementation review (found on the project website). Data governance also requires considering and respecting First Nations, Inuit, and Métis data sovereignty principles. These principles, such as the First Nations principles of OCAP®, Manitoba Métis principles of OCAS, and Inuit Qaujimajatuqangit, must guide the respectful governance of data collected with, from, or about Indigenous Peoples in Canada. The clinical literature included in this review did not provide details on how RapidAI manages, uses, and stores personal data, so the strategies RapidAI has in place to protect personal health data and to what degree they comply with existing privacy legislation or the interests of diverse peoples were not examined. Understanding how patient data are collected, used, and stored by AI-enabled technologies is critical in informing how these technologies impact human autonomy and dignity and whether patients decide to engage with these technologies.

The patient contributor suggested that in emergency situations where someone is experiencing stroke, some people would primarily focus on surviving the health event by being diagnosed and treated as soon as possible and may not consider ethical issues related to their data being analyzed by AI tools like RapidAI. She expressed minimal concerns about sharing her personal, de-identified information with the AI manufacturer for use of the software, for training the AI, quality improvement, or further development of new technologies. She likened this to sharing information with smartphone or computer software manufacturers or sharing your health card information with medical clinics. While she generally had few qualms about sharing her data, she mentioned that the privacy and reliability of the manufacturer's storage of personal health information were important. It is worth considering that patients may have limited concerns about the privacy of their data until a breach or other compromise occurs, potentially leading to detrimental consequences.

The patient contributor suggested that there might be a divide regarding privacy concerns, with some individuals potentially being more protective of their privacy and wanting more testing and validating of newer technologies before embracing them, while others might be more accustomed to sharing their data and interacting with newer technologies. Respecting patient autonomy requires that patients are fully informed of whether and how their data may be used so that they can make these decisions.

Mitigating Harms

AI technologies should minimize the risk of harm to people, including by meeting regulatory requirements for safety, accuracy, and efficacy before being used to inform patient care decisions,[57,58] which are described in the accompanying implementation review (found on the project website). In the case of RapidAI, this could refer to data that establishes its clinical performance (e.g., how it may benefit the diagnostic process) and provides information on its potential harms. After deployment, measures should be in place to ensure quality control and quality improvement, as well as to monitor the performance of the AI in real-world settings. Low and very low certainty evidence summarized in this review provided some information on clinical efficacy, but limited information on the downstream impacts of using RapidAI for stroke detection on patient harms.

In the absence of direct measures of harm (e.g., rates of symptomatic intracranial hemorrhage or procedure-related complications), the rates of false positives and false negatives can serve as additional sources of potential concern, as they can lead to undue worry or inappropriate treatment decisions resulting in unnecessary interventions or missed opportunities for timely care (and thus forgone benefits). Many of the included studies reported on the rates of false positives and false negatives when RapidAI was used as a standalone intervention; however, it is challenging to extrapolate this data to risk for direct harms to patients. Inaccurate diagnoses

by RapidAI only become problematic if (a) the diagnoses are not corrected or reinterpreted by clinicians before making care decisions or (b) clinicians overturn a correct diagnosis after reviewing the results of RapidAI analyses.

Transparency and Explainability

AI transparency, explainability, and intelligibility require that information is published or documented throughout the lifecycle of a technology, ensuring that decisions are clear and justifiable, and that operations are understandable to all relevant parties.[59,76] These concepts highlight the importance of ensuring an AI system's decision-making process is comprehensible to technology developers, patients, clinicians, and other users. Explainability refers to the ability to understand why an AI system arrives at a particular decision, while intelligibility involves how well the reasoning process can be understood by humans.[76,77] Achieving both explainability and intelligibility can be facilitated through model transparency.[78] Although these aspects are not often reported in studies evaluating the effectiveness or accuracy of AI-enabled digital health interventions, they are nonetheless important features for addressing the ethics of AI across the technology lifecycle.

Within the effectiveness and accuracy studies included in this review, there was no information detailing the methods used to develop the machine learning models used by RapidAI. Specifically, the methods for data abstraction, cleaning, and preparation, selecting model architectures, ground truth labelling, data splitting (i.e. into training and testing cohorts), sample size calculations, model training, and hyperparameter tuning during the model development phases were not described. Transparent reporting of these aspects may help to foster trust between RapidAI and its users, such as patients and clinicians.

Equity and Access

The principle of justice asserts that everyone should have fair and equal access to the benefits of health care, without discrimination against any individual or social group.[79] This has equity-related implications for AI-enabled clinical decision support tools, by implying the tool should consistently improve clinical decision-making regardless of patients' personal characteristics, such as place of residence, race, ethnicity, culture, language, occupation, gender, sex, religion, education, socioeconomic status, or social capital. To determine if the effectiveness and accuracy of RapidAI are consistent and robust across diverse populations, we used PROGRESSS-Plus criteria to guide data extraction and our reporting of findings. However, the included studies did not provide details on the characteristics of study populations and did not conduct subgroup analyses based these criteria, preventing us from evaluating how RapidAI might perform across different groups.

As previously noted, the literature we reviewed did not describe the methods used to develop RapidAI's machine learning models. As such, we were unable to assess the representativeness and diversity of the training dataset and comment on considerations related to inclusivity. Furthermore, it was unclear if RapidAI has undergone bias assessment to determine if certain patient subgroups (e.g., based on age, gender, ethnicity) are disproportionately affected by the model outputs. As a result, we are unable to comment on the potential bias risks in the stroke detection algorithms.

The patient contributor raised access concerns about the availability of AI technologies in hospitals outside urban stroke centers. She questioned whether AI-enabled stroke detection software would be available to all major hospitals for assisting in triaging and potentially transferring patients more quickly, or if its use would be restricted to certain facilities. Limited availability in community hospitals could raise equity issues, particularly in rural or remote regions with large populations of older adults who are at higher risk for stroke due to their age and health status. Even if the risk of bias in an algorithm's performance is low, limited access based on geographic location could exacerbate existing health inequities.

Conversely, the implementation of RapidAI has the potential to improve access to stroke care by increasing diagnostic efficiency. In health care settings with limited radiology expertise, RapidAI could help prioritize cases of highest urgency, potentially leading to quicker diagnosis and intervention for those who need it most. However, based on the evidence summarized in this review, it is very uncertain whether this translates into improved patient outcomes. Reducing the time needed to make a stroke diagnosis may not necessarily lead to increased access to care. Other aspects of the health system infrastructure, including the availability of emergency medical services, stroke care specialists, operating rooms, imaging equipment, radiology technologists, and other emergency care resources may still limit the speed of treatment. Consequently, while RapidAI has the potential to improve diagnostics efficiency, it is unlikely to address inequities in access to stroke care that arise from these broader systemic constraints.

In summary, the use of AI for detecting stroke raises ethical implications related to protecting human autonomy, data privacy, mitigating harms, transparency, equity, and access. These considerations are relevant and important to how AI-enabled digital health technologies are assessed. However, they tend to be underreported in evidence that is generally examined when evaluating the effectiveness and accuracy of interventions. A holistic view of these various dimensions is needed across the lifecycle of AI-enabled digital health technologies (i.e., through initial development, clinical testing, implementation, and ongoing monitoring) to better inform decision-making.

Strengths and Limitations

Our review employed robust methods that were guided by the current methodologic standards for Cochrane rapid reviews.[32] We integrated ethics and equity considerations throughout the review process, which could help to guide policy-making and clinical practice by highlighting some of the issues that may arise during the implementation of RapidAI for detecting stroke. Additionally, we incorporated the perspectives, experiences, and priorities of a patient contributor with lived experience of a stroke, and we sought feedback from clinical and ethics experts, the manufacturer, and other interested parties to ensure that multiple perspectives were considered. Despite these strengths, the review also has several limitations, which are described below.

Evidence Gaps

In addition to the limitations in the evidence noted above, no evidence was found for the following; therefore, no conclusions can be formed on these aspects:

- The impact of RapidAI for detecting ICH on clinical outcomes, such as time to intervention or direct patient outcomes (e.g., functional status).
- The diagnostic accuracy of RapidAI as an adjunct or aid to clinician interpretation (i.e., as it would be used in clinical practice) for detecting LVO.
- The cost-effectiveness of RapidAI for stroke detection.

Furthermore, none of the included studies reported mortality, length of hospital stay, health-related quality of life, or health care resource implications as outcomes; therefore, no conclusions can be formed on the impact of RapidAI on these outcomes. The clinical expert that we consulted for this project considered data on patient harms to be critically important for informing the use of RapidAI in practice, but these were also not reported. In the absence of clinical harms data, decision-makers may wish to reflect on the rates of false positives and false negatives reported in the diagnostic accuracy studies, as these have the potential to lead to harms for patients and health systems.

The evidence summarized in this review was predominantly of very low certainty, with the exception of 1 outcome-comparison, which was of low certainty. This indicates that the reported estimated effects may be very different from the true effect. Therefore, the findings should be interpreted with consideration for the limitations noted herein.

Generalizability

This rapid review summarizes the results of 2 cohort studies[64,65] and 11 diagnostic accuracy studies[64-74] evaluating the effectiveness and accuracy of RapidAI for detecting stroke. However, only 3 of these studies[64,65,70] directly evaluated RapidAI as it is used in clinical practice (i.e., as an adjunct or aid to clinician interpretation of CT imaging results). The remaining 10 included studies[64-69,71-74] evaluated RapidAI as a standalone intervention. Consequently, much of the evidence summarized in this review is indirect and does not provide a clear indication of how the results may apply to clinical practice.

None of the included studies were conducted in Canada. Additionally, both cohort studies[64,65] and 7 diagnostic accuracy studies[64,67,68,70-72,74] recruited participants from a single institution. While there was no strong indication that the findings from studies conducted in Australia,[66,67,74] the UK,[68,71] the US,[64,65,70,72] or studies that included patient data from a variety of sources[69,73] would

not apply to Canadian settings, differences in stroke diagnosis approaches or patient populations may limit the generalizability of the evidence to the Canadian context.

We used PROGRESS-Plus[34,35] to guide data extraction and report writing to attempt to gain insights into how the effectiveness and accuracy of RapidAI may vary across different populations. However, the included studies did not report on many criteria that were identified by the clinical and ethics experts that we consulted for this project as important in the context of stroke care or AI, such as place of residence, race, ethnicity, culture, language, education, socioeconomic status, and disability status. Some of the included studies[64-68,70-74] provided limited information on the demographics of study participants, including age and sex or gender. Race of study participants was reported in 2 studies (described in 1 publication).[64] However, no included study provided information on how sex, gender, and race were defined or measured. Due to the limited reporting of these characteristics, it is unclear if study populations included people from equity-deserving groups and whether the effects of RapidAI are generalizable, consistent, and robust across diverse patient populations.

We did not identify any evidence on the effectiveness or accuracy of RapidAI for stroke detection in pediatric populations or younger adults (the participants of included studies had mean or median ages between 61 and 75 years). As such, the appropriateness of RapidAI in evaluating younger patients with suspected stroke is unclear.

Heterogeneity of the Evidence

There was considerable clinical heterogeneity among the included studies with respect to the types of LVO examined, the reference standards used, and the versions of RapidAI being evaluated. Most studies described LVO as occlusions involving the M1 segment of the middle cerebral artery or the internal carotid artery. However, some studies also included occlusions of the M2 segment of the middle cerebral artery, the basilar artery, or the intracranial vertebral artery. For reference standards, some studies used interpretation by a single neuroradiologist while others used consensus interpretation by a panel of radiologists (with up to 4 members). There was also variability in the training level of reference standard interpreters (e.g., years of experience), their blinding status (i.e., whether they were aware of RapidAI results or the clinical history of patients), and the types of imaging used to make a reference standard diagnosis (e.g., CT angiography alone versus multimodal imaging). The included studies assessed various iterations and components of the RapidAI software, including Rapid CTA, Rapid LVO, the Rapid NCCT Stroke platform, or described the intervention as RapidAI without any further details. Altogether, these sources of heterogeneity limited our ability to synthesize results across studies, as for the individual comparisons often only 1 study was available.

Limitations of Our Approach

For feasibility reasons, only studies published in English were eligible for inclusion in this review. As a result, we may have introduced language bias and missed key data from studies evaluating the effectiveness, accuracy, or cost-effectiveness of RapidAI published in non-English languages. Within our literature search, we identified 2 studies[80,81] published in Spanish that may have contained relevant data. However, only the titles and abstracts of these publications[80,81] were available in English, and they were not considered further.

We acknowledge that our literature review was specific to comparative studies examining the effectiveness, accuracy, and cost-effectiveness of RapidAI for stroke detection. We did not search for other sources of information that may provide more details on the process for developing, training, and validating the machine learning models used by RapidAI, such as pre-clinical studies or product information sheets. While we did attempt to consider elements of ethics and equity in our primary literature review, more directed searching on ethical considerations related to AI in the context of medical imaging, radiology, or stroke detection may have augmented our consideration of ethical dimensions of the technology.

Finally, although we intended to engage with 2 patients and a clinician, we had limited response to our outreach for interested individuals. In the end, our engagement was limited to 1 patient contributor. While our conversation was helpful for understanding her perspectives, experiences, and priorities for the use of AI in stroke detection and contextualize the findings of our report, her input is unlikely generalizable to all patients who have experienced stroke. We are also unable to provide clinician perspectives without their participation. Our approach also required individuals to have access to reliable technology, phone, and internet access to view our recruitment initiatives and participate as contributors, which would have excluded some voices.

Conclusions and Implications for Decision- or Policy-Making

This review included 2 cohort studies and 11 diagnostic accuracy studies regarding the effectiveness and accuracy of RapidAI for stroke detection. No relevant evidence was identified regarding the cost-effectiveness of RapidAI for stroke detection.

Summary of Evidence

Findings from the included cohort studies suggest that evaluation of CT angiography imaging for detecting LVO by RapidAI with clinician interpretation, compared to clinician interpretation alone, for patients with suspected stroke may result in clinically important reductions in radiology report turnaround time. The evidence is very uncertain about the effect of evaluating CT angiography images with RapidAI on other time to intervention outcomes (e.g., time from door to intubation, time from door to revascularization), measures of functional status (i.e., mRS scores, NIHSS scores), and response to therapy (i.e., TICI scores). The certainty of evidence for all effectiveness outcomes was low or very low, primarily because of critical risk of bias due to confounding and imprecision, as results for each efficacy outcome were based on single studies, often with small sample sizes.

Low certainty evidence from 1 diagnostic accuracy study suggests that Rapid ICH with clinician interpretation, compared to a reference standard of concordant results between Rapid ICH and neuroradiologist interpretation or consensus diagnosis by a panel of 3 neuroradiologists, has a sensitivity of 92% (95% CI, 78 to 98%) and specificity of 100% (95% CI, 98 to 100%) for detecting ICH using non-contrast CT in people with suspected acute stroke. The certainty of evidence for this finding was rated down due to risk of bias and imprecision, as the lower bound of the 95% CI for sensitivity suggests a different conclusion regarding the diagnostic value of RapidAI.

As a standalone intervention, evidence from 10 diagnostic accuracy studies indicates that the sensitivity of RapidAI for detecting LVO ranges from 62% to 96%, while estimates of specificity ranging from 65% to 98%. There was heterogeneity in the types of LVO assessed, the versions of RapidAI used, the type of CT image analyzed (i.e., non-contrast CT or CT angiography), and the methods for determining reference standard diagnoses, likely contributing to the variability in results for sensitivity and specificity. The certainty of the evidence for these findings was moderate, low, very low, or there was insufficient information to judge certainty, primary due to risk of bias and imprecision. These results have unclear applicability to clinical practice, as the accuracy of RapidAI by itself does not directly answer how much it might improve the accuracy of a clinician reader, improve access to care, or impact patient-important outcomes.

Our findings are generally aligned with a 2024 NICE assessment[22] of AI-derived software to help clinical decision making in stroke, which included but was not specific to RapidAI for stroke detection. The NICE assessment[22] concluded that the clinical evidence on AI-derived software to help clinical decision making in stroke is limited in quality. However, some studies evaluating RapidAI (and 2 other AI-derived tools) suggested that people had faster access to treatment after using the software, although it was unclear to what extent this was an effect of the software. The diagnostics advisory committee recommended that RapidAI can be used in the National Health Service while more evidence is generated and that it should only be used alongside health care professional interpretation of CT scans to reduce the risk of incorrect results.[22]

Economic Information

Due to the lack of available economic evidence that met our criteria for this review, we cannot draw any conclusions on the cost-effectiveness of RapidAI for stroke detection. However, Ontario Health published a health technology assessment[31] in 2020 on the use of automated CT perfusion imaging to aid in the selection of patients with acute ischemic stroke for mechanical thrombectomy. They estimated that mechanical thrombectomy informed by automated CT perfusion imaging was likely to be cost-effective for eligible patients up to 24 hours after a stroke. Publicly funding automated CT perfusion imaging across 42 hospitals in Ontario was estimated to result in additional costs of $1.3 million in the first year and $0.9 million per year thereafter.[31]

According to their budget impact analysis, the annual cost of a licence for the RapidAI neuroimaging platform in 2019 was between $27,500 and $32,500 per hospital, depending on the number of connected scanners.[31] The increased cost in the first year was due to 1-time fees related to the initial implementation and optimization of RapidAI, as well as for training staff ($12,350 for hospitals with

2 or more scanners).[31] This costing information was for the full RapidAI platform, which includes several modules for conducting various cerebrovascular diagnostic procedures, in addition to automated CT perfusion. Their analysis focused on 1 component of the platform, which does capture the full value of the technology.

We reached out to the manufacturer but were unable to obtain updated pricing information. Although the current report does not address RapidAI's capacity to inform treatment selection (e.g., selecting patients for reperfusion therapy), the findings of the health technology assessment[31] by Ontario Health may provide insights into the potential costs and benefits of implementing the RapidAI platform.

Considerations for Future Research

Robustly designed and transparently reported studies of low risk of bias that enroll diverse patient populations are needed to increase the certainty of the evidence. Investigators should aim to consistently and transparently report the versions and algorithms of RapidAI being tested, the types of LVO included, and the methods for enrolling participants. The evidence from observational studies summarized in this review was limited by a critical risk of bias due to confounding. Future studies should apply methods to control for confounding factors (e.g., age, smoking status, diabetes, hypertension) through experimental design or through statistical analysis. When assessing diagnostic accuracy, future studies should evaluate the performance of RapidAI used to assist clinician interpretation, rather than used alone, as it is applied in clinical practice.

Because the studies included in our review were focused on applying RapidAI in clinical settings (rather than earlier stages of development, such as when the machine learning models were trained and refined), we found not all items of the APPRAISE-AI tool[43] and the CLAIM guideline[44] were addressed in the studies included in our assessment. However, we encourage researchers, manufacturers, and technology developers to consider transparently reporting these aspects in studies evaluating the effectiveness and accuracy of AI-enabled digital health technologies, as this is the type of information evaluators are interested in.

Stroke disproportionately affects women,[82] people at lower levels of socioeconomic status,[83] people living in rural or remote areas,[13] and people with various racial and ethnic identities.[84] To help address health equity concerns, researchers should consider collecting and reporting equity-relevant population characteristics, such as place of residence, race, ethnicity, culture, language, occupation, gender, sex, religion, education, socioeconomic status, social capital, and disability status. Although these types of information are often underreported in medical imaging literature due to the common practice of dataset anonymization, including these characteristics and performing subgroup analyses where appropriate would generate evidence on how the performance of RapidAI may vary across different populations. This could provide insight into the potential impact of RapidAI on existing health inequities and help ensure that it does not create new ones. The literature we reviewed did not indicate whether equity-deserving groups were included in the training data for RapidAI algorithms; future research should explicitly report on this, especially as the stroke detection algorithms are updated or new ones are developed.

Our conversation with the patient contributor engaged in this project identified several patient-important outcomes, including time to intervention or diagnosis, diagnostic accuracy, mortality, and measures of physical and cognitive function. While some of these outcomes were reported in the included studies, the evidence was generally very uncertain. Future studies should evaluate the effects of RapidAI on patient-important outcomes (including benefits and harms) as well as its potential impact on health care resource utilization to assess its potential for alleviating health human resource constraints. This would provide a more comprehensive understanding of the role of RapidAI in stroke detection.

As the number of available AI-enabled digital health technologies continues to increase, evidence evaluating their effectiveness and accuracy could evolve to provide detailed descriptions of the process used to develop machine learning models. This includes the methods for data collection, data preprocessing techniques, model selection criteria, and the strategies used to validate models and mitigate the risk for algorithmic bias. Such reporting would improve transparency, enabling thorough assessments by regulators and evaluators. It would also support more informed discussions about the ethical implications of AI technologies and could help to ensure technologies meet the needs of diverse patient populations.

Implications for Clinical Practice and Policy-Making

The findings of this report suggest RapidAI shows promise for assisting clinicians in detecting LVO or ICH on CT images for patients with suspected acute stroke. The use of RapidAI for stroke detection may result in clinically important reductions in radiology report turnaround time. The evidence is very uncertain about the effect of RapidAI on other clinical outcomes, including several time to intervention metrics, measures of physical and cognitive function, and response to therapy (e.g., reperfusion rates). Evidence of the accuracy of RapidAI when used as intended (i.e., alongside clinician judgment) was limited to 1 study.[70] Decision-makers who intend to use these findings to inform decisions should consider that the current evidence is limited and of low or very low certainty. The uncertain potential for clinical benefit secondary to reduced radiology report turnaround times needs to be balanced with the risk for undesirable effects resulting from false positives and false negatives. Although false positives and false negatives were infrequent in the 1 study,[70] the certainty of evidence was low.

The implementation of AI-enabled digital health technologies, including RapidAI, raises ethical concerns that need to be addressed to ensure responsible and equitable implementation. In this review, we discussed how factors such as autonomy, privacy, safety, transparency, explainability, and equity relate to RapidAI and could influence its acceptability by clinicians, patients, and health care institutions as they decide whether to adopt it. Decision-makers should reflect on the implications of these equity and ethics considerations within their local context when making decisions about the use of RapidAI for stroke detection.

Based on mean radiology report turnaround times and the point estimate for the between-group difference from 1 study,[64] RapidAI may result in quicker radiology report turnaround times. While this estimate is from low certainty evidence and our confidence in the effect estimate is limited (the true effect may be substantially different from the estimate of effect), the use of RapidAI could lead to an increased number of assessments that can be undertaken or less radiologist time required to undertake the same number of assessments. For example, based on assumptions of a standard 8-hour shift for a radiologist and that their time is solely allocated to radiology assessments and the mean radiology report turnaround times reported (22 minutes per assessment with RapidAI and 30.5 minutes per assessment without RapidAI), the use of RapidAI may lead to approximately 22 assessments over an 8-hour period, compared with 16 assessments without the use of RapidAI. Such gains in efficiency could lead to increased clinician capacity for other tasks and faster triage, decision-making, and treatment initiation and ultimately in better patient outcomes. Future research addressing the limitations of the current evidence could provide greater clarity on how RapidAI impacts patients, clinicians, and decision-makers.

References

1. Kothari R, Sauerbeck L, Jauch E, et al. Patients' Awareness of Stroke Signs, Symptoms, and Risk Factors. *Stroke.* 1997;28(10):1871-1875.
2. Public Health Agency of Canada. Stroke in Canada. Ottawa (ON): Government of Canada; 2022: https://www.canada.ca/en/public-health/services/publications/diseases-conditions/stroke-in-canada.html. Accessed 2024 May 3.
3. FACT SHEET: Stroke Statistics. Ontario Stroke Network; 2012: https://www.ontariostrokenetwork.ca/pdf/Final_Fact_Sheet_Stroke_Stats_3.pdf. Accessed 2024 May 15.
4. Holodinsky JK, Lindsay P, Yu AYX, Ganesh A, Joundi RA, Hill MD. Estimating the Number of Hospital or Emergency Department Presentations for Stroke in Canada. *Can J Neurol Sci.* 2023;50(6):820-825.
5. Muller CJ, Alonso A, Forster J, et al. Stroke Incidence and Survival in American Indians, Blacks, and Whites: The Strong Heart Study and Atherosclerosis Risk in Communities Study. *J Am Heart Assoc.* 2019;8(12):e010229.
6. Hyldgård VB, Johnsen SP, Støvring H, Søgaard R. Socioeconomic Status And Acute Stroke Care: Has The Inequality Gap Been Closed? *Clin Epidemiol.* 2019;11:933-941.
7. Thompson SG, Barber PA, Gommans JH, et al. Geographic Disparities in Stroke Outcomes and Service Access: A Prospective Observational Study. *Neurology.* 2022;99(4):e414-e426.
8. Ikeme S, Kottenmeier E, Uzochukwu G, Brinjikji W. Evidence-Based Disparities in Stroke Care Metrics and Outcomes in the United States: A Systematic Review. *Stroke.* 2022;53(3):670-679.
9. Cruise C, Mfoafo M'Carthy N, Ganesh A, Lashewicz B. Imperfect Patients: Disparities in Treatment of Stroke Patients with Premorbid Disability. *CJNS.* 2023;50(6):826-837.
10. Stroke and mental health: The invisible and inequitable effects on women. *2023 Stroke report*. Toronto (ON): Heart and Stroke Foundation of Canada; 2023: https://issuu.com/heartandstroke/docs/womens_stroke_report_june_2023/5. Accessed 2024 Jun 4.
11. Lives disrupted: the impact of stroke on women. *2018 Stroke report*. Toronto (ON): Heart and Stroke Foundation of Canada; 2018: https://www.heartandstroke.ca/-/media/pdf-files/canada/stroke-report/strokereport2018.pdf?rev=8491d9c349f7404491f36be67f649c0b. Accessed 2024 Jun 4.
12. Towfighi A, Boden-Albala B, Cruz-Flores S, et al. Strategies to Reduce Racial and Ethnic Inequities in Stroke Preparedness, Care, Recovery, and Risk Factor Control: A Scientific Statement From the American Heart Association. *Stroke.* 2023;54(7):e371-e388.
13. Fleet R, Bussières S, Tounkara FK, et al. Rural versus urban academic hospital mortality following stroke in Canada. *PLoS One.* 2018;13(1):e0191151.
14. Tadi P, Lui F. Acute Stroke. *StatPearls*. Treasure Island (FL): StatPearls Publishing LLC.; 2023: https://www.ncbi.nlm.nih.gov/books/NBK535369/. Accessed 2024 May 15.
15. Heran M, Lindsay P, Gubitz G, et al. Canadian Stroke Best Practice Recommendations: Acute Stroke Management, 7th Edition Practice Guidelines Update, 2022. *Can J Neurol Sci.* 2024;51(1):1-31.
16. Saver JL. Time is brain--quantified. *Stroke.* 2006;37(1):263-266.
17. Huang Y, Zheng H, Mo M. Effect of different operation time on surgical effect and quality of life in patients with severe hypertensive intracerebral hemorrhage. *Am J Transl Res.* 2021;13(8):9538-9545.
18. Powers WJ, Rabinstein AA, Ackerson T, et al. 2018 Guidelines for the Early Management of Patients With Acute Ischemic Stroke: A Guideline for Healthcare Professionals From the American Heart Association/American Stroke Association. *Stroke.* 2018;49(3):e46-e110.
19. Musuka TD, Wilton SB, Traboulsi M, Hill MD. Diagnosis and management of acute ischemic stroke: speed is critical. *CMAJ.* 2015;187(12):887-893.
20. RapidAI. https://www.rapidai.com/. Accessed 2024 Jun 13.
21. *RAPID User's Manual.* Meno Park (CA): iSchemaView, Inc.
22. Artificial intelligence (AI)-derived software to help clinical decision making in stroke. London (UK): NICE; 2024: https://www.nice.org.uk/guidance/dg57.
23. Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices. In: Post TW, ed. *UpToDate*. Waltham (MD): UpToDate; 2024: https://www.uptodate.com/. Accessed 2025 Jul 17.
24. 2024–2025 Annual Business Plan. Ottawa (ON): CADTH; 2024: https://www.cadth.ca/sites/default/files/corporate/planning_documents/2024-2025_Annual_Business_Plan.pdf. Accessed 2024 Jul 23.
25. News: International Health Technology Assessment Collaboration Expands. 2023; https://www.cadth.ca/news/international-health-technology-assessment-collaboration-expands. Accessed 2024 Jul 24.
26. Scottish Health Technologies Group (SHTG). An evidence framework for the assessment of health technologies. Edinburgh (GB): Healthcare Improvement Scotland; 2023: https://shtg.scot/what-we-do/evidence-framework/. Accessed 2024 Jul 24.

27. Transformation Directorate. Digital Technology Assessment Criteria (DTAC). London (GB): NHS England; 2021: https://transform.england.nhs.uk/key-tools-and-info/digital-technology-assessment-criteria-dtac/. Accessed 2024 Jul 24.
28. Westwood M, Ramaekers B, Grimm S, et al. Software with artificial intelligence-derived algorithms for analysing CT brain scans in people with a suspected acute stroke: a systematic review and cost-effectiveness analysis. *Health Technol Assess.* 2024;28(11):1-204.
29. Akay EMZ, Hilbert A, Carlisle BG, Madai VI, Mutke MA, Frey D. Artificial Intelligence for Clinical Decision Support in Acute Ischemic Stroke: A Systematic Review. *Stroke.* 2023;54(6):1505-1516.
30. Murray NM, Unberath M, Hager GD, Hui FK. Artificial intelligence to diagnose ischemic stroke and identify large vessel occlusions: a systematic review. *J Neurointerv Surg.* 2020;12(2):156-164.
31. Automated CT perfusion imaging to aid in the selection of patients with acute ischemic stroke for mechanical thrombectomy: A health technology assessment. *Ont Health Technol Assess Ser.* 2020;20(13):1-87.
32. Garritty C, Hamel C, Trivella M, et al. Updated recommendations for the Cochrane rapid review methods guidance for rapid reviews of effectiveness. *BMJ.* 2024;384:e076335.
33. Vivian A Welch JP, Janet Jull, Lisa Hartling, Terry Klassen, Elizabeth Kristjansson, Jordi Pardo Pardo, Mark Petticrew, David J Stott, Denise Thomson, Erin Ueffing, Katrina Williams, Camilla Young, Peter Tugwell. Chapter 16: Equity and specific populations. In: Higgins JPT TJ, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, ed. *Cochrane Handbook for Systematic Reviews of Interventions*: Cochrane; 2023.
34. O'Neill J, Tabish H, Welch V, et al. Applying an equity lens to interventions: using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. *J Clin Epidemiol.* 2014;67(1):56-64.
35. Oliver S, Kavanagh J, Caird J, et al. Health Promotion, Inequalities and Young People's health: a Systematic Review of Research. *EPPI-Centre report no. 161*. London (GB): University of London; 2008: https://eppi.ioe.ac.uk/cms/Portals/0/PDF%20reviews%20and%20summaries/Inequalities%20Young%20People%20R2008Oliver.pdf?ver=2010-12-22-123934-167. Accessed 2024 Jun 4.
36. Pexman JH, Barber PA, Hill MD, et al. Use of the Alberta Stroke Program Early CT Score (ASPECTS) for assessing CT scans in patients with acute stroke. *AJNR Am J Neuroradiol.* 2001;22(8):1534-1542.
37. PM B. Chapter 3: Understanding the design of test accuracy studies. In: Deeks JJ BP, Leeflang MM, Takwoingi Y, ed. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy.* Vol Version 2.0: Cochrane; 2023.
38. Mathes T, Pieper D. An algorithm for the classification of study designs to assess diagnostic, prognostic and predictive test accuracy in systematic reviews. *Syst Rev.* 2019;8(1):226.
39. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ.* 2016;355:i4919.
40. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011;155(8):529-536.
41. Reitsma JB RA, Whiting P, Yang B, Leeflang MM, Bossuyt PM, Deeks JJ. Chapter 8: Assessing risk of bias and applicability. In: Deeks JJ BP, Leeflang MM, Takwoingi Y (editors), ed. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*: Cochrane; 2023.
42. Higgins JPT, Morgan RL, Rooney AA, et al. A tool to assess risk of bias in non-randomized follow-up studies of exposure effects (ROBINS-E). *Environ Int.* 2024;186:108602.
43. Kwong JCC, Khondker A, Lajkosz K, et al. APPRAISE-AI Tool for Quantitative Evaluation of AI Studies for Clinical Decision Support. *JAMA Network Open.* 2023;6(9):e2335377-e2335377.
44. Mongan J, Moy L, Charles E. Kahn J. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A Guide for Authors and Reviewers. *Radiology: Artificial Intelligence.* 2020;2(2):e200029.
45. Popay J, Roberts H, Aj S, et al. Guidance on the Conduct of Narrative Synthesis in Systematic Reviews Final Report. *J Epidemiol Community Health.* 2005;59(Suppl 1)::A7.
46. Julian PT Higgins TL, Jonathan J Deeks. Chapter 6: Choosing effect measures and computing estimates of effect. In: Higgins JPT TJ, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, ed. *Cochrane Handbook for Systematic Reviews of Interventions*: Cochrane; 2023.
47. Stevenson M, Sergeant E. epiR: Tools for the Analysis of Epidemiological Data. R package version 2.0.74. 2024; https://CRAN.R-project.org/package=epiR. Accessed 2024 Jun 5.
48. Scherer R. PropCIS: Various Confidence Interval Methods for Proportions. 2018; https://cran.r-project.org/web/packages/PropCIs/index.html. Accessed 2024 Jun 20.
49. Zeng L, Brignardello-Petersen R, Hultcrantz M, et al. GRADE guidelines 32: GRADE offers guidance on choosing targets of GRADE certainty of evidence ratings. *J Clin Epidemiol.* 2021;137:163-175.
50. Hultcrantz M, Rind D, Akl EA, et al. The GRADE Working Group clarifies the construct of certainty of evidence. *J Clin Epidemiol.* 2017;87:4-13.
51. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol.* 2011;64(4):383-394.

52. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. *BMJ.* 2004;328(7454):1490.
53. Schünemann HJ, Mustafa RA, Brozek J, et al. GRADE guidelines: 21 part 1. Study design, risk of bias, and indirectness in rating the certainty across a body of evidence for test accuracy. *J Clin Epidemiol.* 2020;122:129-141.
54. Schünemann HJ, Mustafa RA, Brozek J, et al. GRADE guidelines: 21 part 1. Study design, risk of bias, and indirectness in rating the certainty across a body of evidence for test accuracy. *J Clin Epidemiol.* 2020;122:129-141.
55. Murad MH, Mustafa RA, Schünemann HJ, Sultan S, Santesso N. Rating the certainty in evidence in the absence of a single estimate of effect. *Evidence Based Medicine.* 2017;22(3):85-87.
56. Santesso N, Glenton C, Dahm P, et al. GRADE guidelines 26: informative statements to communicate the findings of systematic reviews of interventions. *J Clin Epidemiol.* 2020;119:126-135.
57. United Nations Educational SaCO. Recommendation on the Ethics of Artificial Intelligence. Paris (FR): UNESCO; 2022: https://unesdoc.unesco.org/ark:/48223/pf0000381137. Accessed 2024 Jun 20.
58. Guidance W. Ethics and governance of artificial intelligence for health. *World Health Organization.* 2021.
59. Herington J, McCradden MD, Creel K, et al. Ethical Considerations for Artificial Intelligence in Medical Imaging: Deployment and Governance. *J Nucl Med.* 2023;64(10):1509-1515.
60. EUnetHTA Joint Action 2, Work Package 8. HTA Core Model ® version 3.0 (Pdf). Diemen (NLD): EUnetHTA; 2016: https://www.eunethta.eu/wp-content/uploads/2018/01/HTACoreModel3.0.pdf. Accessed 2024 Jun 20.
61. Benkhalti M, Espinoza M, Cookson R, Welch V, Tugwell P, Dagenais P. Development of a checklist to guide equity considerations in health technology assessment. *Int J Technol Assess Health Care.* 2021;37:e17.
62. Staniszewska S, Brett J, Simera I, et al. GRIPP2 reporting checklists: tools to improve reporting of patient and public involvement in research. *BMJ.* 2017;358:j3453.
63. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ.* 2021;372:n71.
64. Soun JE, Zolyan A, McLouth J, et al. Impact of an automated large vessel occlusion detection tool on clinical workflow and patient outcomes. *Front Neurol.* 2023;14:1179250.
65. Adhya J, Li C, Eisenmenger L, et al. Positive predictive value and stroke workflow outcomes using automated vessel density (RAPID-CTA) in stroke patients: One year experience. *Neuroradiol J.* 2021;34(5):476-481.
66. Delora A, Hadjialiakbari C, Percenti E, et al. Viz LVO versus Rapid LVO in detection of large vessel occlusion on CT angiography for acute stroke. *J Neurointerv Surg.* 2024;16(6):599-602.
67. Slater LA, Ravintharan N, Goergen S, et al. *RapidAI* Compared With Human Readers of Acute Stroke Imaging for Detection of Intracranial Vessel Occlusion. *Stroke: Vascular and Interventional Neurology.* 2024;4(2):e001145.
68. Chan N, Sibtain N, Booth T, et al. Machine-learning algorithm in acute stroke: real-world experience. *Clin Radiol.* 2023;78(2):e45-e51.
69. Yedavalli V, Heit JJ, Dehkharghani S, et al. Performance of RAPID noncontrast CT stroke platform in large vessel occlusion and intracranial hemorrhage detection. *Front Neurol.* 2023;14:1324088.
70. Eldaya RW, Kansagra AP, Zei M, et al. Performance of Automated RAPID Intracranial Hemorrhage Detection in Real-World Practice: A Single-Institution Experience. *Journal of Computer Assisted Tomography.* 2022;46(5):770-774.
71. Mallon DH, Taylor EJR, Vittay OI, Sheeka A, Doig D, Lobotesis K. Comparison of automated ASPECTS, large vessel occlusion detection and CTP analysis provided by Brainomix and RapidAI in patients with suspected ischaemic stroke. *J Stroke Cerebrovasc Dis.* 2022;31(10):106702.
72. Schlossman J, Ro D, Salehi S, et al. Head-to-head comparison of commercial artificial intelligence solutions for detection of large vessel occlusion at a comprehensive stroke center. *Front Neurol.* 2022;13:1026609.
73. Dehkharghani S, Lansberg M, Venkatsubramanian C, et al. High-Performance Automated Anterior Circulation CT Angiographic Clot Detection in Acute Stroke: A Multireader Comparison. *Radiology.* 2021;298(3):665-670.
74. Amukotuwa SA, Straka M, Smith H, et al. Automated Detection of Intracranial Large Vessel Occlusions on Computed Tomography Angiography: A Single Center Experience. *Stroke.* 2019;50(10):2790-2798.
75. Park HJ. Patient perspectives on informed consent for medical AI: A web-based experiment. *Digit Health.* 2024;10:20552076241247938.
76. Amann J, Blasimme A, Vayena E, Frey D, Madai VI. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Med Inform Decis Mak.* 2020;20(1):310.
77. Linardatos P, Papastefanopoulos V, Kotsiantis S. Explainable AI: A Review of Machine Learning Interpretability Methods. *Entropy (Basel).* 2020;23(1).
78. Ennab M, McHeick H. Designing an Interpretability-Based Model to Explain the Artificial Intelligence Algorithms in Healthcare. *Diagnostics (Basel).* 2022;12(7).
79. Varkey B. Principles of Clinical Ethics and Their Application to Practice. *Med Princ Pract.* 2021;30(1):17-28.
80. Cirio JJ, Diluca P, Ciardi C, et al. Impact of artificial intelligence on therapeutic metrics of cerebrovascular attack during the COVID-19 pandemic. *Medicina (Argentina).* 2023;83(5):705-718.

81. Cirio JJ, Ciardi C, Buezas M, et al. Implementation of artificial intelligence in hyperacute arterial reperfusion treatment in a comprehensive stroke center. *Neurologia Argentina.* 2021;13(4):212-220.
82. Alimi IO, Archibald A, Nwankwo O, Palanichami D, Oladimeji T, Keku E. Increasing the Awareness of Stroke Among Canadian Women. *Cureus.* 2022;14(3):e23159.
83. Huang K, Khan N, Kwan A, Fang J, Yun L, Kapral MK. Socioeconomic status and care after stroke: results from the Registry of the Canadian Stroke Network. *Stroke.* 2013;44(2):477-482.
84. Hosman FL, Engels S, den Ruijter HM, Exalto LG. Call to Action for Enhanced Equity: Racial/Ethnic Diversity and Sex Differences in Stroke Symptoms. *Front Cardiovasc Med.* 2022;9:874239.

ISSN: TBD

About CADTH: CADTH is an independent, not-for-profit organization responsible for providing Canada's health care decision-makers with objective evidence to help make informed decisions about the optimal use of drugs, medical devices, diagnostics, and procedures in our health care system.

Funding: CADTH receives funding from Canada's federal, provincial, and territorial governments, with the exception of Quebec.

Questions or requests for information about this report can be directed to **Requests@CADTH.ca**