Canadian Journal of Health Technologies

Canada's Drug Agency
L'Agence des médicaments du Canada

November 2024 Volume 4 Issue 11

Drugs Health Technologies Health Systems

Health Technology Review

RapidAI for Stroke Detection: Main Report

Table of Contents

RapidAI Review

List of Tables

Abbreviations

AI	artificial intelligence
CDA-AMC	Canada’s Drug Agency
CI	confidence interval
CLAIM	Checklist for Artificial Intelligence in Medical Imaging
DICOM	Digital Imaging and Communications in Medicine
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HTA	health technology assessment
ICH	intracranial hemorrhage
LVO	large-vessel occlusion
mRS	modified Rankin Scale
NIHSS	US National Institutes of Health Stroke Scale
PACS	picture archiving and communication systems
SD	standard deviation
TICI	Thrombolysis in Cerebral Infarction

Key Messages

What Is the Issue?

- Stroke is a sudden loss of neurologic function caused by poor or interrupted blood flow within the brain. It is 1 of the leading causes of death and a major cause of disability in Canada. For patients with suspected stroke, prompt evaluation using CT imaging and other tests can help to determine the type of stroke, to assess the severity of damage, and to guide treatment decisions.
- RapidAI is an artificial intelligence (AI)–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 (LVO) (i.e., ischemic stroke) and intracranial hemorrhage (ICH) (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. Our methods were guided by the Scottish Health Technologies Group's health technology assessment (HTA) framework.
- 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 improving speed and accuracy of diagnosis, minimizing the damaging effects of stroke, and reducing mortality rates. She also highlighted ethical considerations regarding the use of AI in health care, such as providing data privacy and equitable access, as well as 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 ICH, low-certainty evidence suggests that Rapid ICH combined with clinician interpretation, using clinician interpretation as a reference standard, has a sensitivity of 92% (95% confidence interval [CI], 78% to 98%) and a specificity of 100% (95% CI, 98% to 100%). However, estimates of sensitivity and specificity for detecting LVO 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, or 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 life cycle when using RapidAI for detecting stroke.

What Does This Mean?

- RapidAI has the potential to improve acute stroke care by creating efficiencies in the diagnostic process. However, the impact of RapidAI on many outcomes, including those that are important to patients, is uncertain due to limitations of the available evidence.
- To improve the certainty of findings, there is a need for evidence from robustly conducted studies at lower risk of bias that enrol 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 neurologic 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 a major cause of disability (e.g., limitations to physical functioning, sensory

impairment),[2-4] with more than 100,000 stroke events resulting in hospital or emergency department presentations each year.[5]

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.[6-13] 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.[11,12] 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.[14]

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.[15] Hemorrhagic strokes occur when a blood vessel ruptures, causing bleeding into the brain.[15]

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), neurologic examination, laboratory tests (e.g., blood work including electrolyte levels, random glucose levels, complete blood count, coagulation status, and creatinine levels), and cerebrovascular imaging with CT scans or MRI.[16] Decision-making in stroke assessment is often facilitated by data-based software programs, such as image postprocessing software packages.[17]

Stroke diagnosis and intervention is time sensitive. The 2022 Canadian Stroke Best Practice Recommendations[16] 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.[18-20] Following ischemic stroke, an estimated 1.9 million neurons die every minute it remains untreated.[18] Accurately determining whether a stroke is ischemic or hemorrhagic with neuroimaging studies, often CT scans, is crucial for selecting appropriate treatment options.[15] Some of the main interventions for ischemic stroke include IV thrombolysis and mechanical thrombectomy, whereas hemorrhagic stroke treatments aim to control bleeding, reduce pressure on the brain, and prevent further complications.[16,21] 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 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 noncontrast CT, CT angiography, and CT perfusion) that can be used by clinicians to assist in assessing patients with suspected stroke and deciding appropriate treatment.[22] RapidAI builds upon the original RAPID software, which was initially developed to automate and expedite the postprocessing of CT perfusion imaging and is widely

implemented in some jurisdictions, such as Ontario.[23,24] By reducing the time needed to generate perfusion maps and improving information sharing among the clinical team, RapidAI supports faster clinical decision-making and helps identify patients with ischemic stroke who may benefit from thrombectomy, even beyond the conventional 6-hour treatment window since symptom onset.[25-29] More recently, the RAPID platform has expanded its suite of products to include several static AI-derived algorithms for evaluating the brain's physiological status, such as Rapid ICH, Rapid ASPECTS, Rapid CTA, Rapid LVO, and Rapid HVS.[22,30] These algorithms are considered static because they remain fixed and unchanging after development through machine-learning processes, performing their tasks based on pre-established rules and training data. This differs from dynamic or adaptive algorithms, which continue to learn and update themselves based on new data over time. While RapidAI has numerous features and functionalities, what is relevant for this report is its capability to detect suspected ICH and LVO, which is used to inform stroke diagnosis. 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.[31] Potential benefits may include reducing crucial time intervals in stroke care, enhancing workflow efficiencies in emergency settings, improving stroke assessment in centres with limited access to stroke care specialists (e.g., in regional, rural, or remote communities without stroke centres and where nonstroke providers encounter stroke patients), and facilitating faster triage, decision-making, and treatment initiation, which could result in better patient outcomes and improved geographic equity in access to stroke care.

As of March 2024 (i.e., when we checked 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 [PACS]). It can be installed on a single server (virtual or physical) within the hospital firewall that can be remotely configured.[32] While RapidAI should be compatible with every scanner make and model,[32] hospital sites may need to consider the age of their current MRI and CT scanners. With certain older scanners, patients may require 2 scans, because data cannot be obtained on the entire brain with a single data acquisition.

In terms of training, the manufacturer offers a combination of online training webinars and in-person vendor-hosted sessions. These include clinical training for physicians (e.g., stroke neurologists, neurointerventional physicians, interventional radiologists, radiologists) on the automated software outputs, image interpretation, artifact recognition, and troubleshooting, as well as training for technologists on CT perfusion, automated software outputs, stroke CT perfusion protocol, IV placement, artifact recognition, and troubleshooting.[24]

Why Is It Important to Do This Review?

Globally, we are seeing an increase in medical devices relying on software incorporating AI.[33] Given the inherent nature of AI as a disruptive technology in health care, its comprehensive assessment through 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.

Canada's Drug Agency (CDA-AMC) has recently established information-sharing and collaborative relationships with various organizations, including an international partnership with HTA bodies.[34] From this partnership, CDA-AMC learned of the Scottish Health Technology Group's evidence framework, which outlines an approach to digital HTA. This includes an HTA framework[35] and Digital Technology Assessment Criteria as an add-on component.[36] This review of RapidAI applied the Scottish Health Technology Group's HTA framework.[35] The accompanying implementation review applied the Digital Technology Assessment Criteria.[36] The application of the Scottish framework allowed us to leverage existing work to ensure alignment and harmonization across organizations and to gain efficiency and sustainability. We plan 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 are 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[26,37-39] 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). Due to time and other resource constraints, 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.[40]

To address health equity in our review process, we followed guidance outlined in the *Cochrane Handbook for Systematic Reviews of Interventions*.[41]

- 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.
- We 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.

- We investigated subgroup effects, as informed by PROGRESS-Plus criteria,[42,43] that were identified by the experts consulted during this project as important in the context of stroke care or AI.
- We used PROGRESS-Plus[42,43] 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.[35] When possible, we provided information addressing the items for each of the domains set out in the framework.[35]

This rapid review was conducted alongside an implementation review for AI-enabled medical devices.

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 major international health technology agencies and health technology agencies in Canada, as well as a focused internet search. The search approach was customized to retrieve a limited set of results, balancing comprehensiveness with relevance. The search strategy comprised both controlled vocabulary, such as the National Library of Medicine's MeSH (Medical Subject Headings), and keywords. The 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 August 19, 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 the detection of both 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 provided outcome data only 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 provided outcome data only for patients who underwent tissue plasminogen activator therapy or mechanical thrombectomy following diagnosis, 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

mRS = modified Rankin Scale; QALY = quality-adjusted life-year; ROC = receiver operating characteristic.

[a]RapidAI as a standalone intervention was considered relevant only 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 preprints. 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,[44] 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[42,43] to guide data extraction. Each included publication was checked to determine whether 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).[42,43] 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*[45] and Mathes and Pieper.[46]

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[47] for nonrandomized studies and the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) checklist[48] for diagnostic accuracy studies. Published guidance for each tool was followed.[47-49] 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),[50] we did not proceed with assessing other domains of bias for studies deemed at critical risk of bias due to baseline confounding or the use of inappropriate methods of outcome measurement. 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)[51] and medical imaging (e.g., Checklist for Artificial Intelligence in Medical Imaging [CLAIM] guideline).[52] Studies were included in this review regardless of 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 judgments 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.[53] 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 formulas[54] to streamline the presentation of population characteristics across included studies. When they were not reported in publications, we calculated diagnostic accuracy metrics (e.g., sensitivity, specificity, negative predictive value, positive predictive value, concordance rates) with Clopper-Pearson exact 95% CIs from the available data (e.g., true-positives, false-positives, true-negatives, false-negatives) using the EpiR package in R.[55] Concordance rates are presented to estimate the agreement between index tests and reference standards. However, they do not differentiate disagreements due to false-positives and false-negatives, which can have different clinical implications. Because of this, concordance rates alone are generally not used to inform decision-making and should be considered alongside other performance metrics, such as sensitivity and specificity.[56] 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.[57]

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.[58-63] 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 was 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 nonnull 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,[64] 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).[65]

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 in the Scottish Health Technology Group's HTA framework.[35] These included items related to the use and impact of the technology in health care systems; the safety, acceptability, and credibility of the technology; and expectations for 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,[66] *WHO Guidance: Ethics and Governance of Artificial Intelligence for Health*),[67] the ethical considerations proposed by the AI Task Force of the Society of Nuclear Medicine and Molecular Imaging,[68] as well as the EUnetHTA Core Model 3.0 Ethics Domain[69] and the Equity Checklist for HTA.[70]

Prompts and guiding principles from these sources were used to identify and reflect on ethics and equity considerations of RapidAI throughout the technology life cycle 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 was 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: the patient contributor's lived experience with a stroke and 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,[71] which is outlined the Supporting Information document (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 reviewed the draft version of this report, and their feedback was incorporated into the final version of this report.

Manufacturer Review

The manufacturer reviewed a preliminary list of included studies and the draft version of this report, and its feedback was incorporated into this final version of the report.

Feedback Opportunity

The draft version of this report was posted on the CDA-AMC website to allow interested parties to provide feedback. We considered all feedback received and revised the report accordingly. Therefore, this final version of the report reflects the comments and suggestions from the peer reviewers, the manufacturer, and the feedback opportunity.

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 consisted of 2 cohort studies and 11 diagnostic accuracy studies described in 11 publications (2 publications assessed clinical outcomes before and after implementing RapidAI using a nonconcurrent cohort study design and diagnostic accuracy outcomes using a cross-sectional study design). A Preferred Reporting Items for Systematic Reviews and Meta Analyses (PRISMA)[72] 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 (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 (refer to Characteristics of Included Publications).

Included Studies for Question 1: Effectiveness and Accuracy

We identified 2 retrospective nonconcurrent cohort studies[73,74] and 11 cross-sectional diagnostic accuracy studies[73-83] to address this research question. The 2 cohort studies[73,74] and 1 diagnostic accuracy study[79] examined the effectiveness or accuracy of RapidAI for supporting health care providers' review of imaging results, while 10 diagnostic accuracy studies[73-78,80-83] 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[73] or including patients from a large multihospital network with several comprehensive stroke centres.[74] Seven diagnostic accuracy studies were single-centre and conducted in Australia,[76,83] the UK,[77,80] or the US.[73,79,81] The remaining diagnostic accuracy studies were multicentre and included patients from several hospitals in Australia[75] or the US[74] or included patient data from a variety of sources (such as recent cerebrovascular trials and hospitals in multiple countries).[78,82]

Patient Population

Across all studies, populations consisted of patients presenting with neurologic deficit or suspected acute stroke, all of whom underwent diagnostic imaging such as noncontrast 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[78] to 75[78] years. One cohort study[74] and 1 diagnostic accuracy study[78] 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)[73,74,77,80-83] reported the proportion of female participants, which ranged from 41.3%[77] to 55.2%.[73]
- Five studies[74,75,80,82,83] reported the proportion of male participants, which ranged from 44.4%[75] to 56.8%.[83]
- Two studies[77,83] reported the proportion of participants who were women, which ranged from 41.3%[77] to 43.2%.[83]
- Two studies[76,83] reported the proportion of participants who were men, which ranged from 49.6%[76] to 56.8%.[83]
- Three studies[74,78,79] 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)[73] reported participant race but did not indicate how it was recorded. No other PROGRESS-Plus criteria[42,43] (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[82] presented subgroup analyses for some PROGRESS-Plus criteria,[42,43] 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 Stroke platform (using Rapid HVS to detect LVO). Some of the included studies[73,75,76,83] provided high-level overviews of RapidAI's process for analyzing CT images to detect ICH or LVO. For example, Delora and colleagues[75] 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 were 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. 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[73-83] 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[83] 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 the 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 (refer to Critical Appraisal of Included Studies). Overall, the cohort studies exhibited critical risk of bias across all outcome domains. Ten diagnostic accuracy studies[73-82] 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[83] exhibited low risk of bias across all 4 domains. For applicability concerns, 10 studies[73-78,80-83] were labelled as high risk,

as they evaluated CT images by RapidAI alone, which differs from our primary review question and from how RapidAI is used in clinical practice (i.e., alongside clinician interpretation).

Risk of Bias in Nonrandomized Studies

In both cohort studies,[73,74] 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[73] applied adjustments for some measures of functional status, including for the effects of high cholesterol, heart disease, atrial fibrillation, therapies received, and US National Institutes of Health Stroke Scale (NIHSS) score on admission. However, the authors did not provide a rationale for how these factors were selected or 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[73,74] 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[73,76,77,79-81,83] 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,[82] they applied inappropriate exclusion criteria (e.g., only included patients who tested positive on the index test),[74] or their methods for selecting patients were unclear.[75,78] 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[74,76,78,83] 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 prespecified the relative vessel density thresholds for detecting LVO. Seven studies[73,75,77,79-82] exhibited an unclear risk of bias for the index test because there was insufficient information to determine whether they had prespecified positivity thresholds for detecting LVO or ICH. For applicability concerns, the index test domain for 1 study[79] 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[73-78,80-83] were labelled as high risk because they evaluated CT images by RapidAI alone, which differs from our primary review question and from how RapidAI is used in clinical practice (i.e., alongside clinician interpretation).

The reference standard domains of 5 studies[77,78,81-83] 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[73,75,79] exhibited high risk of bias. In 2 of these studies,[73,75] 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,[79] which is likely to inflate the

estimates of diagnostic accuracy for the index test. Three studies[74,76,80] 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,[73-83] 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 expected 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, the diagnostic performance of RapidAI may not accurately reflect situations where RapidAI correctly classified patients but the reference standard did not.

Seven studies[74-76,80-83] 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[79] or they excluded a considerable number of patients from the analysis.[77] The risk of bias for the flow and timing domain was unclear in 2 studies[73,78] 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.[73-83]

Table 2: Summary of Risk of Bias in the Included Diagnostic Accuracy Studies Using QUADAS-2[48]

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)[75]	Unclear	Unclear	High	Low	Low	High	Low
Slater et al. (2024)[76]	Low	Low	Unclear	Low	Low	High	Low
Chan et al. (2023)[77]	Low	Unclear	Low	High	Low	High	Low
Soun et al. (2023)[73]	Low	Unclear	High	Unclear	Low	High	Low
Yedavalli et al. (2023)[78]	Unclear	Low	Low	Unclear	Low	High	Low
Mallon et al. (2022)[80]	Low	Unclear	Unclear	Low	Low	High	Low
Schlossman et al. (2022)[81]	Low	Unclear	Low	Low	Low	High	Low
Adhya et al. (2021)[74]	High	Low	Unclear	Low	Low	High	Low
Dehkharghani et al. (2021)[82]	High	Unclear	Low	Low	Low	High	Low
Amukotuwa et al. (2019)[83]	Low	Low	Low	Low	Low	High	Low

Study citation	Risk of bias				Applicability concerns		
	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
ICH							
Eldaya et al. (2022)[79]	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[51] and the CLAIM guideline.[52] 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[51] and the CLAIM guideline[52] 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, which can have implications for the generalizability of the model to real-world settings (e.g., the use of inappropriate methods for splitting data can make a model's predictions susceptible to bias and inflate its performance)
- 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[51] and the CLAIM guideline,[52] 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 (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[73,74] (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;[73] ≤ 760 patients; low certainty) by 8.6 (standard deviation [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 the time interval between a patient's arrival at the hospital and the initiation of specific interventions, such as the administration of IV thrombolysis (i.e., door-to-intervention times) (1 study;[73] ≤ 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;[74] 146 patients; very low certainty). Door-to-intervention times are important in stroke care, as the potential benefits of therapies decrease over time.[18-20]

The evidence is very uncertain about the effects of both Rapid CTA (1 study;[74] 141 patients) and Rapid LVO (1 study;[74] ≤ 105 patients) on measures of functional status. It is also very uncertain whether Rapid LVO has any effect on response to therapy (1 study;[73] 80 patients).

Outcomes related to patient harms (e.g., rate of symptomatic ICH) 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
Time to intervention					
Radiology-report turnaround (minutes), mean (SD)[b]	Rapid LVO (Rapid v4.9)	≤ 760[c] (1 NRS)[73]	• Without Rapid LVO: 30.6 (29.9) • With Rapid LVO: 22 (35.1) • Difference: −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)[73]	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)[74]	• Without Rapid CTA: 92 (NR) • With Rapid CTA: 68 (NR) • Difference: −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					
Neurologic deficit (per NIHSS score), median (IQR)[g,h]	Rapid LVO (Rapid v4.9)	NR[c] (1 NRS)[73]	NIHSS scores (without vs. with Rapid LVO): • 36-hours posttreatment: 10 (5 to 18) vs. 11 (2 to 20) • At discharge: 5 (1 to 100) vs. 8 (2 to 20)	Very low *(due to risk of bias and imprecision)*	The evidence is very uncertain about the effect of CTA with Rapid LVO on posttreatment neurologic deficit.

Outcome and follow-up	Intervention	Participants (studies), N	Absolute effects	Certainty[a]	What happens
			• Change from admission to discharge: −7 (−2 to −13) vs. −3 (0 to −7) Between-group differences were NR for any outcome		
Proportion of patients with significant morbidity or mortality (defined as mRS score ≥ 5) at discharge (95% CI)[h,i]	Rapid LVO (Rapid v4.9)	105 (1 NRS)[73]	• Without Rapid LVO: 177 per 1,000 (NR) • With Rapid LVO: 233 per 1,000 (NR) • Difference: 55 more per 1,000 (103 less to 213 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 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)[74]	• Without Rapid CTA: 4.47 (NR) • With Rapid CTA: 3.90 (NR) • Difference: −0.57 (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)[74]	• Without Rapid CTA: 230 per 1,000 (NR) • With Rapid CTA: 343 per 1,000 (NR) • Difference: 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)[73]	TICI scores (without vs. 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.

Outcome and follow-up	Intervention	Participants (studies), N	Absolute effects	Certainty[a]	What happens
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 = CT angiography; IQR = interquartile range; LVO = large-vessel occlusion; mRS = modified Rankin Scale; NA = not applicable; NE = not estimable; NIHSS = US National Institutes of Health Stroke Scale; NR = not reported; NRS = nonrandomized study; SD = standard deviation; TICI = Thrombolysis in Cerebral Infarction; tPA = tissue plasminogen activator; vs. = versus.

[a]Detailed reasons for certainty of evidence ratings are provided in the Supporting Information document (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 being provided for the clinicians.

[c]The sample size for the analysis was not explicitly reported; as a result, 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*,[54] 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 NIHSS is a 15-item neurologic examination stroke scale used for evaluating stroke-related neurologic deficit. Total scores range from 0 to 42, with higher scores indicating more severe neurologic deficit.

[h]Based on the information provided in the publication, this analysis appears to include only 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.[57]

[k]The 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).

Sources: Soun et al.[73] and Adhya et al.[74]

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 (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 noncontrast 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;[79] 307 patients). The certainty of evidence was low, owing to risk of bias and imprecision.

The study by Eldaya and colleagues[79] 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[73,74] 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,[79] 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 noncontrast 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	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Certainty[a]
ICH					
Rapid ICH (version NR) with interpretation by a neuroradiologist	307 (1 study)[79]	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 (refer to Reasons for Certainty of Evidence Ratings).

Source: Eldaya et al.[79]

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., noncontrast CT or CT angiography). There was heterogeneity across the included studies for the types of LVO assessed, the modules and 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;[73,81] 574 patients), Rapid LVO v1.0 (1 study;[82] 217 patients), Rapid CTA as part of RapidAI v4.9 (1 study;[83] 477 patients), and Rapid CTA (version not reported) (1 study;[77] 88 patients) had sensitivity values ranging from 90% to 96% and specificity values ranging from 76% to 98%. Across these 5 studies,[73,77,81-83] the proportion of false-positives and false-negatives ranged from 1% to 13% and 1% to 20% of total tests, respectively (1 study[73] 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 assessed, study populations, reference standards used, and modules and versions of RapidAI being evaluated may have contributed to the variability in results. For example, the inclusion of distal occlusions (e.g., within the M2 segment of the middle cerebral artery) may have influenced measures of diagnostic performance, as they can be more difficult to detect.[84] 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 earlier in this report, 1 study[82] 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. The 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	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Certainty[a]
M1 MCA and ICA LVO					
Rapid LVO (RapidAI v4.9) alone	574 (2 studies)[73,81]	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)[82]	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)[83]	16	94 (86 to 98)	76 (72 to 80)	Moderate *(due to imprecision)*
Rapid CTA (version NR) alone	88 (1 study)[77]	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)[78]	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)[80]	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)[83]	22	92 (85 to 96)	81 (77 to 85)	Low *(due to risk of bias and imprecision)*
M1 MCA LVO					
RapidAI alone (version NR)	73 (1 study)[80]	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)[83]	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)[81]	NR	80 (NR)	NR	Insufficient information to judge

Index test	Participants (studies), N	% positive per reference standard	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Certainty[a]
Rapid CTA (RapidAI v4.9) alone	477 (1 study)[83]	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)[81]	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)[76]	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)[75]	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)[74]	52	80 (73 to 86)	71 (63 to 78)	Very low *(due to risk of bias and imprecision)*

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

[a]Detailed reasons for certainty of evidence ratings are provided in the Supporting Information document (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.[82] used case-control selection; as a result, 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).[78]

[e]Unclear which types of LVO were considered eligible. In Delora et al.,[75] the study population included patients with occlusions of the ICA, M1 MCA segment, and M2 MCA segment. The study population from Adhya et al.[74] included patients with occlusions of the ICA, carotid terminus, M1 MCA segment, and M2 MCA segment.

Sources: Delora et al.,[75] Slater et al.,[76] Chan et al.,[77] Soun et al.,[73] Yedavalli et al.,[78] Mallon et al.,[80] Schlossman et al.,[81] Adhya et al.,[74] Dehkharghani et al.,[82] and Amukotuwa et al.[83]

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 *WHO Guidance: Ethics and Governance of Artificial Intelligence for Health*[67] 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 AI that is responsive and sustainable, which are described in the accompanying implementation review. 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 expectations. Additionally, we drew from the ethical considerations proposed by the AI Task Force of the Society of Nuclear Medicine and Molecular Imaging,[68] 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[35] 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 life cycle. For example, during the initial design phase, issues such as algorithmic bias and lack of representation can arise. Later in the technology life cycle, during widespread implementation in 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.[66,67] The adoption of AI in 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 should be considered as a supportive tool that assists clinicians by providing 1 piece of information for diagnosing patients undergoing assessment for acute stroke. Because RapidAI does not act 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.[31]

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 concerning whether to undergo treatment or testing.[68] 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. Obtaining explicit consent for the use of RapidAI as a part of routine clinical care may be less relevant, as patients are often unaware of the specific software tools used to process their medical information while interacting with health care institutions. However, obligations to disclose the use of AI-enabled technologies where they are collecting and storing patient data to develop or train their machine-learning models remain appropriate and important practices in upholding patient autonomy and privacy.

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 about the use of AI software in their medical assessments, although this may often not be feasible, given the urgency of stroke assessment and intervention. 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[85] indicate that people perceive AI tools differently and support the disclosure of when and how AI tools are used. According to Herington and colleagues,[68] 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.[68]

Data privacy and security are important components of protecting human autonomy that must be safeguarded when implementing AI technologies.[66,67] 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. 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® (ownership, control, access, and possession), Manitoba Métis principles of OCAS (ownership, control, access, and stewardship), and Inuit Qaujimajatuqangit, must guide the respectful governance of data collected with, from, or about Indigenous Peoples in Canada. Additionally, there are other data governance frameworks and principles for equity-deserving groups that help establish data sovereignty and

promote accountability, fairness, and transparency. For example, the Engagement, Governance, Access and Protection Framework[86] guides the collection, management, analysis, and use of race-based data from Black communities. 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 further 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,[66,67] which are described in the accompanying implementation review. 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 ICH 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 of direct harms to patients. Inaccurate

diagnoses by RapidAI become problematic only if the diagnoses are not corrected or reinterpreted by clinicians before making care decisions or 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 life cycle of a technology, ensuring that decisions are clear and justifiable and that operations are understandable to all relevant parties.[68,87] These concepts highlight the importance of ensuring that 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.[87,88] Achieving both explainability and intelligibility can be facilitated through model transparency and by establishing a process to monitor the technology to ensure it performs as expected.[89] While Health Canada monitors postmarket data on the safety of medical devices (e.g., through the collection of medical devices incident reports),[90] we did not find information on who holds accountability if issues arise during the use of RapidAI and how institutions and regulators could respond to potential issues. 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 life cycle.

In 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.[91] 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 whether the effectiveness and accuracy of RapidAI are consistent and robust across diverse populations, we used PROGRESS-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 a result, we were unable to assess the representativeness and diversity of the training dataset and comment on considerations related to inclusivity. Furthermore, it was unclear whether RapidAI has undergone bias assessment to determine whether 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 for bias in the stroke-detection algorithms.

The implementation of RapidAI has the potential to improve access to stroke care and mitigate geographic inequities by increasing diagnostic efficiency, particularly when care is provided outside of centres with well-resourced stroke units. In health care settings with limited radiology expertise, imaging results are often initially reviewed by nonspecialists, and there may be considerable delays before radiologists with stroke expertise review the findings. RapidAI could help identify potential occlusions or ICH that may go undetected during the initial review and help to prioritize cases of highest urgency, which could lead to quicker diagnosis and intervention for those who need it most.

However, based on the evidence summarized in this review, it remains 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. For instance, in settings with limited access to CT scanners, the time required to perform imaging studies may have a greater impact on access to stroke interventions than the time needed to analyze and interpret CT images. Additionally, depending on the age of imaging equipment, patients may require multiple scans to obtain sufficient imaging data, resulting in further delays before intervention can occur. Consequently, while RapidAI has the potential to improve diagnostic efficiency, it is unlikely to fully remediate inequities in access to stroke care that arise from these broader systemic constraints.

The patient contributor raised access concerns about the availability of AI technologies in hospitals outside urban stroke centres. 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 whether its use would be restricted to certain facilities. The budget, personnel, infrastructure, and training requirements needed to implement RapidAI may limit its use to better-resourced hospitals or health care centres, despite its potential to improve some access to stroke care in rural and remote settings. Even if there were data to suggest the performance of RapidAI is robust across diverse patient populations (i.e., low risk of bias in the algorithm's performance), limited access based on geographic location could exacerbate existing health inequities.[92]

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 life cycle 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.[40] 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 in the following.

Evidence Gaps

In addition to the limitations in the evidence noted earlier in this report, 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.

Generalizability

This rapid review summarizes the results of 2 cohort studies[73,74] and 11 diagnostic accuracy studies[73-83] evaluating the effectiveness and accuracy of RapidAI for detecting stroke. However, only 3 of these studies[73,74,79] 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[73-78,80-83] 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[73,74] and 7 diagnostic accuracy studies[73,76,77,79-81,83] recruited participants from a single institution. While there was no strong indication that the findings from studies conducted in Australia,[75,76,83] the UK,[77,80] the US,[73,74,79,81] or studies that included patient data from a variety of sources[78,82] would not apply to settings in Canada, differences in stroke diagnosis approaches or patient populations may limit the generalizability of the evidence to the context in Canada.

We used PROGRESS-Plus[42,43] 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[73-77,79-83] 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).[73] 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 whether study populations included people from equity-deserving groups and whether the effects of RapidAI are generalizable, consistent, and robust across diverse patient populations. Mismatches between the study populations and target populations could lead to a risk that the performance of RapidAI, as described in the included studies, may not be applicable in all clinical settings (i.e., spectrum bias).[93]

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 a result, 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, and the Rapid NCCT Stroke platform, or described the intervention as RapidAI without any further details. These sources of heterogeneity limited our ability to synthesize results across studies, as often only 1 study was available for the individual comparisons.

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. In our literature search, we identified 2 studies[94,95] published in Spanish that may have contained relevant data. However, only the titles and abstracts of these publications[94,95] 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 preclinical 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 or representative of all patients who have experienced stroke. We were also unable to provide clinician perspectives or their experiences with AI-enabled stroke-detection software, such as RapidAI, without their participation in the engagement activities. 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 noncontrast 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 range 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., noncontrast 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, or very low, or there was insufficient information to judge certainty, primarily 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. It is important for clinicians who use RapidAI to understand the technology's diagnostic performance (e.g., estimates of sensitivity and specificity) and limitations, and to remain cautious of the potential for over-reliance on AI assessments.[96]

Our findings are generally aligned with a 2024 assessment by the UK National Institute for Health and Care Excellence (NICE)[30] on AI-derived software to help clinical decision-making in stroke, which included but was not specific to RapidAI for stroke detection. The assessment by NICE[30] 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 UK National Health Service while more evidence is generated and that it should be used only alongside health care professional interpretation of CT scans to reduce the risk of incorrect results.[30]

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. We reached out to the manufacturer but were unable to obtain updated pricing information on this technology.

From publicly available sources, we identified resource considerations regarding the implementation of RAPID. Two approaches exist to implement RAPID, each with its own unique resource considerations. The first is a single-site (hub) model that would involve loading the application on a local server housed behind

the organization's firewall. In this approach, the CT and/or MRI scanner and PACS modality are connected to the server, with CT and/or MRI images sent to an automated software server through Digital Imaging and Communications in Medicine (DICOM) connections from the scanners; the automated server then pushes the processing results to PACS and/or other digital medical imaging management systems (e.g., the Emergency Neuro Imaging Transfer System in Ontario) for viewing and storage by a computer or mobile device.[24] The alternative approach is called a hub-and-spoke model. This differs in that a single server, installed behind the firewall of 1 organization, is shared across all sites. The CT and/or MRI scanners at each spoke site are configured to push images through a DICOM connection through a dedicated e-health tunnel unique to each site to a centralizer server; this server then pushes the processed results back through the dedicated e-health tunnel through a DICOM connection to the local spoke PACS for viewing and storage.[24] Costs differ between these approaches. The hub model would incur server-installation costs, an annual hub licensing fee, and a 1-time setup and training fee. In contrast, the hub-and-spoke model would also incur an annual hub licensing fee, but the server-installation costs and 1-time setup and training fee would be shared across sites. The hub-and-spoke model has additional connection requirements (e.g., bandwidth and connection) and requires an external threat risk assessment conducted by an external provider. Furthermore, a data-sharing agreement would need to be in place among hub-and-spoke sites.[24]

Ontario Health published an HTA[26] in 2020 on the use of automated CT perfusion imaging to aid in the selection of patients with acute ischemic stroke for mechanical thrombectomy. According to its budget impact analysis, the annual cost of a licence for the RAPID neuroimaging platform in 2019 depended on the number of connected scanners and was between $27,500 (for 2 scanners) and $32,500 (for unlimited scanners) per hospital.[26] There was an added cost in the first year due to 1-time fees for the initial implementation and optimization of RAPID, as well as for training staff ($12,350 for hospitals with 2 or more scanners).[26] 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.[26] This costing information was for the full RAPID platform, which included several modules for conducting various cerebrovascular diagnostic procedures, in addition to automated CT perfusion. Although this report does not address RapidAI's capacity to inform treatment selection (e.g., selecting patients for reperfusion therapy), the findings of the HTA[26] by Ontario Health may provide insights into the potential costs of implementing the RAPID platform.

Of note, the assessment by Ontario Health led the Hyperacute Stroke Steering Committee to recommend a provincial sourcing strategy for RAPID for all Ontario stroke sites and community hospitals. It included an agreement with the manufacturer for a standardized pricing schedule, with further price decrease possible for all participants in the agreement as tier levels are achieved when new sites join.[97] The Ontario agreement is a single-site model that includes all stroke modules as part of its licence, with the option to include 1 scanner or unlimited scanners. Several software features are included, with add-ons such as unlimited number of cases for licensed scanners, unlimited users on the RAPID online training and certification platform, and live webinar trainings sessions with clinical experts.[97] According to a news release, nearly 50% of Ontario stroke sites are currently using RapidAI as of February 2024.[98]

Applying the Scottish Framework

We used the Scottish Health Technologies Group's HTA framework[35] as a guide throughout this rapid review, helping us determine the types of information to summarize. Overall, we were able to provide information to address most of the items outlined in the framework.

For domain 1 ("The technology and its value"), we provided an overview of RapidAI, including how the technology works and a summary of the health condition (i.e., stroke), a description of the care pathway and how it could change with the introduction of RapidAI, the technology's value proposition, and the clinical need for it. These items were captured in the Context and Policy Issues section of this report. For domain 2 ("Safety, acceptability and credibility"), we included a description of RapidAI's compliance with relevant regulatory standards (it is licensed for sale in Canada as a Class III medical device), conducted an assessment of the potential harms of RapidAI, and described how RapidAI could create or exacerbate health inequities through our analysis of ethics and equity considerations. Domain 3 items ("Demonstrating the performance of the technology") were addressed through our review of the evidence on the effectiveness and accuracy of RapidAI for stroke detection (i.e., research question 1).

However, we were unable to address several items in Scottish Health Technologies Group's HTA framework.[35] We did not determine how intended users, including clinicians and patients, were involved in the design and development of the technology. Our review focused on comparative studies, none of which described the initial phases of developing RapidAI. Other published information, such as preclinical studies, might provide this information. We were unable to provide information about service users' and caregivers' experiences with RapidAI. We attempted to engage a clinician with professional experience using RapidAI to learn about their perceptions and whether the technology met their expectations. However, this engagement was unsuccessful due to time constraints, so we could not comment on users' experiences of RapidAI. We did not address the environmental considerations related to using RapidAI. Our review did not assess the potential environmental impacts of changes to care pathways due to the implementation of RapidAI. We were unable to determine a measurement plan for ongoing data collection and evidence generation. The clinical evidence summarized in this review did not provide any information on how the developer plans to monitor the performance and safety of RapidAI as it is adopted in clinical practice. We were also unable to determine the cost-effectiveness of RapidAI and the potential budget impact of implementing RapidAI for stroke detection in Canada as per domain 4 ("Capturing the cost and value for money of the technology"). Although our review included a research question specific to cost-effectiveness (i.e., research question 2), we did not identify any economic evidence that met our eligibility criteria, so we cannot draw conclusions on the cost-effectiveness of RapidAI for stroke detection.

Considerations for Future Research

Robustly designed and transparently reported studies with low risk of bias that enrol 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, race, smoking status, diabetes, hypertension) through experimental design or through statistical analysis.

Clinicians exploring the implementation of RapidAI should consider conducting silent testing trials before adopting it in routine clinical practice. In these trials, AI tools are evaluated in real-time on prospective patients, while care decisions continue to be made independently of the tool's outputs.[99] This approach could gather data effectively on RapidAI's performance in real-world settings, accounting for potential differences in imaging equipment, patient characteristics, and other parameters that could create a risk for data drift or model drift, leading to performance degradation.[99,100] Additionally, independent assurance bodies could play a role in the continued postmarket testing and monitoring of AI-enabled software in clinical settings, to validate their performance and manage risks of overfitting, unintended bias, or degradation of the model.[101-103]

To better understand the diagnostic accuracy of RapidAI for its intended use (i.e., to complement, rather than replace clinicians), additional studies that evaluate the performance of RapidAI alongside clinician interpretation are needed, particularly for detecting LVO. Future studies that evaluate the performance of RapidAI when used by clinicians with varying levels of expertise or specialties (e.g., resident, clinical fellow, radiologist, neuroradiologist) could establish whether RapidAI offers value as a training tool or as an overall improvement to standard clinical practice. Of note, diagnostic accuracy studies of RapidAI alone are informative for testing and monitoring its performance over time (e.g., to detect any performance degradation) and are needed as well.

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[51] and the CLAIM guideline[52] 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,[104] people at lower levels of socioeconomic status,[105] people living in rural or remote areas,[14] and people with various racial and ethnic identities.[106] 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 detect algorithmic bias, provide insights 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 outcomes important to patients, 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 such 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.[79] 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 from reduced radiology-report turnaround times needs to be balanced with the risk of undesirable effects resulting from false-positives and false-negatives. Although false-positives and false-negatives were infrequent in the 1 study,[79] 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 in 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,[73] 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 a standard 8-hour shift for a radiologist whose 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; faster triage, decision-making, and treatment initiation; and, ultimately, better patient outcomes. However, real-world efficiency gains may be tempered based on the type and amount of radiologist workload, which in turn depends on various factors, such as disease prevalence in specific communities of care and whether RapidAI is used across multiple sites with centralized reporting. Future research addressing the limitations of the current evidence could provide greater clarity on how RapidAI impacts patients, clinicians, and decision-makers.

Acknowledgements

This document was externally reviewed by the following content experts who have granted permission to be cited:

Eduardo Portela, MD
Neuroradiologist, The Ottawa Hospital
Assistant Professor, University of Ottawa

Jaron Chong, MD, MHI, FRCPC
Assistant Professor, Western University

Jay Shaw, PT, PhD
Research Director, Artificial Intelligence, Ethics & Health, Joint Centre for Bioethics, University of Toronto

We acknowledge with gratitude the contributions of a patient who preferred to remain anonymous.

Conflicts of Interest

Jaron Chong is the chair of the Artificial Intelligence Standing Committee with the Canadian Association of Radiologists and an ad hoc member of the Scientific Advisory Committee on Digital Health Technologies with Health Canada. No other conflicts of interest were identified.

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. PubMed
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. Gittins M, Lugo-Palacios D, Vail A, et al. Stroke impairment categories: A new way to classify the effects of stroke based on stroke-related impairments. *Clin Rehabil.* 2021;35(3):446-458. PubMed
5. 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. PubMed
6. 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. PubMed
7. 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. PubMed
8. 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. PubMed
9. 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. PubMed
10. 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. PubMed
11. 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.
12. 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.
13. 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. PubMed
14. Fleet R, Bussières S, Tounkara FK, et al. Rural versus urban academic hospital mortality following stroke in Canada. *PLoS One.* 2018;13(1):e0191151. PubMed
15. 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.
16. 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.
17. Kawano H, Hirano T. Minimal Imaging Requirements. *J Neuroendovasc Ther.* 2023;17(11):243-256. PubMed
18. Saver JL. Time is brain--quantified. *Stroke.* 2006;37(1):263-266. PubMed
19. 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. PubMed
20. 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. PubMed

21. Musuka TD, Wilton SB, Traboulsi M, Hill MD. Diagnosis and management of acute ischemic stroke: speed is critical. *CMAJ.* 2015;187(12):887-893. PubMed

22. RapidAI. https://www.rapidai.com/. Accessed 2024 Jun 13.

23. Straka M, Albers GW, Bammer R. Real-time diffusion-perfusion mismatch analysis in acute stroke. *J Magn Reson Imaging.* 2010;32(5):1024-1037. PubMed

24. CorHealth Ontario. Automated Stroke Imaging Software Package (Ischemaview RAPID) Implementation Frequently Asked Questions and Lessons Learned. Toronto (ON): Ontario Health; 2020: https://www.corhealthontario.ca/AutomatedSoftware-Lessons-Learned-and-Frequently-Asked-Questions-Version-8.0-Final-Version.pdf. Accessed 2024 Aug 26.

25. Zhang Y, Song S, Li Z, Huang B, Geng Y, Zhang L. The Application of Software "Rapid Processing of Perfusion and Diffusion" in Acute Ischemic Stroke. *Brain Sci.* 2022;12(11). PubMed

26. 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. PubMed

27. Campbell BC, Yassi N, Ma H, et al. Imaging selection in ischemic stroke: feasibility of automated CT-perfusion analysis. *Int J Stroke.* 2015;10(1):51-54. PubMed

28. Albers GW, Marks MP, Kemp S, et al. Thrombectomy for Stroke at 6 to 16 Hours with Selection by Perfusion Imaging. *N Engl J Med.* 2018;378(8):708-718. PubMed

29. Nogueira RG, Jadhav AP, Haussen DC, et al. Thrombectomy 6 to 24 Hours after Stroke with a Mismatch between Deficit and Infarct. *N Engl J Med.* 2018;378(1):11-21. PubMed

30. Artificial intelligence (AI)-derived software to help clinical decision making in stroke. *Diagnostics guidance (DG57].* London (GB): NICE; 2024: https://www.nice.org.uk/guidance/dg57. Accessed 2024 Jun 13.

31. *RAPID User's Manual.* Menlo Park (CA): iSchemaView, Inc.

32. RapidAI. Frequently Asked Questions. Menlo Park (CA): iSchemiaView, Inc.: https://www.rapidai.com/faq#technical.

33. 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.

34. News: International Health Technology Assessment Collaboration Expands. Ottawa (ON): Canada's Drug Agency; 2023: https://www.cadth.ca/news/international-health-technology-assessment-collaboration-expands. Accessed 2024 Jul 24.

35. 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.

36. 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.

37. 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. PubMed

38. 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. PubMed

39. 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. PubMed

40. 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. PubMed

41. 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.

42. 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. PubMed

43. 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.

44. 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. PubMed

45. 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.

46. 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. PubMed

47. 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. PubMed

48. 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. PubMed

49. 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.

50. 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. PubMed

51. 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. PubMed

52. 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. PubMed

53. 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.

54. 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.

55. 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.

56. Deeks JJ TY, Macaskill P, Bossuyt PM. Chapter 4: Understanding test accuracy measures. In: Deeks JJ BP, Leeflang MM, Takwoingi Y (editors), ed. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*: Cochrane; 2023.

57. Scherer R. PropCIS: Various Confidence Interval Methods for Proportions. 2018; https://cran.r-project.org/web/packages/PropCIs/index.html. Accessed 2024 Jun 20.

58. 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. PubMed

59. 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. PubMed

60. 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. PubMed

61. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. *BMJ.* 2004;328(7454):1490. PubMed

62. 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. PubMed

63. Schünemann HJ, Mustafa RA, Brozek J, et al. GRADE guidelines: 21 part 2. Test accuracy: inconsistency, imprecision, publication bias, and other domains for rating the certainty of evidence and presenting it in evidence profiles and summary of findings tables. *J Clin Epidemiol.* 2020;122:142-152. PubMed

64. 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. PubMed

65. 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. PubMed

66. 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.

67. Guidance W. *Ethics and governance of artificial intelligence for health.* World Health Organization. 2021.

68. 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. PubMed

69. 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.

70. 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. PubMed

71. 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. PubMed

72. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ.* 2021;372:n71. PubMed

73. 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. PubMed

74. 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. PubMed

75. 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. PubMed

76. 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.

77. Chan N, Sibtain N, Booth T, et al. Machine-learning algorithm in acute stroke: real-world experience. *Clin Radiol.* 2023;78(2):e45-e51. PubMed

78. 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. PubMed

79. Eldaya RW, Kansagra AP, Zei M, et al. Performance of Automated RAPID Intracranial Hemorrhage Detection in Real-World Practice: A Single-Institution Experience. *J Comput Assist Tomogr.* 2022;46(5):770-774. PubMed

80. 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. PubMed

81. 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. PubMed

82. 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. PubMed

83. 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. PubMed

84. Bathla G, Pillenahalli Maheshwarappa R, Soni N, et al. CT Perfusion Maps Improve Detection of M2-MCA Occlusions in Acute Ischemic Stroke. *J Stroke Cerebrovasc Dis.* 2022;31(6):106473. PubMed

85. Park HJ. Patient perspectives on informed consent for medical AI: A web-based experiment. *Digit Health.* 2024;10:20552076241247938. PubMed

86. Engagement, Governance, Access, and Protection (EGAP): A Data Governance Framework for Health Data Collected from Black Communities. Toronto (ON): Black Health Equity Working Group; 2021: https://blackhealthequity.ca/wp-content/uploads/2021/03/Report_EGAP_framework.pdf. Accessed 2024 Aug 14.

87. 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. PubMed

88. Linardatos P, Papastefanopoulos V, Kotsiantis S. Explainable AI: A Review of Machine Learning Interpretability Methods. *Entropy (Basel).* 2020;23(1). PubMed

89. Ennab M, McHeick H. Designing an Interpretability-Based Model to Explain the Artificial Intelligence Algorithms in Healthcare. *Diagnostics (Basel).* 2022;12(7). PubMed

90. Health Canada. Incident reporting for medical devices: Guidance document. Ottawa (ON): Government of Canada; 2021: https://www.canada.ca/content/dam/hc-sc/documents/services/drugs-health-products/reports-publications/medeffect-canada/incident-reporting-medical-devices-guidance-2021/incident-reporting-medical-devices-guidance-2021-en.pdf. Accessed 2024 Aug 22.

91. Varkey B. Principles of Clinical Ethics and Their Application to Practice. *Med Princ Pract.* 2021;30(1):17-28. PubMed

92. Ratwani RM, Sutton K, Galarraga JE. Addressing AI Algorithmic Bias in Health Care. *JAMA.* 2024. PubMed

93. Willis BH. Spectrum bias—why clinicians need to be cautious when applying diagnostic test studies. *Fam Pract.* 2008;25(5):390-396. PubMed

94. 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.

95. 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.

96. Quinn TP, Senadeera M, Jacobs S, Coghlan S, Le V. Trust and medical AI: the challenges we face and the expertise needed to overcome them. *J Am Med Inform Assoc.* 2021;28(4):890-894. PubMed

97. Ontario C. *Automated CT Perfusion Imaging in the Diagnosis and Treatment of Acute Ischemic Stroke.* 2023.

98. Media Release: Brightshores Launches Rapid Automated CT Perfusion Software to Elevate Stroke Care in the Region. Owen Sound (ON): Brightshores Health System; 2024: https://www.brightshores.ca/media-release-brightshores-launches-rapid-automated-ct-perfusion-software-to-elevate-stroke-care-in-the-region/. Accessed 2024 Aug 26.

99. Kwong JCC, Erdman L, Khondker A, et al. The silent trial - the bridge between bench-to-bedside clinical AI applications. *Front Digit Health.* 2022;4:929508. PubMed

100. Finlayson SG, Subbaswamy A, Singh K, et al. The Clinician and Dataset Shift in Artificial Intelligence. *N Engl J Med.* 2021;385(3):283-286. PubMed

101. Shah NH, Halamka JD, Saria S, et al. A Nationwide Network of Health AI Assurance Laboratories. *JAMA.* 2024;331(3):245-249. PubMed

102. Canada Health Infoway. Toolkit for Implementers of Artificial Intelligence in Health Care. Toronto (ON): Government of Canada; 2023: https://www.infoway-inforoute.ca/en/component/edocman/resources/artificial-intelligence/3998-toolkit-for-implementers-of-artificial-intelligence-in-health-care?Itemid=101. Accessed 2024 Aug 22.

103. Canada H. Good Machine Learning Practice for Medical Device Development: Guiding Principles. 2021; https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/good-machine-learning-practice-medical-device-development.html. Accessed 2024 Aug 22.

104. 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. PubMed

105. 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. PubMed

106. 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. PubMed

AI Implementation Review

List of Tables

Abbreviations

AI	artificial intelligence
API	application programming interface
CDA-AMC	Canada's Drug Agency
DTAC	Digital Technology Assessment Criteria
DHIEX	digital health information exchange
DHT	digital health technology
HTA	health technology assessment
IT	information technology
MFA	multifactor authentication
ML	machine learning
MLMD	machine learning–enabled medical device
NHS UK	National Health Service
NICE UK	National Institute for Health and Care Excellence
OWASP	Open Web Application Security Project
PHI	personal health information
PIPEDA	*Personal Information Protection and Electronic Documents Act*
SaMD	software as a medical device
SME	small and medium-sized enterprises

Key Messages

What Is the Issue?

- Globally, we are seeing a widespread increase in the interest, development, and use of artificial intelligence (AI)–enabled medical devices. Comprehensive evaluation through health technology assessment (HTA) can ensure that digital health technologies (DHTs), including AI-enabled medical devices, are adequately equipped to balance benefits and harms, while being interoperable and equitably accessible to people living in Canada.
- In the UK, a checklist called Digital Technology Assessment Criteria (DTAC) is used as an add-on component to HTAs to capture additional considerations for the implementation of DHTs. The 5 core areas of DTAC are clinical safety, data protection, technical security, interoperability, and usability and accessibility. In Canada, we currently do not have a DTAC equivalent that can be used as an add-on to traditional HTA.
- This implementation review is needed to assist health systems in Canada in preparing for the uptake of AI-enabled medical devices, as these technologies pose new challenges. We assessed whether the safeguards and assessment criteria captured by DTAC and other AI-related resources are in place to inform decision-making around the digital infrastructure elements of implementation.

What Did We Do?

- We conducted an implementation review, using a phased approach, to determine whether DTAC can be applied to the health care context in Canada to inform the implementation of DHTs and to identify any additional implementation considerations specific to the use of AI-enabled medical devices in Canada. We integrated ethics and equity considerations across both phases of the review.
- In phase 1, we applied DTAC to the health care context in Canada by determining whether we have equivalent or similar measures, strategies, and policies in place to implement DHTs safely.
- In phase 2, an information specialist searched for literature to identify implementation guidance specific to AI and relevant to Canada to supplement DTAC. One reviewer screened publications for inclusion based on predefined criteria, incorporated relevant information into tables, and summarized the findings narratively.
- We leveraged patient engagement activities conducted in a concurrent Canada's Drug Agency review of a specific AI-enabled medical device in stroke detection to learn from a patient contributor with lived experience of a hemorrhagic stroke. We learned about her experience, perspectives, priorities, and thoughts about using AI in clinical decision-making.

What Did We Find?

- With some caveats, we found that many of DTAC's assessment criteria have equivalent or similar guidance for the health care context in Canada. Some exceptions are derived from the differences in Canada's current governance and health care structure. Further investigation is required to

understand whether certain policies in Canada provide sufficient coverage to fulfill DTAC's criteria (e.g., clinical safety).

- We identified several considerations for implementing AI-enabled medical devices, with many having underlying ethical and equity implications. Much of the identified guidance emphasizes implementation considerations that apply to the AI system's entire life cycle, including the most prevalent consideration: ensuring AI-enabled medical devices are monitored, maintained, and sustainable. Examples of additional considerations include AI data governance and data protection; transparency and explainability; and inclusiveness, equity, and minimization of bias.
- The patient contributor highlighted several considerations relevant for this review, such as data protection and privacy as well as accessibility and equity.

What Does This Mean?

- We have identified key considerations for AI-enabled medical devices that health care decision-makers may consider for the safe and successful implementation of AI in health care in Canada.
- While Canada has DTAC-equivalent or similar measures, strategies, or policies in place, we identified a need for a checklist like DTAC that senior decision-makers can use. This checklist could be an adaptation of DTAC and could include additional implementation considerations for AI-enabled medical devices to ensure that these technologies meet the minimum baseline standards set out by DTAC and inform the next steps for the safe and successful implementation of AI-enabled medical devices in Canada.
- This implementation review for all AI-enabled medical devices is to be used alongside reviews of specific AI technologies, including the concurrent review of RapidAI, and will serve as a foundational report to be tailored for each AI topic and updated with the latest developments in the regulation and other aspects of management of AI in the context of Canada.

Context and Policy Issues

What Are AI-Enabled Medical Devices?

AI-enabled medical devices, a type of DHT, are advancing rapidly and generating much hope and hype.[1] AI uses algorithms or models to perform tasks and mimic human behaviours, such as learning, making decisions, and making predictions.[2] When a medical device incorporates AI algorithms and machine-learning (ML) models to enhance its functionality and performance, it is often described as an AI-enabled or ML-enabled medical device (MLMD).[2,3] An example of an AI-enabled medical device is the software platform RapidAI,[4] which was assessed alongside this review. RapidAI provides tools for medical imaging by using AI to facilitate viewing, processing, and analysis of CT images.[4] The RapidAI software broadly aims to help clinicians assess patients with suspected medical conditions and determine the appropriate treatment.[4] While DHTs promise to improve various outcomes (e.g., better access to health care), evaluating and implementing them presents new and unique challenges. AI likely poses all the challenges (e.g., data considerations,

interoperability issues) characteristic of other classes of DHTs and additional ones (e.g., algorithmic fairness and biases, black-box or continuous-learning nature),[5-7] serving as a useful test case for addressing the many challenges DHTs pose.

How Are AI-Enabled Medical Devices Regulated in Canada?

In Canada, medical devices are regulated by the *Medical Devices Regulations*, which are established under the authority of the *Food and Drugs Act*.[8,9] The regulations use a risk-based approach to regulating products within their scope, with devices classified into 1 of 4 classes (i.e., Class I represents the lowest risk and Class IV, the highest).[10] Software has become an increasingly important part of many products, integrated widely into digital platforms, including those used for medical purposes.[10,11] In 2019, Health Canada provided guidance documents to clarify how software as a medical device (SaMD) fits into Health Canada's regulatory framework for medical devices, including examples of software using AI.[10,12] SaMD is defined as software intended to be used for 1 or more medical purposes that performs these purposes without being part of a hardware medical device.[10,12]

All health care organizations, including regulatory bodies and HTA agencies, in Canada and internationally face similar challenges associated with DHTs, including AI-enabled medical devices, that go beyond each organization's or jurisdiction's purview. As a result, there has been a strong interest in working together and aligning approaches across organizations and jurisdictions,[13] and work has been done to align guidance from regulatory agencies in Canada, the UK, and the US (e.g., MLMD guiding principles).[14,15]

Why Is It Important to Do This Review?

Globally, we are seeing an increase in medical devices relying on software incorporating AI.[3] Given the inherent nature of AI as a disruptive technology in health care, its comprehensive assessment through HTA is essential to ensure that DHTs are adequately equipped to balance benefits and harms, while being interoperable and equitably accessible for people living in Canada.

Canada's Drug Agency (CDA-AMC) has recently established information-sharing and collaborative relationships with various organizations, including an international partnership with HTA bodies.[16] From this partnership, CDA-AMC learned of the Scottish Health Technology Group's evidence framework, which outlines an approach to DHT assessment.[17] This includes an HTA framework[18] and DTAC[19] as an add-on component. The Scottish group's approach to add-on DTAC is consistent with the UK National Institute for Health and Care Excellence (NICE)'s approach (i.e., it mandates DTAC for use in the UK National Health Service [NHS]). This review of implementation considerations applied DTAC[19] to the health care setting in Canada. The accompanying review of RapidAI applied the Scottish Health Technology Group's HTA framework.[18] The application of the Scottish framework allowed us to leverage existing work to ensure alignment and harmonization across organizations and to gain efficiency and sustainability. We plan to share our experiences with the framework and the criteria with the international partners.

DTAC is a checklist that captures additional considerations for the implementation of DHTs not captured by traditional HTA; in the UK setting, the checklist represents the NHS's minimum baseline standards for the use of a DHT.[19] For example, it includes 5 core areas to establish whether the product is clinically safe to use

from risk-management perspectives (i.e., under “clinical safety”); collects, stores, and uses data compliantly (i.e., under “data protection”); meets industry best-practice security standards (i.e., under “technical security”); exchanges data with other systems well (i.e., under “interoperability”); and follows best practice and meets user needs (i.e., under “usability and accessibility”).[19] As of September 2024, Canada does not have a DTAC equivalent for health care systems to use at the point of procurement or as part of an ongoing monitoring process for DHTs, including AI-enabled medical devices. While DTAC questions are generally specific to the UK setting, we expected DTAC could be applied to the health care context in Canada to inform the implementation of DHTs.

Objectives

This report aims to assist health systems in Canada in preparing for the uptake of AI-enabled medical devices by assessing whether appropriate safeguards and assessment criteria are in place to inform implementation decision-making. The key objectives of this implementation review are to identify and describe the digital infrastructure elements of implementation considerations specific to the use of AI-enabled medical devices in Canada.

The target audience of this work is large and includes regulators, government officials, health care decision-makers, clinicians, health care professionals, DHT developers, and researchers, among others.

Research Question

What are the implementation considerations for the use of AI-enabled medical devices in Canada?

Methods

Study Design

We conducted an implementation review using a 2-phase approach:

1. We applied DTAC to the health care context in Canada by determining whether we have equivalent or similar measures, strategies, and policies in place to implement DHTs in Canada safely (phase 1).
2. We conducted a literature review to identify implementation toolkits, guidance, and recommendations specific to AI and relevant to Canada to supplement DTAC, in case there are any additional considerations important for AI-enabled medical devices use in Canada (phase 2).

We integrated patient input and ethics and equity considerations across both phases of the implementation review.

This report is not a systematic review and did not involve a critical appraisal of the literature. It is not intended to provide recommendations for or against the use of AI in health care. Thus, conclusions or recommendations about the value of or place in therapy for AI are outside of this report’s scope.

Phase 1: Apply DTAC

For this review, we focused on applying all assessment criteria (i.e., Sections C and D) of DTAC to the health care context in Canada. Section C focuses on DTAC's core assessment criteria, which includes technical questions about clinical safety, data-protection, technical security, and interoperability criteria. Section D focuses on key usability and accessibility principles.

For each component of Section C and D of DTAC, 1 researcher reviewed each question and:

- determined what the question is asking and who is responsible for addressing it in the context of the UK
- investigated who is responsible for overseeing the element identified by the question in the context of Canada (e.g., federal government, provincial/territorial government, international organization)
- investigated what measures, strategies, and policies are in place in Canada, to address whether they are regulated or mandated in some capacity (e.g., by Health Canada) and whether they are equivalent or similar to those in the UK.

One reviewer obtained publicly accessible documents from government websites (e.g., the NHS in the UK, Government of Canada) and organizational websites (e.g., International Organization for Standardization). One reviewer extracted and populated these data into tables in Microsoft Word and progressed to phase 2.

Phase 2: Conduct Literature Review

Literature Search Methods

An information specialist conducted a focused internet search using keywords describing AI in health care and AI-enabled medical devices, by browsing and searching sources listed in relevant sections of Grey Matters: A Practical Tool For Searching Health-Related Grey Literature.[20] Sections included websites of health technology agencies in Canada, major international health technology agencies, regulatory agencies, and international health organizations to identify information on the implementation considerations for AI in health care. Searching was completed on May 21, 2024. The information specialist limited the search to English-language documents.

Screening and Guidance Selection

One reviewer screened the literature search results in Microsoft Word and reviewed the full text of all potentially relevant publications. Table 1 presents the selection criteria.

Table 1: Selection Criteria

Criteria	Description
Concept	Implementation considerations for the use of AI-enabled medical devices[a]
Context	Canada and health systems comparable to those in Canada[b]
Types of Sources	Implementation toolkits, guidance, and recommendations (e.g., guidance from reputable health organizations or professional associations)
Publication Date	No date limits

Criteria	Description
Language	English

AI = artificial intelligence.

[a]We define AI-enabled medical devices as those used for medical purposes that have a direct involvement in and direct impact on patient care to achieve specific health outcomes for those patients, aligned with Health Canada's approach to defining software as a medical device[10] and NICE's Evidence Standards Framework for Digital Health Technologies and Tier C technologies.[21]

[b]Countries that qualified as having health systems comparable to those in Canada for this review included (in alphabetical order): Australia, France, Germany, Netherlands, and the UK, all of which are Organisation for Economic Co-operation and Development (OECD) members and have a mix of public and private health care.[22]

Exclusion Criteria

One reviewer excluded publications that did not meet the selection criteria outlined in Table 1. If a guidance document included outdated evidence (e.g., describes a law that no longer applies), we included only the current and relevant information for this review.

Data Extraction

We used a pragmatic approach to extract and synthesize the data. One reviewer extracted and populated all relevant data into tables in Microsoft Word, including bibliographic details, such as organization name, year of publication, country, target audience (e.g., regulators, DHT developers), and principal components addressed (e.g., clinical safety, data privacy, interoperability).

Analysis, Synthesis, and Reporting of Findings

During phase 1, the reviewer first populated Table 2 in the Supporting Information document (refer to Main Findings) with details about what each assessment question is asking, who is responsible in the context of the UK and Canada, and any equivalent measures, strategies, and policies that are in place in Canada. During phase 2, the reviewer conducted a descriptive analysis to supplement the data table from phase 1 (i.e., Table 2 in Supporting Information document [refer to Main Findings]) by adding implementation consideration themes for using AI-enabled medical devices in Canada next to the corresponding sections (i.e., grouping similar considerations together, determining theme names based on these groupings). Table 3 in the Supporting Information document (refer to Main Findings) provides detailed examples of these additional implementation considerations for AI-enabled medical devices. The reviewer produced a narrative summary of the results presented in the Supporting Information document (refer to Main Findings). Finally, the reviewer provided general conclusions about whether DTAC's assessment considerations are already in place in the various jurisdictions in Canada.

Ethics and Equity Considerations

The integration of ethics and equity considerations was driven by an analysis of key items with ethics and equity implications in DTAC, augmented by key tools and frameworks (e.g., UNESCO Recommendation on the Ethics of AI;[23] *WHO Guidance: Ethics and Governance of Artificial Intelligence for Health*),[24] as well as the European Network for Health Technology Assessment (EUnetHTA) Core Model 3.0 Ethics Domain[25] and the Equity Checklist for HTA.[26]

We used prompts and guiding principles from these sources to help identify and reflect on ethics and equity considerations in the implementation of AI and DHTs relevant to patients, providers, and health systems.

With the assistance of a reviewer with ethics expertise, we synthesized the findings of these prompts and descriptive analyses into analytic categories representing the key ethical and equity considerations related to AI and the implementation of DHTs 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 leveraged the engagement activities conducted for the RapidAI review to incorporate input from patients and clinicians in this project.

For the RapidAI review, 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 was 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: the patient contributor's lived experience with a stroke and AI in interpreting image results. Patient involvement was guided by the Guidance for Reporting Involvement of Patients and the Public (version 2) Short Form reporting checklist,[27] which is outlined in Table 4 of the Supporting Information document (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 AI implementation 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 consulted for the RapidAI review reviewed the project plan. The same experts reviewed the draft version of this report, and their feedback was incorporated into the final version of this report.

Feedback Opportunity

The draft version of this report was posted on the CDA-AMC website to allow interested parties to provide feedback. We considered all feedback received and revised the report accordingly. Therefore, this final version of the report reflects the comments and suggestions from the peer reviewers and from the feedback opportunity.

Summary of Findings

Summary of Included Guidance

For phase 1 of this review, we used DTAC as our reference document and applied it to the health care context in Canada. We identified most of the relevant information to contextualize each assessment question in DTAC to Canada from search engine results of government websites (e.g., provincial, territorial, and federal) and websites of international organizations (e.g., International Organization for Standardization).

For phase 2 of this review, our literature search results included 13 relevant guidance documents that described implementation considerations for AI-enabled medical devices. Among these documents, we focused our data extraction, analysis, and synthesis of the general health guidance documents from Canada (i.e., directly applicable to the health care context in Canada, not specialty-specific). We also used *WHO Guidance: Ethics and Governance of Artificial Intelligence for Health* and other identified international guidance to provide additional examples of implementation considerations as supporting evidence from international sources.

Guidance from Canada

We identified 8 relevant guidance documents from Canada.

Canada Health Infoway has published a comprehensive toolkit for implementers of AI in health care.[28] This toolkit provides an overview of the issues related to implementing and using AI solutions in health care. It offers strategic and operational guidance for designing responsible AI projects and governance programs. The toolkit includes thorough checklists for identifying and addressing bias in AI systems, assessing AI vendors, and data testing and monitoring.[28] For example, the AI vendor assessment checklist could be

used as a guide when evaluating vendors for any AI-related medical devices in the health care context. The checklist asks several questions under the following categories: commitment to responsible innovation, value alignment, inclusivity, nature of technology, risk management, data quality testing, data access, data sharing, transparency and explainability, human oversight, accountability, and knowledge transfer.[28] Canada Health Infoway has also published an additional digital health solutions guidance document regarding privacy and security for vendors, health care organizations, and patients.[29] This guideline contains a small subsection about health care AI, and its content largely overlaps with Canada Health Infoway's *Toolkit for Implementers of AI*.[29]

Health Canada, the UK Medicines and Healthcare products Regulatory Agency, and the US FDA jointly published 10 guiding principles that can inform the development of good ML practices, promoting safe, effective, and high-quality AI- and MLMDs.[15] These agencies have also released related documents for guiding principles for transparency for MLMDs[14] and predetermined change control plans.[30] Health Canada has since released draft premarket guidance for an ML system of an MLMD.[31] Health Canada's premarket guidance targets manufacturers submitting a new or amendment application for Class II, III, or IV MLMDs under the regulations.[31] Health Canada also expects applications for an MLMD to include information about how the manufacturer has adopted good ML practices[15] within the organization and implemented them throughout the product life cycle.[31] For example, transparency requirements should consider those involved in patient health care (e.g., patients, users, health care providers, and regulators) across the medical device's life cycle.[31]

The Vector Institute, an Ontario-based AI research institute, produced a *Health AI Implementation Toolkit*[32] to highlight common deployment barriers for those looking to implement innovative health AI research in a clinical context. This guidance[32] includes a health AI implementation checklist organized according to 4 primary considerations:

- Do you have a mature, validated AI model?
- What is your technical integration strategy?
- What is your change management strategy?
- Do you have a data-governance and sustainability plan?

This guidance primarily targets individuals (e.g., researchers, clinicians, health professionals) who have developed a robust, mature health AI model or application and are looking to deploy their solution in a clinical environment.[32]

In 2022, the Canadian Law & HTA Working Group published *Legal Guidance for Canadian HTA Bodies*.[87] This publication is intended to support HTA bodies in Canada in incorporating legal analysis into their evaluations. Its primary target audience is nonlawyers working in HTA bodies. For example, it mentions that AI poses an increased risk of privacy breaches and acknowledges that the existing legislation attempts to balance the patient's right to privacy with the needs of health care workers, administrators, and policy-makers to use and disclose the information to provide treatment, conduct research, perform billing, and deliver effective health care. This guidance provides trigger questions for HTA bodies to consider around

such topics. Additionally, we identified other specialty-specific guidance from reputable associations (e.g., the Registered Nurses' Association of Ontario for clinical practice in a digital health environment and the Canadian Association of Radiologists for ethical and legal issues related to AI in radiology).[34,35]

International Guidance

We identified 5 international guidance documents aimed at health systems comparable to Canada,[22] including Australia, France, and the UK.

WHO published a report that endorses key ethical principles for the use of AI for health.[24] WHO[36] also developed a framework that provides an overview of considerations in evaluating clinical evidence regarding AI-SaMD, aiming to help formulate a consensus for guiding validation, generating evidence, and reporting across the total product life cycle in a global health context. Given the nature of this guidance, we focused on extracting information related to implementation considerations (e.g., deployment, postdeployment monitoring and surveillance) rather than clinical evaluation.[36]

The UK's Medicines and Healthcare products Regulatory Agency produced guidance to ensure that medical device regulation is fit for purpose for software, including AI.[37] Moreover, the UK has further guidance on the regulation of AI as a medical device,[38] and NICE's evidence standards framework for DHTs for the HTA setting now includes evidence requirements for AI and data-driven technologies with adaptive algorithms.[21] The Australian government regulates AI as a medical device when it is used for diagnosis, prevention, monitoring, prediction, prognosis, treatment, or alleviation of disease, injury, or disability.[39] It also has evidence requirements for AI software throughout the product life cycle and includes robust postmarket monitoring practices to ensure continued device performance and model accuracy.[39] The Haute Autorité de Santé in France has a guide for submitting a medical device and health technology for HTA.[40] It provides a 7-page table for applicants to complete when the medical device has at least 1 ML process.[40]

We provided details regarding the publications included in the Supporting Information document (refer to Characteristics of Included Guidance). We described additional details about the relevant AI considerations in the subsequent section and Table 3 of the Supporting Information document (refer to Main Findings).[28] We also provided a list of potential references of interest that did not meet our eligibility criteria in the Supporting Information document (refer to References of Potential Interest), which includes publications focusing on knowledge-user perspectives and examples of preliminary frameworks or guidance as additional resources that may be useful to others.

Implementation Considerations

Implementation considerations identified from DTAC and guidance from Canada and international sources were organized by the 5 core areas of DTAC: clinical safety, data protection, technical security, interoperability, and usability and accessibility.

Clinical Safety (Section C1 of DTAC)

Applying DTAC's Section C1 to the Health Care Context in Canada

The clinical safety section of DTAC provides considerations to establish that the DHT product is clinically safe to use beyond the typical clinical benefits and harms considerations in HTA from risk-management perspectives (e.g., clinical risk-management activities for the development and maintenance of health information technology [IT] systems).[19] In the UK, the clinical safety considerations fall under the national government (i.e., NHS, Medicines and Health care products Regulatory Agency). The identified considerations for Canada are federally governed (i.e., Health Canada).

Specifically, for the UK, DTAC asks whether clinical risk-management activities undertaken for the DHT product comply with DCB0129, a standard for organizations responsible for health IT development and maintenance. DCB0129 outlines the clinical risk-management standards for manufacturers of health IT systems. In Canada, an international standard similar to DCB0129 is ISO 14971, which applies to manufacturers of medical devices worldwide but is not limited to cybersecurity considerations.[41] Certification to the ISO 13485 Quality Management Systems standard is also required by Health Canada's regulations (i.e., a standard that looks to ISO 14971 for risk-management guidance).[42] For medical devices in Canada, manufacturers of Class II, III, and IV devices must obtain a medical device licence (i.e., a Medical Devices Active Licence) before selling a medical device in Canada.[43] The application for a medical device licence is completed through Health Canada.[44] Class I medical devices do not require a licence but are monitored through the establishment licensing process.[43] Health Canada has *Medical Devices Regulations* (SOR/98 to 282) outlining safety and effectiveness requirements;[9] SaMD guidance, including risk classification of medical devices; and guidance for incident reporting.[31,43,45-47] The NHS provides a clinical risk-management system template that outlines the processes to ensure that all health care IT used to support care within the organization is developed, implemented, and used safely. We did not identify an equivalent clinical risk-management system template for the context in Canada. We did identify a landing page on Health Canada's website devoted to compliance and enforcement of medical devices (e.g., forms, guidance, policies, and laws).[48] Additional investigation is required to determine whether the forms, guidance, policies, and laws that Health Canada provides on its website provide sufficient coverage to fulfill DTAC's criteria for clinical safety.

Additional Considerations for AI-Enabled Medical Devices

Monitoring, Maintenance, and Sustainability Throughout the AI Product Life Cycle

WHO Guidance: Ethics and Governance of Artificial Intelligence for Health highlights the importance of promoting human well-being, human safety, and the public interest in AI technologies.[24] WHO emphasizes that AI technologies should satisfy regulatory requirements of safety, accuracy, and efficacy before deployment and that measures should be in place to ensure quality control and quality improvement (i.e., monitoring, maintenance, and sustainability).[24]

We identified additional guidance that provides AI considerations for establishing clinical safety before deployment and as a part of maintenance and monitoring after deployment. The fundamental reason for regularly monitoring and maintaining AI products is to ensure their relevance, accuracy, efficacy, and safety.[32] Thus, AI considerations about monitoring and maintenance have been included under this clinical safety

section unless they were explicitly linked to data protection, technical security, interoperability, or usability and accessibility. This concept also coincides with the idea of iterative implementation, which is described in the literature as a way to overcome AI implementation barriers in primary health care in Canada.[49] These considerations are presented together (i.e., life cycle approach) to showcase specific examples from the identified evidence.

- **Performance and validation:** Health Canada's draft guidance reiterates the need for manufacturers of MLMDs to provide performance or bench testing or software verification and validation information (e.g., descriptions of the chosen performance metrics, acceptance criteria, and operating point or threshold with clinical and risk-based justifications; evidence to demonstrate that the ML system performs as intended and meets expected performance requirements when integrated as part of the medical device system or software). Health Canada states manufacturers should also provide the appropriate clinical evidence, including clinical validation studies, to support their device's safe and effective clinical use.[31] The Vector Institute[32] provides further considerations to ensure that the AI model validation process is comprehensive and rigorous and the model is robust and generalized well to new and unseen data. They state that this can improve outcomes while maximizing patient safety and privacy.[32] They also state that AI systems should incorporate feedback and continuously improve based on user needs and technological advancements to remain relevant and beneficial in the long-term.[32] Performance monitoring is a component of the sustainability strategy for evaluating the AI tool.[32] It involves regularly monitoring the system's performance and impact on patient outcomes to identify any issues that need to be addressed and to demonstrate the system's value to its users.[32,34] When the model has a "human in the loop," emphasis should be placed on the performance of the human-plus-AI team, including human-factor and human-interoperability considerations, rather than on the model's performance in isolation.[15,34]
- **Risk assessment and management:** The identified guidance emphasizes providing evidence of risk management to address risk of harms (e.g., by addressing questions such as "If the system does not operate as intended, could it negatively impact patients' standard of care?" and "Could the output from the AI system result in denying services for a patient?" and monitoring performance degradation, such as data drift).[15,28,31,39]
- **Data quality testing:** The Vector Institute suggests implementing "processes to regularly check, clean, and improve the data quality to ensure its accuracy and consistency" (e.g., by addressing questions such as "Does the vendor have documented processes to test data for completeness, representativeness, and accuracy?" and "Has an assessment been done to see if the source of data is suitable for the intended purpose?").[28,32]

Table 3 in Supporting Information document provides additional examples (refer to Main Findings).

During the interview with the patient contributor, she discussed clinical safety. Specifically, the patient contributor hoped AI could result in a faster and more accurate diagnosis. She hoped this might help clinicians initiate the most appropriate treatment sooner, which could reduce the damage caused by stroke (i.e., her lived experience with a medical condition) and improve outcomes (e.g., mortality). The patient

contributor posited that, while ensuring the accuracy of the AI technology was a concern, she was curious about whether AI could accurately identify issues earlier than a clinician or prevent human error.

Data Protection (Section C2 of DTAC)

Applying DTAC's Section C2 to the Health Care Context in Canada

The data-protection section of DTAC provides considerations to establish that the DHT product collects, stores, and uses data compliantly (e.g., ensuring assessment and mitigation strategies and a Data Protection Officer are in place).[19] In the UK, data-protection laws for personal information and health information are nationally regulated. In Canada, privacy legislation is governed at the federal level and at the provincial and territorial level, depending on the data type and jurisdiction. Through the Office of Privacy Commissioner of Canada, privacy legislation (*Personal Information Protection and Electronic Documents Act [PIPEDA],* S.C. 2000, c 5) generally applies to private-sector organizations that collect, use, and disclose personal health information (PHI) in the course of a commercial activity.[50] There is also guidance on how *PIPEDA* applies to processing personal data across borders.[51] Alberta, British Columbia, and Quebec have their own private-sector privacy laws that have been deemed substantially similar to *PIPEDA*.[28,52] Ontario, New Brunswick, Nova Scotia, and Newfoundland and Labrador have also adopted "substantially similar legislation" regarding the collection, use, and disclosure of PHI.[50] Collection, use, and disclosure of PHI by a physician is governed by that jurisdiction's health privacy legislation, if such laws exist. Freedom of information and protection of privacy and health information acts at the provincial and territorial level are additional health privacy laws in Canada.[28] In the UK, some businesses require registering with the Information Commissioner and paying a data-protection fee. In Canada, all businesses that handle personal information are subject to privacy laws (e.g., *PIPEDA*), but there does not appear to be a registration or data-protection fee requirement.[50] Businesses may complete a *PIPEDA*'s voluntary self-assessment tool to evaluate and improve their personal information management systems and practices.[53]

The Government of Canada has tabled Bill C-27, Digital Charter Implementation Act, 2022.[54,55] This act aims to strengthen Canada's private-sector privacy law, create new rules for responsible AI development and deployment, and continue advancing the implementation of *Canada's Digital Charter*.[55] Bill C-27 would repeal Part 1 of the *PIPEDA* and enact the *Consumer Privacy Protection Act,* the *Personal Information and Data Protection Tribunal Act,* and the *AI and Data* Act.[55] The Office of the Privacy Commissioner of Canada's submission on Bill C-27 states that it follows the former Bill C-11, the Digital Charter Implementation Act, 2020, which also proposed amendments to *PIPEDA* and died on the Order Paper when Parliament was dissolved on August 15, 2021, in advance of the 2021 federal election.[56] Thus, until Bill C-27 receives Royal Assent, *PIPEDA* remains Canada's current federal privacy law.

DTAC states that Data Protection Officers are required in the UK, depending on the type of organization and core activities.[19] For example, a Data Protection Officer is needed if an organization's core activities involve processing health data.[19] For Canada, the federal and provincial personal information protection acts all suggest designating a privacy representative.[53,57-59] The UK government has a *Data Security and Protection Toolkit* that all organizations must use if they have access to NHS patient data. Since the provincial and territorial governments in Canada govern the management of PHI, there is no centralized federal toolkit

for data security and protection of patient data. DTAC also identifies data-protection impact assessments, which help organizations assess and demonstrate compliance with data-protection obligations.[19] In the UK, guidance for such assessments is provided by the Information Commissioner at the national level. In Canada, it is more nuanced, given that guidance for personal information would be at the federal level and guidance for health information would be at the provincial and territorial level, and data-protection impact assessments are mandated in some jurisdictions (e.g., British Columbia, Quebec) but not others (e.g., Alberta).

Because of the nuances of privacy laws in Canada, mobilizing high-quality, diverse datasets of PHI from across Canada for AI training can be challenging.[28] Despite this, it appears that Canada has equivalent policies to address DTAC's considerations for data protection that reflect laws and the structure of the health care systems in Canada.

Additional Considerations for AI-Enabled Medical Devices

AI Data Governance and Data Protection

A critical part of deploying an AI application is having a robust data-governance strategy.[32] The Vector Institute,[32] Canada Health Infoway,[29] and Health Canada[15] suggest establishing comprehensive organizational policies and procedures for AI data governance to ensure AI's responsible use and deployment, such as data quality, privacy, security, access, sharing, life cycle management, and regulatory compliance. Appropriate data governance should enable access to the right data for the right people at the right time.[29,32]

Several ethical considerations relate to AI data governance, training, and operationalization, including consent, de-identification, privacy law compliance, and data transmission security.[28] The Canadian Association of Radiologists also emphasizes the ownership of electronic medical records and the secondary use of de-identified medical data are complex issues that will likely depend on the type of use.[35] It mentions that tools and policies are required to facilitate and standardize the anonymization of medical images, which is likely relevant for other specialties as well.[35] Moreover, it highlights that emerging legislation advocates for robust de-identification methodologies and cautions against the risk of re-identification.[28]

Canada Health Infoway suggests that AI system implementers create a code of ethics and principles, provide ethics training, and establish an ethics board or subcommittee.[28] Further considerations include protecting human autonomy (i.e., maintaining human control of health care systems and medical decisions) and allowing individuals to opt out of being included in the data used to train or run the AI system.[28] The Canadian Association of Radiologists states that "respect for data privacy requires balancing the principle of beneficence and justice (to improve medical care for others via secondary use of an individual's data) versus the principle of respect for autonomy (as regards the concept of free and ongoing informed consent)."[35] Individual data participants can request access and correction to their personal information under privacy laws in Canada. These rights can be difficult to implement where data are co-mingled in large databases for AI algorithm training (i.e., it is technically challenging to segment data). To this end, the Canadian Association of Radiologists describe that, to facilitate the development of AI applications in health care, a transition from "informed consent" for specific data uses to "broad consent," "opt-out consent," and/or "presumed consent"

for more general data uses is required.[35] Using examples from other countries, the Canadian Association of Radiologists describes what each consent model might look like. For example, the European Union's *General Data Protection Regulation* allows patients to give general "consent to certain areas of scientific research, when in keeping with recognized ethical standards" (i.e., broad consent). It mentions that the data custodian and the general parameters of potential data use are identified and agreed to, but the exact projects and users of the data are not known. It presents the example of "opt-out" consent (e.g., allowing certain items to be excluded) as an option that will likely be difficult to feasibly implement, as it would be cumbersome and logistically challenging. The Canadian Association of Radiologists provide a further recommendation "to advocate for general adoption of revised forms of consent (such as "broad consent") for appropriately safeguarded secondary use of data for AI in health care in Canada."[35] Importantly, the Canadian Association of Radiologists also recommends advocating for programs to educate the public and increase awareness of the benefits of sharing personal health data that is fully anonymized as well as harm-reduction strategies.[35] Thus, the use of AI in health care has prompted consideration of how models of data governance and informed consent (including modified, dynamic, and broad consent, or waivers of consent) can evolve in an attempt to balance considerations of privacy and autonomy, the purported benefits of the use of AI technologies, and public interest.[24]

Multidisciplinary Data-Governance Team Throughout the Product Life Cycle

The identified guidance emphasizes the importance of a multidisciplinary data-governance team that is leveraged throughout the life cycle of the AI product. The team can be enhanced by including members from diverse areas such as privacy, security, IT, legal, clinical, ethics, industry, science, and senior management.[28,32] This concept also coincides with participatory co-design, which the literature describes as a way to overcome AI implementation barriers in primary health care in Canada.[49]

Monitoring, Maintenance, and Sustainability

Certain considerations must be made to ensure proper data protection, such as conducting regular audits to ensure ongoing compliance with all relevant regulations, developing a security control profile to monitor and mitigate the identified risks to privacy and security, and maintaining a process for assessing risks of re-identification.[29,32]

There are, however, situations where some patients may be less concerned about privacy and data protection considerations and believe that these may be superseded by access to care. Specifically, the patient contributor suggested that, in a crisis, she felt that most people would not be thinking of privacy and data sharing; instead, she would be focused on diagnosis, treatment, and survival. 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. In addition, the patient contributor expressed curiosity about whether patients will be informed that their clinicians used AI in their diagnosis. She did not know whether the technology was in use at the time of her stroke. Moreover, she briefly shared concerns about the safety of information and storage reliability.

Technical Security (Section C3 of DTAC)

Applying DTAC's Section C2 to the Health Care Context in Canada

The technical security section of DTAC provides considerations to establish that the DHT meets industry best-practice security standards (e.g., ensuring that testing and measures against cyber threats and proper requirements are in place).[19] In the UK, the related considerations for technical security are nationally governed through the National Cyber Security Centre. The technical security considerations where we identified an equivalent measure, strategy, or policy in Canada were federally governed through Innovation, Science and Economic Development Canada, Communications Security Establishment Canada, or Health Canada.

For the UK, DTAC asks for a Cyber Essentials certificate. CyberSecure Canada from Innovation, Science and Economic Development Canada is a federal cyber certification program that aims to raise the cyber security baseline among small and medium-sized enterprises (SMEs) in Canada, increase consumer confidence in the digital economy, promote international standardization, and better position SMEs to compete globally.[60] To be eligible for certification, the organization must implement the security controls in the National Standard CAN/CIOSC 104:2021 Baseline cyber security controls for SMEs.[61] Moreover, the Canadian Centre for Cyber Security (part of Communications Security Establishment Canada) has published updated cyber security guidance for SMEs.[62]

DTAC also asks for a summary report of an external penetration test of the product that includes Open Web Application Security Project (OWASP) top 10 vulnerabilities from within the previous 12-month period. The OWASP Foundation is a not-for-profit organization with educational resources, guidelines, and open-source tools to help improve the security of SMEs' software. It provides guidance on the OWASP top 10 vulnerabilities[63] and penetration testing methodologies applicable to both the UK and Canada.[64] For additional direction, Health Canada has published a guidance document outlining premarket requirements for medical device cyber security, which mentions structured penetration testing.[41] The Canadian Centre for Cyber Security guides the top measures to enhance cyber security for SMEs[65] (ITSAP.10.035[66]). Compared with guidance from Canada, the UK's National Cyber Security Centre provides more details on producing clean and maintainable code.[67]

DTAC asks to confirm whether all privileged accounts have appropriate multifactor authentication (MFA). In Canada, we have guidance on MFA for organizations and individuals,[68] and strong user identification (e.g., MFA) is enforced as a top measure to enhance cyber security for SMEs (i.e., recommended but not mandated).[65] DTAC also requests that logging and reporting requirements be clearly defined.[19] The Canadian Centre for Cyber Security provides information about network security logging and monitoring (ITSAP.80.085), including a checklist of best practices (i.e., recommended not mandated).[69] It is unclear whether Canada has recommended or mandated guidance on load testing. The details about load testing in the DTAC guidance are sparse, limiting our ability to find equivalent measures. Overall, it appears that we have technical security measures equivalent to DTAC in the health care context in Canada.

Additional Considerations for AI-Enabled Medical Devices

Monitoring, Maintenance, and Sustainability

The Vector Institute emphasizes that measures to secure data should be defined (e.g., encryption methods, access controls, firewalls) and, to ensure data security, regular audits should be conducted (i.e., a component of the sustainability strategy of the AI tool).[32]

Interoperability (Section C4 of DTAC)

Applying DTAC's Section C4 to the Health Care Context in Canada

The interoperability section of DTAC establishes how well the product exchanges data with other systems.[19] In the UK, the related considerations for interoperability are nationally governed through the NHS and Government Digital Services. In Canada, we identified similar measures, strategies, and policies related to interoperability criteria at the federal (e.g., Treasury Board of Canada Secretariat) and provincial and territorial levels (e.g., Ontario Health). Moreover, organizations such as the Healthcare Information and Management Systems Society use maturity models at the organizational level to measure interoperability and support digital health implementation.[70]

DTAC includes questions about application programming interfaces (APIs) and whether the AI product exposes any API or integration channels for other consumers. APIs provide an efficient and controlled way to make data accessible to other systems remotely.[71,72] Still, direct system-to-system access enabled through APIs increases the risk and impact of a security breach.[71,72] The Government of Canada provides API guidance[73] and API security best practices[74] that describe industry standards for secure interoperability (e.g., OAuth 2.0,[75] among others). We also identified the digital health information exchange (DHIEX) standard, a regulatory framework allowing Ontario Health to define and implement the health information standards and requirements for interoperability specifications.[76] It is unclear whether other provinces or territories have similar frameworks. In the UK, an NHS number is used to identify patient record data, which does not apply to Canada's current health care structure.[77] Relatedly, the federal Minister of Health introduced Bill C-72 in June 2024, An Act Respecting the Interoperability of Health Information Technology and to Prohibit Data Blocking by Health Information Technology Vendors, also called the Connected Care for Canadians Act.[78,79] Bill-72 is intended to enable a more connected health system by ensuring that health IT that is licensed, sold, or supplied as a service by a vendor is interoperable and to prohibit data blocking by the vendor.[78,79] Bill-72 seeks to establish a framework that allows for the safe and secure exchange of health data, enabling patients better access to, and more control over, their PHI.[78] DTAC includes the assessment question, "Is your product a wearable or device, or does it integrate with them? If yes, provide evidence of how it complies with ISO/IEEE 11073 Personal Health Data (PHD) Standards." The International Organization for Standardization is an industry standard that applies in Canada.[80] Moreover, Canada Health Infoway has been responsible for licensing, defining, and maintaining pan-Canadian standards that promote interoperability,[81] and produced a shared pan-Canadian interoperability roadmap.[82] This roadmap highlights systemic interoperability challenges within Canada, including the interoperability (or lack thereof) of systems to share electronic health information, a component of implementation that affects providers, patients, care coordination, and collaboration.[82] The roadmap includes pan-Canadian interoperability strategic goals,

including reducing data blocking and easing portability; improving provider access to patient data at point-of-care; enabling patient access to their health record; and improving care coordination and collaboration.[82] Canada Health Infoway also produced a primer for sharing PHI for interoperability, which includes an overview of privacy laws in Canada and practical approaches to privacy for interoperability.[83] Overall, Canada appears to have similar API guidance for interoperability considerations.

Additional Considerations for AI-Enabled Medical Devices

Technical Infrastructure and Integration

We identified guidance highlighting the importance of having a technical integration strategy and defining integration points, as applicable.[32] The Vector Institute describes an integration point as where an ML solution interfaces with existing health care infrastructure.[32] The guidance emphasizes that it is essential to have a proper integration point with the clinical workflow and systems for health care deployment of AI solutions because it ensures that these systems "effectively augment medical professionals' decision-making, enhancing efficiency, and improving patient outcomes without disrupting existing processes that need to be preserved."[32] For example, the guidance advises asking the following question: "How does AI product fit into the current clinical workflow and IT systems?"[32]

Monitoring, Maintenance, and Sustainability

Health Canada's draft guidance for MLMDs reiterates the need for manufacturers to provide performance or bench testing or software verification and validation information (e.g., evidence to support intercompatibility with all supported input and output devices).[31] The following are a few examples of interoperability considerations identified by the Vector Institute as part of a sustainability strategy for an AI tool: the AI system should fit seamlessly into the existing clinical workflows to minimize disruption to the current system (to reduce resistance and improve the likelihood of long-term adoption) and AI applications should be user-friendly, following best practices for user-centred design, should integrate smoothly with existing systems to encourage widespread adoption, and should be easy to use and understand for end-users (e.g., clinicians, medical staff, patients).[32] The Vector Institute also suggested that designing the simplest model possible with the fewest features can improve model interoperability.[32]

Usability and Accessibility (Section D1 of DTAC)

Applying DTAC's Section D1 to the Health Care Context in Canada

The usability and accessibility section of DTAC provides considerations to ensure the DHT product has followed best practices and meets user needs. In the UK and Canada, the related considerations for usability (e.g., simplicity and reliability of service) and accessibility (e.g., engaging with users and understanding their needs, ensuring everyone can use the service and can use and contribute to open standards) are governed at the national level. This core area of DTAC draws on the UK's NHS Service Standard.[84] The Government of Canada's Treasury Board of Canada Secretariat published Digital Standards[85] that generally complement the points described in the NHS standard. These digital standards intend to improve government services in the digital age and are targeted for government practice (i.e., not directed to industry or consumer products specifically).[85] However, the content, headings, and descriptions provided in these standards largely overlap with the NHS Service Standard. For example, DTAC emphasizes the need to understand users and their

needs in the context of health and social care by asking, “Do you engage users in the development of the product?”[19] This question aligns with NHS Service Standard Point 1[84] and the Government of Canada’s Digital Standards, Design with Users.[85] There are some instances where the wording in the UK Service Standard and Canada’s Digital Standards do not fully overlap (e.g., choose the right technology, D1.9; operate a reliable service, D1.11). However, this seemed to be covered by the information presented in the interoperability section of DTAC and by specific AI considerations from the identified guidance (e.g., technical infrastructure and integration, monitoring, maintenance, and sustainability). The Government of Canada Digital Standards are the closest equivalent identified for the health care context in Canada, and it is unclear whether there are other service standards in place for industry and/or consumer products.

DTAC also describes the need for the service to be usable by everyone and comply with the Web Content Accessibility Guidelines, an international standard, which are used in the UK and Canada. It is noted that the Treasury Board of Canada Secretariat is currently reviewing its Standard on Web Accessibility. As a part of a commitment to an accessible and barrier-free Canada, it is currently recommended that organizations adopt the Harmonized European Standard (EN 301 549) and adhere to the guidance available in the *Guideline on Making Information Technology Usable by All*.[86] Therefore, this DTAC section seems to have counterparts in Canada’s health care context.

Additional Considerations for AI-Enabled Medical Devices

Transparency, Explainability, and Intelligibility

WHO Guidance: Ethics and Governance of Artificial Intelligence for Health[24] highlights the importance of ensuring transparency, explainability, and intelligibility. WHO states transparency requires sufficient information to be published or documented when designing and deploying an AI technology that considers the individuals involved in patient health care across the product’s life cycle (e.g., patients, health care providers, users, regulators).[24,28] Canada Health Infoway emphasizes the value of explainable AI in health care as enhancing trust (e.g., “providing explanations to patients will help foster trust and reassure them about the AI system’s safety and equity”); compliance (e.g., “regulations are now moving toward requiring explanations be provided when AI systems decide or support decision-making. Knowing how to effectively explain AI decisions to a variety of stakeholder [user] groups will help mitigate the risks of non-compliance.”); and enhanced service delivery (e.g., “providing explanations to patients helps ensure that AI-supported care delivery is patient-centric”).[28] Further, Canada Health Infoway[28] provides an extensive list of considerations regarding:

- best practices for transparency and explainability policies and procedures (e.g., “establish a process of stakeholder [user] engagement to assess what would constitute a meaningful explanation,” “identify within your organization those that will be accountable to manage and oversee explainability requirements”)
- questions for AI vendor assessment (e.g., “What mechanisms does the vendor propose to make the system more transparent and explainable? For example, disclosing what training data was used, which variables contributed most to a specific outcome, and what data quality tests were conducted to ensure that the system performs as intended.”)

- transparency- and explainability-related questions organizations can ask when assessing the potential risks of an AI system (e.g., “Is the system’s technique compatible with the required level of explainability [e.g., consider that using deep learning techniques may yield lower levels of explainability]?”).[28]

Inclusiveness, Equity, and Minimizing Bias

WHO Guidance: Ethics and Governance of Artificial Intelligence for Health[24] highlights the importance of inclusiveness, equity, and mitigation of bias. Bias threatens inclusiveness and equity because it represents a departure, often arbitrary, from fair and equitable treatment and outcomes.[24] WHO states “inclusiveness requires that AI used in health care is designed to encourage the widest possible appropriate, equitable use and access, irrespective of age, gender, income, ability or other characteristics.”[24] Strategies must be in place to identify and reduce bias in AI technologies. For example, if health-related AI systems are trained using nonrepresentative data (e.g., they exclude equity-deserving populations), this could reinforce or worsen existing discriminatory treatment in the health care system.[87]

We identified guidance that showcases some examples of necessary risk management for bias. It should be ensured that clinical study participants and datasets represent the intended patient population.[15,32] For example, “data collection protocols should ensure that the relevant characteristics of the intended patient population (for example, in terms of age, gender, sex, race, and ethnicity), use, and measurement inputs are sufficiently represented in a sample of adequate size in the clinical study and training and test datasets so that results can be reasonably generalized to the population of interest.”[15] Mismatches between the study populations and target populations could lead to a risk that the performance of the AI technology may not be applicable in all clinical settings (i.e., spectrum bias).[88] The Vector Institute states that representativeness helps ensure the “model can generalize well to new and unseen data, improving predictive accuracy.”[32] AI developers should ensure that AI data, especially training data, do not include sampling bias and are accurate, complete, and diverse.[24] This also helps assess usability and identify circumstances where the model may underperform.[15] According to Health Canada’s draft guidance, manufacturers should apply sex- and gender-based analysis and consider the unique anatomic, physiological, and identity characteristics of patients over the MLMD’s life cycle (i.e., design, risk management, data selection and management, development and training, testing and evaluation, clinical validation, transparency, and postmarket performance monitoring).[31] “This includes considering sex and gender, racial and ethnic minorities, elderly and pediatric populations, and pregnant people and collecting and analyzing disaggregated data on subpopulations in clinical studies, training data, and test data, as appropriate.”[31] WHO also states “the effects of use of AI technologies must be monitored and evaluated, including the disproportionate impact on specific groups of people when they mirror or exacerbate existing forms of bias and discrimination” (i.e., reiterates monitoring, maintenance, and sustainability considerations).[24]

We acknowledge that algorithmic bias can be considered a clinical safety or usability consideration. For this review, we aligned bias considerations with usability, as discussed in the Government of Canada’s Digital Standards,[85] described in the usability and accessibility section. Nevertheless, we recognize the importance of mitigating algorithmic biases for the overall clinical safety of patients. Moreover, Canada Health Infoway includes a thorough checklist for identifying and addressing bias in AI systems to promote responsible and

ethical AI development and use.[28] The checklist provides several questions at each stage of the life cycle, including ideation and feasibility, design and development, validation, and deployment and monitoring. For example, the following questions are considerations during deployment and monitoring: “define triggers that will automatically alert those responsible for oversight and monitoring of the AI system should the AI system begin to behave unexpectedly; if automatic triggers are not possible, define how often the AI system should undergo re-validation to ensure it remains free from bias and robust; document acceptable use criteria for the AI system to ensure the system is only deployed in an appropriate context; consider algorithmic auditing by third parties to ensure that your AI system remains free from bias; if indicators of unwanted bias are found in the system, immediately retrain or redevelop the system.”[28]

An element of equity also involves equitable access to technologies. The patient contributor discussed the topic of equitable access to AI technologies by expressing concern about the accessibility of RapidAI (i.e., the AI-enabled medical device used as an example) outside of major stroke centres. She wondered whether all major hospitals could benefit from this technology to assist in quickly triaging (and potentially transferring) patients. She also had concerns about access to services in rural and remote community hospitals. This example aligns with Section D1.4 of DTAC: “make sure everyone can use the service.”[19]

Responsibility and Accountability

WHO states that responsibility can be assured by the application of “human warranty.”[24] This implies evaluation by patients and clinicians in developing and deploying AI technologies. WHO states that there should be accountability when something does go wrong in an application of an AI technology.[24] The Canadian Association of Radiologists also asserts that guidelines are required before deploying AI tools in hospitals to reduce the potential harm and liability for malpractice (e.g., in case of medical error that includes AI) and that contingency plans must be in place to deal with temporary pause or cessation of systems use (e.g., a triage screening program is paused for re-calibration; the backlog of clinical work must be processed in a timely fashion).[35]

User Buy-In and Organizational Readiness

With the disruptive nature of AI, a change management strategy for implementers of AI-enabled medical devices is important for the success of its deployment. The Vector Institute described the importance of a change management strategy and dedicated an entire section on how to effect change.[32] The Vector Institute emphasizes the need to have buy-in from:

- clinical users (e.g., patients, clinicians, organizations)
- informatics team (e.g., a multidisciplinary team of in the areas of data science and analytics, IT infrastructure, data management, application development, information security, clinical informatics, and quality assurance)
- senior leadership from multiple departments.[32]

A change management strategy touches on many of the AI considerations already discussed (e.g., risk assessment, regulatory compliance, ethics, and compliance) but is described at the organizational level with additional considerations, including clinical champions, shared vision, budget allocation, and return on

investment.[32] They provide examples of risk assessment questions regarding organizational readiness (e.g., "Will the AI system affect current employee roles and responsibilities? Will it lead to workforce redundancies? Will employees need new training? Are existing policies adequate to cover the safe and effective operationalization of the system?").[32]

Monitoring, Maintenance, and Sustainability

To strengthen the regulation of AI in or as medical devices, the Canadian Law & HTA Working Group suggests that an increased premarket review and continuous postapproval surveillance are needed.[87] We identified considerations for monitoring AI systems to ensure they perform as intended, including monitoring for bias and human accountability, especially after deployment. This broadly extends from DTAC's Section D1.7 consideration: iterate and improve frequently. For instance:

- Examples of questions from Canada Health Infoway for bias and nondiscrimination assessment during deployment and monitoring: "Define triggers that will automatically alert those responsible for oversight and monitoring of the AI system should the AI system begin to behave unexpectedly; if automatic triggers are not possible, define how often the AI system should undergo re-validation to ensure it remains free from bias and robust; document acceptable use criteria for the AI system to ensure the system is only deployed in an appropriate context; consider algorithmic auditing by third parties to ensure that your AI system remains free from bias; if indicators of unwanted bias are found in the system, immediately retrain, or re-develop the system."[28]
- An example of a question from Canada Health Infoway for AI vendor assessment: "What human oversight mechanisms does the vendor have in place (e.g., are there measures in place that would enable a human to intervene in, override, or reverse system outputs effectively?")[28]

Limitations

For phase 1 of this review, we mainly relied on publicly available information from government websites. For phase 2 of this review, we limited our eligibility criteria to Canada and selected health systems comparable to those in Canada (i.e., Australia, France, Germany, Netherlands, and the UK). We acknowledge that other countries may offer unique considerations for AI in health care, including countries with health systems different from those in Canada, and that we would have missed such considerations in this review. For both phases of the review, 1 reviewer conducted all data extraction and synthesis, and there is a possibility that relevant information or guidance was missed. We identified guidance with a lot of relevant content, and this report aimed to showcase the breadth of implementation considerations for AI-enabled medical devices. We acknowledge that implementation is complex and often involves many considerations (e.g., training, personnel, infrastructure, social and economic factors) that may fall outside the scope of this report. Instead, we have focused on the digital infrastructure elements of implementation and extensively referenced the identified guidance to allow our readers to refer to them in more detail, as needed. Thus, this review aims to serve as a foundational report that can be tailored for each AI topic. Moreover, this report is not a systematic

review and does not involve a critical appraisal of the literature. Thus, conclusions or recommendations about the value of or place in therapy for AI are outside this report's scope.

We leveraged patient engagement from the concurrent RapidAI review. Although we had intended to engage with 3 individuals (i.e., 2 patients and a clinician), we had limited response to our outreach during the RapidAI review and engaged 1 patient contributor in the end. 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 could have excluded some voices. It is worth considering whether people without reliable access to communication technologies would also face unique challenges or considerations around access to or use of DHTs that may not be captured in this review. Moreover, we cannot presume the opinions of 1 patient contributor engaged in the review of a specific AI technology reflects other patients' perspectives on all AI technologies in health care, and we cannot provide clinician perspectives without their participation. Therefore, while this implementation review covers all AI-enabled medical devices, to be applicable to reviews of specific AI technologies beyond RapidAI, it will likely require some tailoring.

Conclusions and Implications for Decision- or Policy-Making

Application of DTAC to the Health Care Context in Canada

We conducted an implementation review using a phased approach. In the first phase, we sought to determine whether and how DTAC applies to the health care context in Canada by exploring whether we have equivalent or similar measures, strategies, and policies to implement DHTs safely. Focusing on all assessment criteria of DTAC, we provided a high-level summary according to each of the 5 core areas of DTAC: clinical safety, data protection, technical security, interoperability, and usability and accessibility. Taken together, DTAC generally applies to the health context in Canada, with equivalent or similar guidance in place with some important caveats, which are summarized in this section.

For clinical safety, Canada appears to have similar medical device regulations, conformity declaration policies, and certain risk-management activities. However, there may be differences in what the UK requires compared with Canada, such as documentation. For example, the NHS provides templates for a "clinical risk management system," which we did not identify on Health Canada's website. Additional investigation is required to determine whether the forms, guidance, policies, and laws that Health Canada provides on its website provide sufficient coverage to fulfill DTAC's clinical safety criteria.

When considering the data-protection domain, the UK's personal and health information data-protection laws fall under the national government. Canada has many privacy laws in place, and the level of governance depends on the data type and jurisdiction. For example, personal information is governed at the federal level, and health information is governed at the provincial or territorial level. In addition, the Government of Canada has tabled Bill C-27, Digital Charter Implementation Act, 2022, which would update or replace current laws.[54,55] Although Canada's data-protection laws are nuanced, it appears that our nation has equivalent laws to address DTAC's considerations for data protection.

For technical security, it appears that DTAC has equivalents in the health care context in Canada. Specifically, we identified guidance and policies that correspond well with the DTAC assessment criteria for this domain (e.g., CyberSecure Canada for a cyber security certificate and guidance on MFA for organizations). We did not identify policies about load testing, but this may be because there was little context on this in the DTAC guidance, limiting our ability to look for equivalents.

Canada appears to have similar API guidance for interoperability considerations as the UK and complies with international standards, including ISO/IEEE11073 Personal Health Data Standards and OAuth 2.0. Compared to the UK, interoperability regarding electronic health records is more complex in Canada, given that, in Canada's current health care structure, electronic health records are not managed at the federal level.[77]

For usability and accessibility strategies, DTAC drew primarily from the UK's NHS Service Standard,[84] and Canada's published Digital Standards[85] that generally complement the points described in the NHS standard. There are some instances where the wording of the UK and Canada standards do not fully overlap, but this seemed to be covered by the information presented in the interoperability section of DTAC and by specific AI considerations from the identified guidance. DTAC describes the need for compliance with the Web Content Accessibility Guidelines, which are used in both the UK and Canada. Therefore, this section of DTAC also has equivalents in the health care context in Canada.

Ethics and Equity Considerations

Considering the overall purpose of DTAC, many ethics and equity considerations for DHTs are already embedded in the assessment criteria. The ethical considerations for DHTs are mainly found in the data protection domain, which relates to data privacy, management, and ownership concepts (i.e., Domains C2.3, C2.5, C2.5.1 of DTAC). WHO guidance also reflects similar ethical considerations.[17,89] The equity considerations for DHTs are mainly found under the usability and accessibility domain, which primarily relates to the involvement of relevant users (e.g., patients, caregivers, providers) in technology design, and whether their needs are incorporated into elements of technology design (i.e., Domains D1.1, D1.1.1, and D1.2 of DTAC). For example, DTAC asks to consider whether a diverse range of users were engaged, and, if not, who may not have been included and whether that has implications for how the subsequent questions may be answered (e.g., are there aspects of the user journey that may not have been considered?). Similar equity considerations are reflected in the Scottish Health Technology Group's HTA framework.[17]

WHO Guidance: Ethics and Governance of Artificial Intelligence for Health has 6 fundamental ethical principles for AI use for health that align with certain DTAC criteria but offer more specific considerations for AI technologies versus broader considerations for DHTs. WHO's fundamental ethical principle 1 (protect human autonomy) falls largely under DTAC's data-protection domain. This can be investigated by looking at human control over the technology, including full informed consent to all aspects of the technology and its application and elements of how personal data are used and managed to protect privacy. Data governance in Canada also requires considering and respecting First Nations, Inuit, and Métis data sovereignty principles (e.g., the First Nations principles of OCAP® [ownership, control, access, and possession],[90] Manitoba Métis principles of OCAS [ownership, control, access, and stewardship],[91] and Inuit Qaujimajatuqangit[92]),

which have implications for guiding the respectful governance of data collected with, from, or about Indigenous Peoples in Canada. Additionally, there are other data-governance frameworks and principles for equity-deserving groups that help establish data sovereignty and promote accountability, fairness, and transparency. For example, the Engagement, Governance, Access and Protection Framework[93] guides the collection, management, analysis, and use of race-based data from Black communities. WHO's principle 2 (promote human well-being, human safety, and the public interest) largely falls under DTAC's clinical safety domain and invites us to consider both impacts on individuals and groups or communities. Principles 3 to 6 (ensure transparency, explainability, and intelligibility; foster responsibility and accountability; ensure inclusiveness and equity; and promote responsible and sustainable AI) relate to DTAC's usability and accessibility domain. As emphasized in the Scottish Health Technology Group's HTA framework[17] and DTAC,[19] WHO's principle 4 reflects the importance of including patients and providers (and their values) in the technology design and considering who or what is accountable for the AI algorithm's decisions in its design and application.

Summary of Identified AI Considerations

In the second phase, our review aimed to identify implementation toolkits, guidance, and recommendations specific to AI and relevant to Canada to supplement DTAC, in case there are any important additional considerations for AI-enabled medical device use in Canada. We identified guidance from Canada and international guidance from WHO, Australia, France, and the UK. We found many additional considerations for AI-enabled medical device use in Canada, which we describe under the most relevant of the 5 core areas identified by DTAC. The AI consideration themes include monitoring, maintenance, and sustainability; AI data governance and data protection; multidisciplinary data-governance team throughout the product life cycle; technical infrastructure and integration; transparency, explainability, and intelligibility; inclusiveness, equity, and bias; responsibility and accountability; and user buy-in and organizational readiness. Each of these AI considerations serves its own purpose and is described in the findings section to explain how they can be considered for AI-enabled medical device implementation.

The most recurring AI considerations discussed across all DTAC core assessment criteria were monitoring, maintenance, and sustainability across the life cycle of the AI product, highlighting its importance. We believe this theme extends beyond DTAC's Section D1.7 consideration to "Iterate and improve frequently," as it has implications for all 5 core areas of DTAC, especially clinical safety. This coincides with the concept of iterative implementation, which is described in the literature as a way to overcome AI implementation barriers in primary health care in Canada.[49] A key AI data-governance consideration is protecting human autonomy (i.e., maintaining human control of health care systems and medical decisions) and allowing individuals to opt out of being included in the data used to train or run the AI system.[28] Transparency, explainability, and intelligibility are considerations specific to AI systems and require close consideration. AI models must be interpretable and explainable for all knowledge users to understand their decision-making process.[32] A known concern for AI-enabled medical devices is algorithmic bias, and the identified evidence highlighted some ways to conduct risk-management strategies for bias, such as ensuring AI data are representative of the intended population and are therefore accurate, complete, and diverse.[15,24] Another key AI consideration was

responsibility and accountability, which are needed if something goes wrong when applying AI technology (i.e., “human warranty”).[24]

Recent and Future Developments

Expectedly, guidance for AI-enabled medical devices will continue to evolve as more AI-enabled medical devices develop and more health care organizations consider and deploy these products. A recent perspective paper published in the *New England Journal of Medicine AI* discussed a new role some health systems have created, the Chief Health AI Officer, to provide specialized leadership.[94] This role could include developing a comprehensive AI strategy aligned with the organization’s goals, identifying high-impact use cases, ensuring effective implementation, and ensuring responsible AI application and compliance with evolving regulations.[94] Given the number of implementation considerations for AI-enabled medical devices, creating the role of Chief Health AI Officer may be a practical decision for health care organizations.

We also noted some emerging regulations in privacy law in Canada (e.g., Bill C-27, Digital Charter Implementation Act, 2022;[54,55] Connected Care for Canadians Act, 2024[78,79]), and there are also emerging regulations in AI, such as the right to contest decisions made by AI systems.[28] This idea was reinforced by the patient contributor engaged during the RapidAI review, who asked whether patients would be informed that AI is involved in making their diagnosis. These examples exemplify the ever-changing landscape and perhaps the need for future frameworks to be published as living documents to allow for adaptions as new guidance becomes available.

Outside of Canada, the European Union’s landmark *Artificial Intelligence Act*, the first comprehensive AI law, was enacted on August 1, 2024. This act lays the foundation for the regulation of AI in the European Union and classifies AI according to its risk: unacceptable risk, high risk, and minimal risk.[95] Also, the STANDING Together collaboration has developed international consensus-based recommendations to highlight and mitigate potential harms caused by bias in data and algorithms to ensure the future of AI-enabled health care is inclusive and equitable.[96] They provide recommendations intended to address inequitable AI health technology performance across the life cycle. Specifically, they list recommendations for documenting and using health datasets.[96] The extended scientific paper on these recommendations, including detailed explanatory text giving context and rationale for each item, is still in development.[96] These recent and future developments in Canada and internationally will likely help shape the future guidance for AI-enabled medical devices.

Implications for Clinical Practice and Policy-Making

In this review, we found that many of the safeguards and assessment criteria captured by DTAC in the UK setting have equivalent or similar measures, strategies, or policies in place in the health care context in Canada, with the exceptions previously noted. Through this work, we identified key implementation considerations for health care decision-makers to address for the safe and successful implementation of AI-enabled medical devices. While Canada has DTAC-equivalent or similar measures, strategies, or policies in place, we do not have a checklist like DTAC that decision-makers can use at the point of procurement or implementation of AI. This report helps identify the need for a checklist, like DTAC, for use in Canada.

This checklist could be an adaptation of DTAC and include additional implementation considerations for AI-enabled medical devices, such as those we identified in this review from organizations like the Vector Institute[32] and Canada Health Infoway.[28] Senior health care decision-makers can use this checklist to ensure that AI technologies of interest meet the minimum baseline standards set out by DTAC and inform the next steps for the safe and successful implementation of AI-enabled medical devices in Canada.

This implementation review for all AI-enabled medical devices is to be used alongside reviews of specific AI technologies, including the concurrent review of RapidAI, and will serve as a foundational report to be tailored for each AI topic and updated with the latest developments in the regulation and other aspects of the management of AI in the context of Canada.

Acknowledgements

This document was externally reviewed by the following content experts who have granted permission to be cited:

Eduardo Portela, MD
Neuroradiologist, The Ottawa Hospital
Assistant Professor, University of Ottawa

Jaron Chong, MD, MHI, FRCPC
Assistant Professor, Western University

Jay Shaw, PT, PhD
Research Director, Artificial Intelligence, Ethics & Health, Joint Centre for Bioethics, University of Toronto

We acknowledge with gratitude the contributions of a patient who preferred to remain anonymous.

Conflicts of Interest

Jaron Chong is the chair of the Artificial Intelligence Standing Committee with the Canadian Association of Radiologists and an ad hoc member of the Scientific Advisory Committee on Digital Health Technologies with Health Canada. No other conflicts of interest were identified.

References

1. Bélisle-Pipon J-C, Couture V, Roy M-C, Ganache I, Goetghebeur M, Cohen IG. What makes artificial intelligence exceptional in health technology assessment? *Front Artif Intell.* 2021;4:736697. PubMed
2. Artificial Intelligence Medical Devices (AIMD) Working Group. Machine Learning-enabled Medical Devices: Key Terms and Definitions. Washington (DC): International Medical Device Regulators Forum; 2022: https://www.imdrf.org/sites/default/files/2022-05/IMDRF%20AIMD%20WG%20Final%20Document%20N67.pdf. Accessed 2024 Jun 19.
3. Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices. In: Post TW, ed. *UpToDate*. Waltham (MA): UpToDate; 2024: https://www.uptodate.com/. Accessed 2024 Jun 19.
4. RapidAI. Menlo Park (CA): iSchemaView, Inc.; 2024: https://www.rapidai.com/. Accessed 2024 Jun 5.
5. Gerke S, Minssen T, Cohen G. *Ethical and legal challenges of artificial intelligence-driven healthcare. Artificial intelligence in healthcare*: Elsevier; 2020:295-336.
6. Mumtaz H, Riaz MH, Wajid H, et al. Current challenges and potential solutions to the use of digital health technologies in evidence generation: a narrative review. *Front Digit Health.* 2023;5:1203945. PubMed
7. Poon AI, Sung JJ. Opening the black box of AI□Medicine. *Journal of gastroenterology and hepatology.* 2021;36(3):581-584. PubMed
8. Department of Justice Canada. 2021, Food and Drugs Act: RSC 1985, c F-27. Ottawa (ON): Government of Canada; 2021: https://laws-lois.justice.gc.ca/eng/acts/f-27/. Accessed 2024 Jun 16.
9. Department of Justice Canada. 2021, Medical Devices Regulations: SOR/98-282. Ottawa (ON): Government of Canada; 2021: https://laws-lois.justice.gc.ca/eng/regulations/sor-98-282/. Accessed 2024 Jun 16.
10. Health Canada. Guidance Document: Software as a Medical Device (SaMD): Definition and Classification. Ottawa (ON): Government of Canada; 2019: https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/application-information/guidance-documents/software-medical-device-guidance-document.html. Accessed 2024 Jun 16.
11. Software as a Medical Device (SaMD). Silver Spring (MD): U.S. Food & Drug Administration; 2018: https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd. Accessed 2024 Jun 20.
12. Health Canada. Guidance Document: Software as a Medical Device (SaMD): Classification Examples. Ottawa (ON): Government of Canada; 2019: https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/application-information/guidance-documents/software-medical-device-guidance/examples.html. Accessed 2024 Jun 16.
13. Unsworth H, Wolfram V, Dillon B, et al. Building an evidence standards framework for artificial intelligence-enabled digital health technologies. *The Lancet Digital Health.* 2022;4(4):e216-e217. PubMed
14. Health Canada. Transparency for machine learning-enabled medical devices: Guiding principles. Ottawa (ON): Government of Canada; 2024: https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/transparency-machine-learning-guiding-principles.html. Accessed 2024 Jul 24.
15. Health Canada. Good Machine Learning Practice for Medical Device Development: Guiding Principles. Ottawa (ON): Government of Canada; 2021: https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/good-machine-learning-practice-medical-device-development.html. Accessed 2024 Jun 19.
16. International Health Technology Assessment Collaboration Expands. Ottawa (ON): CADTH; 2023: https://www.cadth.ca/news/international-health-technology-assessment-collaboration-expands. Accessed 2024 Jun 19.
17. Health Technology Assessment in Scotland: What evidence do we need? Edinburgh (GB): Healthcare Improvement Scotland; 2023: https://shtg.scot/media/2435/2023104-evidence-framework-v14.pdf. Accessed 2024 Jun 5.
18. 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 Jun 19.
19. 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 Jun 19.

20. Grey Matters: A Practical Tool For Searching Health-Related Grey Literature. Ottawa (ON): CADTH; 2022: https://greymatters.cadth.ca/. Accessed 2024 Jul 24.
21. Evidence standards framework (ESF) for digital health technologies. *Corporate document [ECD7].* London (GB): National Institute for Health and Care Excellence; 2022: https://www.nice.org.uk/corporate/ecd7. Accessed 2024 Jun 18.
22. Public and private health care: How do other countries do it? Ottawa (ON): Canadian Medical Association https://www.cma.ca/our-focus/public-and-private-health-care/how-do-other-countries-do-it. Accessed 2024 May 10.
23. Recommendation on the Ethics of Artificial Intelligence. Paris (FR): United Nations Educational, Scientific and Cultural Organization; 2021: https://unesdoc.unesco.org/ark:/48223/pf0000381137. Accessed 2024 Jun 5.
24. Ethics and governance of artificial intelligence for health. Geneva (CHE): World Health Organization; 2021.
25. EUnetHTA Joint Action 2. HTA Core Model, Version 3.0 *Work package 8*. Diemen (NLD): European Network for Health Technology Assessment; 2016: https://www.eunethta.eu/wp-content/uploads/2018/01/HTACoreModel3.0.pdf. Accessed 2024 Jun 5.
26. 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. PubMed
27. 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. PubMed
28. Toolkit for Implementers of Artificial Intelligence in Health Care. 2nd ed. Toronto (ON): Canada Health Infoway; 2023: https://www.infoway-inforoute.ca/en/component/edocman/resources/artificial-intelligence/3998-toolkit-for-implementers-of-artificial-intelligence-in-health-care?Itemid=101. Accessed 2024 Jun 17.
29. Digital Health Solutions Privacy & Security Guideline. Toronto (ON): Canada Health Infoway; 2023: https://www.infoway-inforoute.ca/en/component/edocman/resources/reports/privacy/6475-digital-health-solutions-privacy-security-guideline?Itemid=103#:~:text=This%20Guideline%20sets%20out%20technical,and%20secure%20information%20handling%20processes. Accessed 2024 Jun 17.
30. Health Canada. Predetermined change control plans for machine learning-enabled medical devices: Guiding principles. Ottawa (ON): Government of Canada; 2024: https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/good-machine-learning-practice-medical-device-development/predetermined-change-control-plans-machine-learning-enabled-medical-devices.html. Accessed 2024 Aug 22.
31. Health Canada. Draft guidance document: Pre-market guidance for machine learning-enabled medical devices. Ottawa (ON): Government of Canada; 2023: https://www.canada.ca/content/dam/hc-sc/documents/services/drugs-health-products/medical-devices/application-information/guidance-documents/pre-market-guidance-machine-learning-enabled-medical-devices/pre-market-guidance-machine-learning-enabled-medical-devices.pdf. Accessed 2024 Jun 18.
32. Dhalla A, Akinli Kocak S, Wan F, Da Silva J, Jain T. Vector Institute Health AI Implementation Toolkit. Toronto (ON): Vector Institute; 2023: https://vectorinstitute.ai/health-ai-implementation-toolkit/. Accessed 2024 Jun 14.
33. Perrier E, Rifai M, Terzic A, Dubois C, Cohen JF. Knowledge, attitudes, and practices towards artificial intelligence among young pediatricians: A nationwide survey in France. *Front Pediatr.* 2022;10:1065957. PubMed
34. Best Practice Guideline: Clinical Practice in a Digital Health Environment. Toronto (ON): Registered Nurses' Association of Ontario; 2024: https://rnao.ca/bpg/guidelines/clinical-practice-digital-health-environment. Accessed 2024 Jun 17.
35. Canadian Association of Radiologists Artificial Intelligence Working Group. Canadian Association of Radiologists white paper on ethical and legal issues related to artificial intelligence in radiology. *CAR J.* 2019;70(2):107-118.
36. Generating evidence for artificial intelligence-based medical devices: a framework for training, validation and evaluation. Geneva (CHE): World Health Organization; 2021: https://iris.who.int/bitstream/handle/10665/349093/9789240038462-eng.pdf?sequence=1. Accessed 2024 Jun 18.
37. Medicines and Healthcare Products Regulatory Agency. Guidance: Software and AI as a Medical Device Change Programme. London (GB): Gov.UK; 2023: https://www.gov.uk/government/publications/software-and-ai-as-a-medical-device-change-programme. Accessed 2024 Jun 18.

38. Regulatory Horizons Council. The Regulation of Artificial Intelligence as a Medical Device. London (GB): Department for Science, Innovation and Technology; 2022: https://assets.publishing.service.gov.uk/media/6384bf98e90e0778a46ce99f/RHC_regulation_of_AI_as_a_Medical_Device_report.pdf. Accessed 2024 Jun 18.
39. Therapeutic Goods Administration. Artificial Intelligence (AI) and medical device software: Information for software manufacturers about how we regulate AI medical devices. Canberra (AUS): Commonwealth of Australia; 2024: https://www.tga.gov.au/how-we-regulate/manufacturing/manufacture-medical-device/manufacture-specific-types-medical-devices/software-based-medical-devices/artificial-intelligence-ai-and-medical-device-software Accessed 2024-06-18.
40. LPPR: Dossier submission to the Medical Device and Health Technology Evaluation Committee (CNEDiMTS). Saint-Denis (FR): Haute Autorité de Santé; 2020: https://www.has-sante.fr/upload/docs/application/pdf/2020-10/guide_dm_vf_english_publi.pdf. Accessed 2024 Jun 18.
41. Health Canada. Guidance Document: Pre-market Requirements for Medical Device Cybersecurity. Ottawa (ON): Government of Canada; 2019: https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/application-information/guidance-documents/cybersecurity/document.html. Accessed 2024 Jun 27.
42. Health Canada. Quality Systems ISO 13485. Ottawa (ON): Government of Canada; 2019: https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/quality-systems-13485.html. Accessed 2024 Aug 22.
43. Health Canada. Safe medical devices in Canada. Ottawa (ON): Government of Canada; 2022: https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/activities/fact-sheets/safe-medical-devices-fact-sheet.html. Accessed 2023 Jun 27.
44. Health Canada. Medical device application and report forms. Ottawa (ON): Government of Canada; 2024: https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/application-information/forms.html. Accessed 2024 Jul 2.
45. Health Canada. Incident reporting for medical devices: Guidance document. Ottawa (ON): Government of Canada; 2021: https://www.canada.ca/en/health-canada/services/drugs-health-products/reports-publications/medeffect-canada/incident-reporting-medical-devices-guidance-2021.html. Accessed 2024 Jul 1.
46. Health Canada. Mandatory Medical Device Problem Reporting Form for Industry. Ottawa (ON): Government of Canada; 2018: https://www.canada.ca/en/health-canada/services/drugs-health-products/medeffect-canada/adverse-reaction-reporting/mandatory-medical-device-problem-reporting-form-industry-adverse-reaction-reporting.html. Accessed 2024 Jul 2.
47. Report a medical device problem (for health care professionals). Ottawa (ON): Government of Canada; 2021: https://hpr-rps.hres.ca/side-effects-reporting-form.php?form=medical_devices#:~:text=Reporting%20by%20hospitals%20is%20required,being%20documented%20within%20the%20hospital. Accessed 2024 Jul 2.
48. Health Canada. Compliance and enforcement of medical devices - Forms, guidance, policies and laws. Ottawa (ON): Government of Canada; 2023: https://www.canada.ca/en/health-canada/services/drugs-health-products/compliance-enforcement/information-health-product/medical-devices.html. Accessed 2024 Jul 2.
49. Darcel K, Upshaw T, Craig-Neil A, et al. Implementing artificial intelligence in Canadian primary care: Barriers and strategies identified through a national deliberative dialogue. *PLoS One.* 2023;18(2):e0281733. PubMed
50. PIPEDA requirements in brief. Ottawa (ON): Office of the Privacy Commissioner of Canada; 2024: https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/pipeda_brief/. Accessed 2024 Jun 17.
51. Guidelines for processing personal data across borders. Ottawa (ON): Office of the Privacy Commissioner of Canada; 2009: https://www.priv.gc.ca/en/privacy-topics/airports-and-borders/gl_dab_090127/. Accessed 2024 Jun 17.
52. Provincial laws that may apply instead of PIPEDA. Ottawa (ON): Office of the Privacy Commissioner of Canada; 2020: https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/r_o_p/prov-pipeda/. Accessed 2024 Jun 17.
53. PIPEDA Self-Assessment Tool. Ottawa (ON): Office of the Privacy Commissioner of Canada; 2021: https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/pipeda-compliance-help/pipeda-compliance-and-training-tools/pipeda_sa_tool_200807/. Accessed 2024 Jun 17.

54. Department of Justice Canada. Bill C-27: An Act to enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act and to make consequential and related amendments to other Acts. Ottawa (ON): Government of Canada; 2022: https://www.justice.gc.ca/eng/csj-sjc/pl/charter-charte/c27_1.html. Accessed 2024 Jun 25.
55. Innovation, Science and Economic Development Canada. Bill C-27 summary: Digital Charter Implementation Act, 2022. Ottawa (ON): Government of Canada; 2022: https://ised-isde.canada.ca/site/innovation-better-canada/en/canadas-digital-charter/bill-summary-digital-charter-implementation-act-2020. Accessed 2024 Jun 25.
56. Submission of the Office of the Privacy Commissioner of Canada on Bill C-27, the Digital Charter Implementation Act, 2022. Ottawa (ON): Office of the Privacy Commissioner of Canada; 2023: https://www.priv.gc.ca/en/opc-actions-and-decisions/submissions-to-consultations/sub_indu_c27_2304/. Accessed 2024 Jun 25.
57. A Guide to B.C.'s Personal Information Protection Act for Businesses and Organizations. Victoria (BC): Office of the Information & Privacy Commissioner for British Columbia; 2015: https://www.oipc.bc.ca/documents/guidance-documents/1371. Accessed 2024 Jun 17.
58. PIPA: 10 Steps to Implement PIPA. Edmonton (AB): Office of the Information and Privacy Commissioner of Alberta; 2010: https://oipc.ab.ca/resource/pipa-implementation/. Accessed 2024 Jun 17.
59. National Assembly of Québec. Bill 64, chapter 25: An Act to modernize legislative provisions as regards the protection of personal information. *First Session. Forty-Second Legislature*. Québec (QC): Québec Official Publisher; 2021: https://www.publicationsduquebec.gouv.qc.ca/fileadmin/Fichiers_client/lois_et_reglements/LoisAnnuelles/en/2021/2021C25A.PDF. Accessed 2024 Jun 17.
60. Innovation, Science and Economic Development Canada. CyberSecure Canada. Ottawa (ON): Government of Canada; 2024: https://ised-isde.canada.ca/site/cybersecure-canada/en. Accessed 2024 Jun 28.
61. Baseline Cyber Security Controls for Small and Medium Organizations. Ottawa (ON): Digital Governance Council; 2021: https://dgc-cgn.org/standards/find-a-standard/standards-in-cybersecurity/cybersecurity-smes/. Accessed 2024 Jul 2.
62. Get Cyber Safe Guide for Small and Medium Businesses. Ottawa (ON): Government of Canada; 2022: https://www.getcybersafe.gc.ca/en/resources/get-cyber-safe-guide-small-and-medium-businesses. Accessed 2024 Jun 28.
63. Open Worldwide Application Security Project. OWASP Top Ten. 2021; https://owasp.org/www-project-top-ten/.
64. Penetration Testing Methodologies. *The OWASP Testing Framework*: Open Worldwide Application Security Project: https://owasp.org/www-project-web-security-testing-guide/latest/3-The_OWASP_Testing_Framework/1-Penetration_Testing_Methodologies. Accessed 2024 Jul 2.
65. Canadian Centre for Cyber Security. Top measures to enhance cyber security for small and medium organizations (ITSAP.10.035). Ottawa (ON): Government of Canada; 2024: https://www.cyber.gc.ca/en/guidance/top-measures-enhance-cyber-security-small-and-medium-organizations-itsap10035. Accessed 2024 Jul 2.
66. Canadian Centre for Cyber Security. Security considerations for your website (ITSM.60.005). Ottawa (ON): Government of Canada; 2021: https://www.cyber.gc.ca/en/guidance/security-considerations-your-website-itsm60005#wb-tphp. Accessed 2024 Jul 2.
67. Secure development and deployment guidance. London (GB): National Cyber Security Centre: https://www.ncsc.gov.uk/collection/developers-collection/principles/produce-clean-maintainable-code. Accessed 2024 Jul 2.
68. Canadian Centre for Cyber Security. Secure your accounts and devices with multi-factor authentication (ITSAP.30.030). Ottawa (ON): Government of Canada; 2024: https://www.cyber.gc.ca/en/guidance/secure-your-accounts-and-devices-multi-factor-authentication-itsap30030. Accessed 2024 Jul 2.
69. Canadian Centre for Cyber Security. Network security logging and monitoring - ITSAP.80.085. Ottawa (ON): Government of Canada; 2022: https://www.cyber.gc.ca/en/guidance/network-security-logging-monitoring-itsap80085. Accessed 2024 Jul 2.
70. Digital Health Transformation. Chicago (IL): Healthcare Information and Management Systems Society, Inc. (HIMSS); 2024: https://www.himss.org/what-we-do-solutions/maturity-models. Accessed 2024 Aug 26.

71. Treasury Board of Canada Secretariat. Government of Canada Standards on APIs. Ottawa (ON): Government of Canada; 2020: https://www.canada.ca/en/government/system/digital-government/modern-emerging-technologies/government-canada-standards-apis.html. Accessed 2024 Jul 3.
72. Guidance for Application Programming Interface (API) Gateways. Ottawa (ON): Government of Canada; 2021: https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/enabling-interoperability/api-guidance/gateways.html. Accessed 2024 Jul 3.
73. API Guidance. Ottawa (ON): Government of Canada; 2021 https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/enabling-interoperability/api-guidance.html. Accessed 2024 Jul 3.
74. API Security Best Practices Primer. Ottawa (ON): Government of Canada; 2021: https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/enabling-interoperability/api-guidance/security.html. Accessed 2024 Jul 3.
75. OAuth 2.0. https://oauth.net/2/. Accessed 2024 Jul 3.
76. Ontario Health. Digital Health Information Exchange Standard. Toronto (ON): Government of Ontario; 2022: https://www.ontariohealth.ca/system-planning/digital-standards/digital-health-information-exchange. Accessed 2024 Jul 3.
77. Competition Bureau Canada. Unlocking the power of health data. Ottawa (ON): Government of Canada; 2022: https://competition-bureau.canada.ca/unlocking-power-health-data. Accessed 2024 Jul 3.
78. Siddle D, Caldwell M. Access, Interoperability, and Innovation: An Introduction to Canada's Bill C-72, the Connected Care For Canadians Act. Vancouver (BC): McCarthy Tetrault LLP; 2024: https://www.mccarthy.ca/en/insights/blogs/techlex/access-interoperability-and-innovation-introduction-canadas-bill-c-72-connected-care-canadians-act. Accessed 2024 Jul 24.
79. House of Commons Canada. Bill C-72. *First Session, Forty-fourth Parliament, 70-71 Elizabeth II – 1-2 Charles III, 2021-2022-2023-2024*. Ottawa (ON): Parliament of Canada: https://www.parl.ca/DocumentViewer/en/44-1/bill/C-72/first-reading. Accessed 2024 Jul 24.
80. IEEE Standard - Health informatics -- Device interoperability -- Part 10206: Personal health device communication -- Abstract Content Information Model. Piscataway (NJ): IEEE Standards Association; 2022: https://standards.ieee.org/ieee/11073-10206/10311/#:~:text=Within%20the%20context%20of%20the%20ISO%2FIEEE%2011073%20personal%20health,between%20device%20types%20and%20vendors. Accessed 2024 Jul 3.
81. Enabling connection and communication across the health system. Toronto (ON): Canada Health Infoway; 2024: https://www.infoway-inforoute.ca/en/digital-health-initiatives/interoperability. Accessed 2024 Jul 3.
82. Shared Pan-Canadian Interoperability Roadmap. Toronto (ON): Canada Health Infoway; 2023: https://www.infoway-inforoute.ca/en/component/edocman/6444-connecting-you-to-modern-health-care-shared-pan-canadian-interoperability-roadmap/view-document?Itemid=103. Accessed 2024 Aug 22.
83. Privacy as an Enabler: Sharing Personal Health Information for Interoperability Primer Toronto (ON): Canada Health Infoway; 2022: https://www.infoway-inforoute.ca/en/component/edocman/resources/reports/privacy/6387-privacy-as-an-enabler-sharing-personal-health-information-for-interoperability-primer?Itemid=103. Accessed 2024 Aug 22.
84. NHS service standard. London (GB): NHS England; 2022: https://service-manual.nhs.uk/standards-and-technology/service-standard. Accessed 2024 Jul 22.
85. Digital standards playbook. Ottawa (ON): Government of Canada; 2021: https://www.canada.ca/en/government/system/digital-government/government-canada-digital-standards.html. Accessed 2024 Jul 3.
86. Treasury Board of Canada Secretariat. Standard on Web Accessibility. Ottawa (ON): Government of Canada; 2011: https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=23601. Accessed 2024 Jul 3.
87. Legal Guidance for HTA Bodies. Canadian Law & HTA Working Group; 2022: https://drive.google.com/file/d/1nxsf4HRP9xYZaDi6oXmwB5YyberIbpHv/view. Accessed 2024 Jun 21.
88. Willis BH. Spectrum bias—why clinicians need to be cautious when applying diagnostic test studies. *Family Practice.* 2008;25(5):390-396. PubMed
89. Zhou K, Gattinger G. The Evolving Regulatory Paradigm of AI in MedTech: A Review of Perspectives and Where We Are Today. *Ther Innov Regul Sci.* 2024;58(3):456-464. PubMed

90. The First Nations Principles of OCAP®. Akwesasne (ON): First Nations Information Governance Centre: https://fnigc.ca/ocap-training/. Accessed 2024 Jun 15.

91. Framework for Research Engagement with First Nation, Metis, and Inuit Peoples. Winnipeg (MB): University of Manitoba; 2021: https://umanitoba.ca/health-sciences/sites/health-sciences/files/2021-01/framework-research-report-fnmip.pdf. Accessed 2024 Jun 15.

92. Inuit Qaujimajatuqangit: The Role of Indigenous Knowledge in Supporting Wellness in Inuit Communities in Nunavut Prince George (BC): National Collaborating Centre for Aboriginal Health; 2010: https://www.ccnsa-nccah.ca/docs/health/FS-InuitQaujimajatuqangitWellnessNunavut-Tagalik-EN.pdf. Accessed 2024 Jun 15.

93. Engagement, Governance, Access, and Protection (EGAP): A Data Governance Framework for Health Data Collected from Black Communities in Ontario. Toronto (ON): Black Health Equity; 2021: https://blackhealthequity.ca/wp-content/uploads/2021/03/Report_EGAP_framework.pdf. Accessed 2024 Aug 22.

94. Beecy AN, Longhurst CA, Singh K, Wachter RM, Murray SG. The Chief Health AI Officer—An Emerging Role for an Emerging Technology. *NEJM AI.* 2024;1(7).

95. The EU Artificial Intelligence Act. Future of Life Institute 2024: https://artificialintelligenceact.eu/. Accessed 2024 Aug 22.

96. STANDING Together Working Group. Recommendations for Diversity, Inclusivity, and Generalisability in Artificial Intelligence Health Technologies and Health Datasets. Birmingham (GB): Birmingham AI & Digital Health Research & Policy Group; 2023: https://www.datadiversity.org/recommendations. Accessed 2024 Jun 26.

Canada's Drug Agency
L'Agence des médicaments du Canada
Drugs. Health Technologies and Systems. Médicaments, technologies de la santé et systèmes.

ISSN: 2563-6596

Canada's Drug Agency (CDA-AMC) is a pan-Canadian health organization. Created and funded by Canada's federal, provincial, and territorial governments, we're responsible for driving better coordination, alignment, and public value within Canada's drug and health technology landscape. We provide Canada's health system leaders with independent evidence and advice so they can make informed drug, health technology, and health system decisions, and we collaborate with national and international partners to enhance our collective impact.

Questions or requests for information about this report can be directed to Requests@CDA-AMC.ca.