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

MT9

Assessment of Artificial Intelligence–Enabled Health Technologies: Our Initial Experience at Canada's Drug Agency

Calvin Young, Chantelle Lachance, Joanne Kim, Allison Gates, Renata Axler, Ana Komparic, Angie Hamson, Bernice Tsoi, Caitlyn Ford, Chris Kamel; Canada's Drug Agency, Ottawa, Ontario

Background

- At Canada's Drug Agency (CDA-AMC), our initial experience in assessing artificial intelligence (AI)–enabled health technologies focused on stroke, a leading cause of death and disability in Canada.[1]
- RapidAI is an AI-enabled software platform that facilitates the viewing, processing, and analysis of CT images to aid clinicians in assessing patients with suspected stroke. Although the platform is already in use in some settings in Canada, its application in stroke detection is limited.
- Understanding its potential benefits and harms is key for clarifying its role in stroke detection.

Objectives

- To evaluate 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) and to develop evidence-based recommendations for its implementation in Canada.

Methods

- We conducted an evidence review on the effectiveness, accuracy, and cost-effectiveness of RapidAI for stroke detection.
- Ethics and equity considerations were integrated throughout, and were informed by the literature and patient, clinician, and other expert input.
- The Health Technology Expert Review Panel (HTERP), an advisory body to CDA-AMC, reviewed the evidence and developed recommendations on the appropriate use of RapidAI for stroke detection. HTERP used the new CDA-AMC deliberative framework to guide its recommendations, considering the following domains: unmet clinical need, clinical value, economic considerations, impacts to health systems, and distinct social and ethical considerations.

Results

- We found 2 cohort studies and 11 diagnostic accuracy studies that assessed the effectiveness and accuracy of RapidAI for detecting stroke. Table 1 presents the findings for selected outcomes based on evidence from the 2 cohort studies.
- The sensitivity and specificity of Rapid ICH with clinician interpretation for detecting ICH using noncontrast CT in people with suspected acute stroke were 92% (95% CI, 78% to 98%) and 100% (95% CI, 98% to 100%), respectively (1 study; 307 patients; low certainty).
- As a standalone intervention, the sensitivity of RapidAI for detecting LVO ranged from 62% to 96%, while estimates of specificity ranged from 65% to 98% (10 studies; moderate, low, very low, or unclear certainty).
- We found no relevant evidence regarding the cost-effectiveness of RapidAI for detecting ICH or LVO in people with suspected stroke.
- Ethical considerations related to autonomy, privacy, safety, transparency, explainability of machine-learning models, and equity have implications across the technology life cycle when using RapidAI for detecting stroke and could influence its acceptability by clinicians, patients, and health care institutions.
- Using the available evidence, HTERP deliberated on and answered the question, "Should RapidAI be implemented to detect stroke in Canada, and how?"

Table 1: Summary of Findings for Clinician Interpretation of CTA Imaging With RapidAI Versus Clinician Interpretation of CTA Imaging Without RapidAI

Outcome and follow-up	Intervention	Participants (studies), N	Absolute effects	Certainty
Time to intervention				
Radiology-report turnaround (minutes), mean (SD)	Rapid LVO (Rapid v4.9)	≤ 760 (1 NRS)	• Without Rapid LVO: 30.6 (29.9) • With Rapid LVO: 22.0 (35.1) • Difference: −8.6 (32.8)	Low *(due to risk of bias)*
CTA to groin puncture (minutes), mean (SD)	Rapid CTA (version NR)	146 (1 NRS)	• Without Rapid CTA: 92 (NR) • With Rapid CTA: 68 (NR) • Difference: −24 (NE)	Very low *(due to risk of bias and imprecision)*
Functional status				
Proportion of patients with significant morbidity or mortality (defined as an mRS score ≥ 5) at discharge (95% CI)	Rapid LVO (Rapid v4.9)	105 (1 NRS)	• 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)	Very low *(due to risk of bias and imprecision)*
Proportion of patients considered to be functionally independent (defined as an mRS score ≤ 2) at 90 days (95% CI)	Rapid CTA (version NR)	141 (1 NRS)	• 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)	Very low *(due to risk of bias and imprecision)*

CI = confidence interval; CTA = computed tomography angiography; LVO = large-vessel occlusion; mRS = modified Rankin Scale; NE = not estimable; NR = not reported; NRS = nonrandomized study; SD = standard deviation.

HTERP's Recommendation

In locations where RapidAI has already been implemented for use in detecting suspected LVO and ICH, HTERP recommends:

- RapidAI is used only as indicated, alongside clinician interpretation of CT imaging, to reduce the risk of incorrect results
- the generation of evidence to evaluate its value in health care systems, including its use in less-resourced centres with limited access to stroke care specialists.

In locations considering the implementation of RapidAI for use in detecting suspected LVO and ICH, given the uncertainty and gaps in the evidence regarding clinical, economic, and equity value of RapidAI, HTERP cannot provide recommendations for or against its implementation.

Discussion

The evidence review and deliberations identified several limitations in the clinical literature on RapidAI for stroke detection:

- The reviewed studies did not describe the methods used to develop RapidAI's machine-learning models, preventing an assessment of the training dataset's representativeness and diversity.
- While RapidAI may offer greater utility in rural settings with limited access to stroke care specialists, existing evidence comes from comprehensive stroke centres with high volumes of stroke cases and the resources needed to perform timely imaging studies. As a result, the findings from the included studies may not be generalizable to less-resourced settings.
- Most clinical studies evaluating the diagnostic accuracy of RapidAI assessed it as a standalone tool, despite its intended use as a supportive intervention. Consequently, much of the evidence summarized in the review is indirect and does not provide a clear indication of how the tool performs in clinical practice.

Similar limitations may apply to other AI-enabled diagnostic technologies.

To address these gaps in future assessments, evaluators could incorporate nontraditional sources of information (e.g., preclinical studies, product information sheets) to better understand the processes used for developing, training, and validating the machine-learning models. Early engagement and collaboration among software developers, evidence generators and evaluators, patients, clinicians, technical experts, health care administrators, and other AI ecosystem partners could help align information needs within the health technology assessment landscape.

Conclusions

- While RapidAI has the potential to improve time to diagnosis, its impact on many outcomes, including those that are important to patients, is uncertain.
- The cost-effectiveness of RapidAI for stroke detection is currently unknown.
- Decision-makers may wish to reflect on the ethical and equity considerations that arise during the deployment of AI-enabled technologies (e.g., autonomy, privacy, transparency, and access).
- Lessons learned will support future assessments of AI-enabled health technologies by CDA-AMC.

Reference

1. 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.

Read the full report here: