Artificial Intelligence-Based Clinical Decision Support in the Emergency Department: Bridging Development to Implementation

Hashim Khaliq Kareemi, MD

Thesis submitted to the University of Ottawa in partial fulfillment of the requirements for the Master's of Science in Epidemiology.

School of Epidemiology and Public Health, Faculty of Medicine
University of Ottawa

Author: Dr. Hashim Khaliq Kareemi[1,2]
Thesis Supervisor: Dr. Christian Vaillancourt[1,3,4]
Thesis Co-Supervisor: Dr. Krishan Yadav[1,3.4]
Thesis Advisory Committee: Dr. Christian Vaillancourt,[1,3,4] Dr. Krishan Yadav,[1,3,4] Dr. Wojtek Michalowski[5]

Affiliations:

1. School of Epidemiology and Public Health, University of Ottawa, Ottawa, ON
2. Department of Emergency Medicine, University of British Columbia, Vancouver, BC
3. Department of Emergency Medicine, University of Ottawa, Ottawa, ON
4. Acute Care Research Program, Ottawa Hospital Research Institute, Ottawa, ON
5. Telfer School of Management, University of Ottawa, Ottawa, ON

Table of Contents

Preface

Thesis Abstract

Background: Artificial intelligence (AI)-based clinical decision support (CDS) tools are desired by physicians to augment high-stakes decision-making in the emergency department (ED).

Objective: This thesis evaluates the current field of AI-CDS in the ED and explores barriers and facilitators to their development and implementation.

Methods: We conducted a scoping review of AI-CDS tools for individual ED patient care and categorized them by phase of development. We conducted interviews with expert researchers to identify barriers and potential facilitators for successful implementation.

Results: Despite a rapidly growing number of publications, only 3.5% of AI-CDS tools have been tested or implemented in a live clinical setting. Expert researchers identified challenges regarding data infrastructure, team capacity, defining the clinical problem, regulatory approval, legal and liability concerns, time, and cost.

Conclusion: To bridge the gap between development and implementation, researchers must incorporate implementation science principles in the earliest stages of AI-CDS tool development.

Acknowledgments

I am indebted to several individuals who, together, provided me with the guidance, support, and drive needed to complete this work.

To Christian Vaillancourt, thank you for helping me harness my thoughts into meaningful pursuits, for keeping me intellectually curious, and for showing me the value of honest work done for honest reasons.

To Krishan Yadav, thank you for your unwavering support in research, clinical work, and beyond, for always leading by example, and for encouraging me to dream bigger than I would allow myself.

To Wojtek Michalowski, thank you for your wisdom, patience, and guidance as I explore the world of artificial intelligence and digital technologies.

Thank you to my study co-authors for your enthusiasm and willingness to engage in this work; to my fellow Canadian Association of Emergency Physicians AI Special Interest Group members for helping create an outlet for our shared passion, of which this work is just one representation; to my emergency medicine colleagues in Ottawa for inspiring me to pursue academia; and to my colleagues in Vancouver for nurturing my desire to continue that pursuit.

To Ammi, Abbu, Qasim, and Kaamil, thank you for your love and support from day one.

To Elizabeth, thank you for being my partner, my proofreader, my idea-hearer, and my constant through this journey.

Author Contributions and Required Ethics Approvals

Manuscript 1: Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Scoping Review Protocol

Author Contributions: HK, KY, WM, and CV conceptualized and planned the study. HK devised and refined the search strategy, performed trial searches, reviewed the search with health sciences librarians, and wrote the manuscript. All authors edited, reviewed, and approved the final manuscript.

Required Approval: None.

Manuscript 2: Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Scoping Review

Author Contributions: HK, KY, WM, and CV conceptualized and planned the study. HK, CP, NB, AM, HL, and GG performed the title and abstract screening. HK, SM, LG, and MB performed the full text screening. HK created the tables and figures and wrote the manuscript. All authors edited, reviewed, and approved the final manuscript.

Required Approval: None.

Manuscript 3: Barriers and Facilitators to Developing Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Qualitative Study

Author Contributions: HK, KY, WM, and CV conceptualized and planned the study. HK conducted the interviews. HK and AC performed coding and analysis of the interview transcripts. HK created the tables and figures and wrote the manuscript. All authors edited, reviewed, and approved the final manuscript.

Required Approval: Approved by Ottawa Health Science Network-Research Ethics Board on May 29, 2024 (protocol ID: 20240338-01H).

Glossary

AI: Artificial Intelligence

AI-CDS: Artificial Intelligence-Based Clinical Decision Support

CDS: Clinical Decision Support

CINAHL: Cumulative Index to Nursing and Allied Health Literature

CONSORT-AI: Consolidated Standards of Reporting Trials – Artificial Intelligence

COREQ: Consolidated Criteria for Reporting Qualitative Research

DECIDE-AI: Developmental and Exploratory Clinical Investigations of Decision Support Systems Driven by Artificial Intelligence

ED: Emergency Department

EHR: Electronic Health Record

ICU: Intensive Care Unit

ML: Machine Learning

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

SALIENT: Staged Clinical Artificial Intelligence Implementation

SPIRIT-AI: Standard Protocol Items: Recommendations for Interventional Trial – Artificial Intelligence

SRQR: Standards for Reporting Qualitative Research

STARD-AI: Standards for Reporting of Diagnostic Accuracy Studies – Artificial Intelligence

TRIPOD+AI: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis Or Diagnosis + Artificial Intelligence

Chapter 1. Introduction

This chapter outlines the thesis objectives and provides an overview of the individual chapters.

1.1. Thesis Objectives

The primary objective of this thesis is to understand the current state of artificial intelligence (AI)-based clinical decision support (CDS) tools used in the emergency department (ED) and identify strategies to facilitate adoption of these tools into emergency clinical practice.

Specifically, we sought to:

A. Describe the available literature regarding AI-CDS tools that have been developed for individual patient care in the ED and the phases of development these tools have achieved.
B. Construct a theory-driven explanation for the limited clinical translation of AI-CDS tools in the ED.
C. Describe the barriers to development and implementation of AI-CDS tools in the ED and identify facilitators that may help surmount them.

1.2. Overview of Thesis Chapters

1.2.1. Chapter 2: Background

This chapter introduces the concepts of AI and machine learning (ML) and provides context for their application to medicine and healthcare, particularly in the realm of CDS. It then summarizes contemporary challenges faced by emergency healthcare systems and the mounting crisis driven by increasing patient volumes and resource scarcity. It concludes with an outline of how AI-CDS tools have been proposed as desired solutions to several of those challenges.

1.2.2. Chapter 3: Scoping Review

This chapter summarizes and describes the existing literature for AI-CDS tools for use in the ED. We conducted a scoping review of the literature to identify the types of tools that have been developed and the phases of development they have achieved. We include the scoping review protocol manuscript, "Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Scoping Review Protocol" that was published online in *Open Science Framework*, and the scoping review manuscript, "Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Scoping Review" that was published in *Academic Emergency Medicine*. Reproduced with permission according to Wiley Article Sharing Policy.

Publication: Kareemi H, Yadav K, Michalowski W, et al. Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Scoping Review Protocol [Internet]. OSF; 2025. Available from: osf.io/fdz3y

Publication: Kareemi H, Yadav K, Price C, et al. Artificial intelligence-based clinical decision support in the emergency department: A scoping review. *Acad Emerg Med*. 2025;32(4):386-395. doi:10.1111/acem.15099

1.2.3. Chapter 4: Qualitative Analysis of Expert Interviews

This chapter utilizes a qualitative methodological approach to describe the experiences of researchers who have developed and implemented AI-CDS tools, summarizes the challenges they faced and identifies solutions that can be applied early in the development of these tools. We include the manuscript, "Barriers and Facilitators to Developing Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Qualitative Study" that is under review for publication in *Canadian Journal of Emergency Medicine*.

1.2.4. Chapter 5: Discussion

This chapter summarizes the thesis and outlines its key findings, relevance to the current body of knowledge, strengths and limitations, and implications for future research and practice.

Chapter 2. Background

This chapter describes challenges with high-stakes, time-sensitive decision-making in emergency medicine, complicated by the crisis facing emergency departments (ED) being driven by increasing patient volumes and resource scarcity. It introduces the concepts of artificial intelligence (AI) and machine learning (ML) and outlines how they have been incorporated into healthcare solutions. It concludes by defining clinical decision support (CDS) and describing how AI is being applied to this field to address the challenges facing emergency medicine.

2.1. Emergency Healthcare: From Challenge to Crisis

Emergency medicine is characterized by diagnostic uncertainty, time-sensitive care, and the threat of adverse outcomes. Emergency physicians must often act decisively with incomplete data, navigating complex patient presentations that evolve rapidly and unpredictably. These high-stakes decisions are made more difficult by a worsening crisis in emergency healthcare: EDs globally are facing surging patient volumes, increasing case complexity, and critical shortages of staff and resources.[1-3] In Canada, the number of unscheduled ED visits rose to 15.5 million in 2023-2024, larger than any prior year, with an increasing proportion of patients having higher acuity levels.[4] This confluence of pressures has led to overcrowding, prolonged wait times, and frequent delays in care, which are associated with increased mortality and worsened patient outcomes.[2, 5] The

problem is not simply one of capacity, but of cognitive and operational overload. Emergency clinicians are being asked to do more, faster, and with fewer supports, raising the risk of diagnostic errors, clinical burnout, and ED closures.[2, 6, 7] There is an urgent need for tools that can mitigate this burden by augmenting decision-making, reducing uncertainty, and streamlining workflows. The introduction of CDS systems, particularly those enhanced by AI, represents a potential solution to these systemic challenges.

2.2. Artificial Intelligence and Machine Learning in Healthcare

AI refers to the development of computer systems capable of performing tasks that usually require human intelligence, such as reasoning, pattern recognition, and decision-making.[8, 9] Within the domain of AI, ML describes a subset of techniques that allow systems to learn from data, identify patterns, and improve their performance on specific tasks over time.[9, 10] The line between ML and traditional statistical approaches remains blurry. While some make the distinction that ML models do not require explicit programming to achieve these tasks,[9] in practice many still do. Perhaps a more accurate distinction is that ML models learn through examples whereas traditional statistical models are programmed using predefined rules.[10]

In healthcare, AI and specifically ML have demonstrated considerable promise across a range of applications. Notable successes include image-based diagnostics (e.g. radiology and dermatology),[11, 12] discovery of new pharmacologic agents,[13] and personalized treatment recommendations.[14] In the ED, ML models have been shown to outperform usual care and traditional statistical approaches in areas such as sepsis

prediction, emergency triage, and diagnostic imaging interpretation.[15] These advances have been supported by the increasing availability of large-scale electronic health record (EHR) data and improvements in computational power.

2.3. Applying Artificial Intelligence to Clinical Decision Support

Clinical decision support (CDS) includes tools and systems designed to augment healthcare decision-making by providing clinicians with evidence-based knowledge, patient-specific information, or predictive insights.[16, 17] CDS tools have traditionally included rule-based alerts, reminders, order sets, and diagnostic support. For example, a simple rule-based alert may involve a pop-up window on the EHR interface that warns a clinician that a patient has a penicillin allergy when antibiotics are being ordered. However, the advent of AI has broadened the scope and sophistication of CDS capabilities.

AI-enabled CDS tools can support decisions across the diagnostic, prognostic, and therapeutic spectrum. For example, ML-based algorithms can assist in identifying patients at risk of deterioration, predicting length of stay, or recommending appropriate imaging or treatment strategies.[15] Large language models such as ChatGPT are being increasingly studied for diagnostic and triage tasks in emergency departments.[18] These tools may hold particular promise in the ED, where time and data constraints are acute. They have a theoretical advantage over traditional CDS tools in that they can analyze massive datasets involving complex types of data and provide nuanced, personalized treatment recommendations that automatically incorporate information from the EHR. However,

there is limited evidence to suggest that AI-CDS tools currently improve clinician diagnostic performance when applied in live clinical settings.[19]

To facilitate the creation of safe, efficacious AI-CDS tools, structured development pathways have been proposed. Notably, the *Developmental and Exploratory Clinical Investigations of Decision Support Systems Driven by Artificial Intelligence* (DECIDE-AI) framework outlines a translational roadmap for AI-CDS tools,[20] emphasizing the importance of early-stage clinical evaluation, iterative user feedback, and real-world feasibility testing. It mirrors established pathways for pharmaceutical and surgical innovations, incorporating reporting guidelines for each stage of AI-CDS development: "offline" validation of models for prediction (*Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis Or Diagnosis + Artificial Intelligence*, TRIPOD+AI) and diagnostic test accuracy (*Standards for Reporting of Diagnostic Accuracy Studies – Artificial Intelligence*, STARD-AI);[21, 22] early "live" evaluations (DECIDE-AI);[20] production of trial protocols (*Standard Protocol Items: Recommendations for Interventional Trial – Artificial Intelligence*, SPIRIT-AI);[23] and reporting findings for large-scale effectiveness trials (*Consolidated Standards of Reporting Trials – Artificial Intelligence*, CONSORT-AI).[24]

2.4. Desire for AI-CDS Tools in the ED

The application of AI-CDS in the ED is not merely a technical possibility, it is increasingly viewed as a clinical necessity. Recent research has explored the perspectives of emergency medicine stakeholders regarding the need for and utility of AI-CDS tools. A

2024 international Delphi study identified consensus among emergency clinicians for the application of AI-CDS to therapeutic (e.g. drug interaction) and diagnostic (e.g. sepsis identification) decisions.[25] Similarly, a 2023 survey found that Canadian emergency physicians prioritized triage and diagnostics for AI-CDS application, albeit lower than non-CDS tasks such as automated documentation and computer/EHR utilization.[26]

Despite this interest, major gaps remain. It is currently unclear how many AI-CDS tools have been developed specifically for ED use, whether they have been successfully integrated into real-time clinical workflows, and whether they have demonstrable impacts on outcomes such as diagnostic accuracy, clinician efficiency, or patient safety. Furthermore, it is unclear if any studies have addressed the challenges of user trust, explainability, and alert fatigue, all of which are critical to successful deployment and sustained use.[27, 28]

While the need and desire for AI-CDS tools in the ED are increasingly evident, significant questions remain regarding their real-world effectiveness, scalability, and adoption. Addressing these gaps is essential to realizing the potential of AI in emergency medicine.

References

1. Morley C, Unwin M, Peterson GM, Stankovich J, Kinsman L. Emergency department crowding: A systematic review of causes, consequences and solutions. PLoS One. 2018;13(8):e0203316.
2. Canadian Agency for Drugs and Technologies in Health. Emergency Department Overcrowding in Canada: CADTH Health Technology Review Recommendation [Internet]. Ottawa (ON): Canadian Agency for Drugs and Technologies in Health; 2023.
3. Kelen GD, Wolfe R, D'Onofrio G, Mills AM, Diercks D, Stern SA, et al. Emergency Department Crowding: The Canary in the Health Care System. Catalyst non-issue content. 2021;2(5).
4. Canadian Institute for Health Information. NACRS emergency department visits and lengths of stay. 2023.
5. Sun BC, Hsia RY, Weiss RE, Zingmond D, Liang LJ, Han W, et al. Effect of emergency department crowding on outcomes of admitted patients. Ann Emerg Med. 2013;61(6):605-11.e6.
6. Pines JM, Pollack CV, Jr., Diercks DB, Chang AM, Shofer FS, Hollander JE. The association between emergency department crowding and adverse cardiovascular outcomes in patients with chest pain. Acad Emerg Med. 2009;16(7):617-25.
7. Canadian Association of Emergency Physicians expresses concerns over summer closures of emergency departments across Canada [press release]. Ottawa (ON): Canadian Association of Emergency Physicians2023.

8. Russell SJ, Norvig P, Chang M-W. Artificial intelligence : a modern approach. Fourth Edition. ed. Harlow: Pearson; 2021.
9. Ramlakhan S, Saatchi R, Sabir L, Singh Y, Hughes R, Shobayo O, et al. Understanding and interpreting artificial intelligence, machine learning and deep learning in Emergency Medicine. Emerg Med J. 2022;39(5):380-5.
10. Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. N Engl J Med. 2019;380(14):1347-58.
11. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. Lancet Digit Health. 2019;1(6):e271-e97.
12. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542(7639):115-8.
13. Blanco-González A, Cabezón A, Seco-González A, Conde-Torres D, Antelo-Riveiro P, Piñeiro Á, et al. The Role of AI in Drug Discovery: Challenges, Opportunities, and Strategies. Pharmaceuticals [Internet]. 2023; 16(6).
14. Huang C, Clayton EA, Matyunina LV, McDonald LD, Benigno BB, Vannberg F, et al. Machine learning predicts individual cancer patient responses to therapeutic drugs with high accuracy. Scientific Reports. 2018;8(1):16444.
15. Kareemi H, Vaillancourt C, Rosenberg H, Fournier K, Yadav K. Machine Learning Versus Usual Care for Diagnostic and Prognostic Prediction in the Emergency Department: A Systematic Review. Acad Emerg Med. 2021;28(2):184-96.

16. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. npj Digital Medicine. 2020;3(1):17.
17. Wasylewicz ATM, Scheepers-Hoeks A. Clinical Decision Support Systems. In: Kubben P, Dumontier M, Dekker A, editors. Fundamentals of Clinical Data Science. Cham (CH): Springer Copyright 2019, The Author(s). 2019. p. 153-69.
18. Hoppe JM, Auer MK, Strüven A, Massberg S, Stremmel C. ChatGPT With GPT-4 Outperforms Emergency Department Physicians in Diagnostic Accuracy: Retrospective Analysis. J Med Internet Res. 2024;26:e56110.
19. Vasey B, Ursprung S, Beddoe B, Taylor EH, Marlow N, Bilbro N, et al. Association of Clinician Diagnostic Performance With Machine Learning-Based Decision Support Systems: A Systematic Review. JAMA Netw Open. 2021;4(3):e211276.
20. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med. 2022;28(5):924-33.
21. Sounderajah V, Ashrafian H, Golub RM, Shetty S, Fauw JD, Hooft L, et al. Developing a reporting guideline for artificial intelligence-centred diagnostic test accuracy studies: the STARD-AI protocol. BMJ Open. 2021;11(6):e047709.
22. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. Bmj. 2024;385:e078378.

23. Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ, Spirit AI, et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. Lancet Digit Health. 2020;2(10):e549-e60.
24. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, Spirit AI, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med. 2020;26(9):1364-74.
25. Li H, Hayward J, Aguilar LS, Franc JM. Desired clinical applications of artificial intelligence in emergency medicine: A Delphi study. Am J Emerg Med. 2024;79:217-20.
26. Eastwood KW, May R, Andreou P, Abidi S, Abidi SSR, Loubani OM. Needs and expectations for artificial intelligence in emergency medicine according to Canadian physicians. BMC Health Serv Res. 2023;23(1):798.
27. Amann J, Blasimme A, Vayena E, Frey D, Madai VI, the Precise Qc. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. BMC Medical Informatics and Decision Making. 2020;20(1):310.
28. Ancker JS, Edwards A, Nosal S, Hauser D, Mauer E, Kaushal R, et al. Effects of workload, work complexity, and repeated alerts on alert fatigue in a clinical decision support system. BMC Medical Informatics and Decision Making. 2017;17(1):36.

Chapter 3. Scoping Review

This chapter summarizes and describes the existing literature for AI-CDS tools for use in the ED. We address our first thesis objective: "Describe the available literature regarding AI-CDS tools that have been developed for individual patient care in the ED and the phases of development these tools have achieved." We conducted a scoping review of the literature to identify the types of tools that have been developed and the phases of development they have achieved.

Herein, we include the scoping review protocol manuscript, "Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Scoping Review Protocol" that was published online in *Open Science Framework*, and the scoping review manuscript, "Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Scoping Review" that was published in *Academic Emergency Medicine*. Reproduced with permission according to Wiley Article Sharing Policy.

3.1. Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Scoping Review Protocol

This section presents the protocol for a scoping review of AI-CDS tools for individual patient care in the ED. It includes the original manuscript, search strategy, and eligibility criteria that were published on December 11, 2023 in the Open Science Framework, a non-peer reviewed database that makes scientific work and protocols available for purposes of project management, transparency, and accountability. This section also includes updates to several terms and definitions that were agreed upon by the planning authors after conducting the initial title and abstract screening. The findings of the scoping review are presented in Section 3.2.

3.1.1. Preface

Contributions of Co-Authors: HK, KY, WM, and CV conceptualized and planned the study. HK devised and refined the search strategy, performed trial searches, reviewed the search with health sciences librarians, and wrote the manuscript. All authors edited, reviewed, and approved the final manuscript.

Ethics Approval: Not required.

Thesis-Related Appendices: None.

Citation: Kareemi H, Yadav K, Michalowski W, et al. Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Scoping Review Protocol [Internet]. OSF; 2025. Available from: osf.io/fdz3y

3.1.2. Title Page

Title

Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Scoping Review Protocol

Registration Open Science Framework DOI 10.17605/OSF.IO/FDZ3Y

Authors

Hashim Kareemi[1,2]

Krishan Yadav[3,4]

Wojtek Michalowski[5]

Christian Vaillancourt[3,4]

Affiliations

1. Department of Emergency Medicine, University of British Columbia, Vancouver, British Columbia, Canada

2. School of Epidemiology and Public Health, University of Ottawa, Ottawa, Ontario, Canada

3. Department of Emergency Medicine, University of Ottawa, Ottawa, Ontario, Canada

4. Ottawa Hospital Research Institute, Ottawa, Ontario, Canada

5. Telfer School of Management, University of Ottawa, Ottawa, Ontario, Canada

Corresponding Author

Hashim Kareemi

Emergency Medicine, Vancouver General Hospital

920 W 10th Ave, Vancouver, BC

V5Z 1M9

Support

Hashim Kareemi's thesis work for his Master's of Science in Epidemiology at the University of Ottawa, of which this work is a component, is supported by a Canadian Graduate Scholarship – Master's from the Canadian Institute of Health Research. The funder did not have any role in the development of this research protocol.

3.1.3. Introduction

Emergency medical care involves high-stakes decisions that are often limited by information and time. The added challenges of emergency department (ED) crowding, staffing issues and wait times complicate these decisions further.[1] Strategies to optimize clinician decision-making in these circumstances are needed. Clinical Decision Support (CDS) refers to a variety of resources and tools that provide or synthesize information with the intent of improving decision-making by clinicians and ultimately benefitting patient care and resource utilization.[2, 3] CDS may include non-computerized tools such as clinical guidelines and clinical decision rules, as well as "CDS systems", which refers to software that integrates individual patient information from an electronic health record.[2] In general, these tools provide support related to different components of the clinical decision making process, namely diagnosis, prognosis, and treatment, in the form of recommendations based on different thresholds. These thresholds may be defined as "knowledge-based" if they utilize rules based on literature- or expert-based recommendations.[4] With the advent and rapid advancement of artificial intelligence (AI) technologies and the digitization of massive amounts of previously uncaptured data (referred to as "big data"), there is increasing interest in "non-knowledge based" thresholds, which are based on advanced statistical and non-statistical modeling techniques such as machine learning (ML).[2, 5]

Studies have suggested that AI-CDS tools may match or even outperform clinicians and knowledge-based CDS,[6] including a systematic review of ED-based models done by our team.[7] However, most of these tools remain in preclinical "in silico" phases of

development and only demonstrate improved performance using observational data; there remains little evidence for their effect on patient outcomes when implemented in clinical practice and relatively few have been tested in clinical trials.[8-10] Several explanations for the disparity between the number of AI-CDS tools in development and implementation, the so-called "AI chasm",[11] have been raised. These include practical concerns such as integration with electronic health records and regulatory approvals, as well as ethical concerns regarding bias against marginalized populations and the "black box" of model explainability.[12, 13] A development pathway for AI-CDS tools has been proposed by the authors of the DECIDE-AI guidelines, which mirrors those of pharmaceutical and surgical innovations.[14] It includes the use of specific guidelines for each stage of development, including TRIPOD-AI and STARD-AI for "offline" validation of predictive (prognosis or diagnosis) and diagnostic test accuracy,[15, 16] respectively, DECIDE-AI for early "live" evaluations,[14] and SPIRIT-AI (for protocols) and CONSORT-AI (for reporting findings) for large-scale effectiveness trials.[17, 18] Efforts to operationalize this development pathway have been made by the authors of *Staged Clinical Artificial Intelligence Implementation* (SALIENT), a clinical AI implementation framework that identifies and proposes solutions to specific issues mapped to stages of the DECIDE-AI pathway.[19]

The current usage of AI-CDS tools and their translation to clinical trials and implementation in the ED setting have not been investigated adequately. Understanding the current landscape of these tools is a necessary step in exploring the barriers to their effective and appropriate use. A preliminary search of PubMed was conducted on November 1, 2023 for existing systematic and scoping reviews on this topic. Two scoping

reviews of AI technologies in the ED setting were found, however neither focused specifically on CDS and instead included studies related to non-CDS such as administrative task automation.[20, 21] Furthermore, there were methodological limitations in both studies that we feel precluded a comprehensive mapping of the current AI-CDS literature, including requiring terms related to "ethics" in the electronic search[21] and excluding conference abstracts and grey literature.[20, 21] To our knowledge, no ED-focused studies have incorporated the DECIDE-AI development pathway, and there remains limited clear, actionable guidance for researchers and policy makers that considers the specific resource constraints of emergency clinical decisions. At the same time, incorporating development guidelines that are agnostic to specific settings or medical subspecialties would allow translating solutions proposed by groups outside of the scope of the emergency department, such as the authors of SALIENT, who went on to validate their implementation framework using a systematic review of AI-based sepsis prediction models deployed primarily in ICU settings.[22]

Our objective is to review and synthesize the current evidence for artificial intelligence-based clinical decision support in the ED setting. Specifically, we intend to answer the following primary and secondary research questions: 1) what AI-CDS tools have been studied in the ED setting; and 2) what barriers and enablers have been identified regarding the translation of AI-CDS from pre-clinical development to clinical implementation in the ED setting.

3.1.4. Methods

We will perform a scoping review according to the Joanna Briggs institute methodology[23] and report our findings following the *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) Extension for Scoping Reviews.[24] This protocol has been reported according to the PRISMA Extension for Protocols.[25]

Eligibility Criteria

Inclusion Criteria

We will include studies involving AI- or ML-based CDS tools applied in the ED. Given the anticipated variability of these terms in the literature, we have created specific definitions based on our collective experience and study objectives for the following: *definition of AI and ML, definition of CDS, timing of output,* and *phase of development*. These can be found in the appendix. We will include full-text articles, conference abstracts and trial protocols of primary research including observational, interventional or implementation studies. In keeping with objectives and scoping nature of this review, we want to capture any relevant CDS tools that are in development or deployed. In our analysis, we will distinguish articles with full text publications from those with only conference abstracts and trial protocols. Studies must be published from the year 2010 to date of search. Since AI technologies are advancing rapidly, other scoping reviews of AI technologies have used a variety of year cut-offs to maintain study relevance, including 2000,[26] 2010,[27] 2014,[21, 28] and 2015.[29] A systematic review of AI-CDS systems

that subsequently informed the derivation of the DECIDE-AI guidelines used a cut-off date of 2010.[30] We have therefore selected 2010 as the cut-off for our search. For feasibility, we will only include English studies.

Exclusion Criteria

We will exclude studies not focusing on AI- or ML-based CDS, such as operational and departmental tasks (e.g. work scheduling, surge prediction, database construction, note generation) not related to the clinical decision making for an individual patient. Studies not directly applicable to ED care, including those focusing on patients in pre-hospital and hospitalized settings, will also be excluded. If the study involves a mixed cohort (e.g. ED and intensive care unit), then the inclusion criteria regarding *timing of output* must be met to warrant inclusion.

Information Sources

We will perform a comprehensive search of five electronic databases: MEDLINE (OVID), Embase (OVID), Scopus, Cumulative Index to Nursing and Allied Health Literature (CINAHL), and Cochrane Central (OVID). Conference abstracts will be included in the electronic database searches. Reference lists of relevant articles will be hand-searched for additional studies. Grey literature will be explored using a hand search of Google Scholar (first 50 results) and two trial registries (ClinicalTrials.gov and World Health Organization International Clinical Trials Registry Platform), using the following search:

(artificial intelligence or *machine learning)* and *emergency*. Dates of coverage will be January 1, 2010 to the search date (anticipated December 7, 2023).

Search Strategy

An electronic search for MEDLINE (OVID) was developed with the assistance of two health sciences librarians and then translated to the other databases. The searches are provided in the appendix. We used an additive search using three themes: artificial intelligence or machine learning; emergency department; and clinical decision support. We utilized search filters for studies related to artificial intelligence[31] and EDs.[32] For the CDS theme, we incorporated terms used in a systematic review of AI-CDS by Vasey et al.[30] We added search filters[33, 34] and terms related to components of the clinical decision making process (i.e. diagnosis, prognosis, treatment) in order to capture studies of AI-CDS tools that may not yet be defined as "clinical decision support" or a "clinical decision support system".

Data Management and Selection Process

All titles will be entered into COVIDENCE for deduplication and screening.[35] Title and abstract screening will be performed by two independent reviewers (HK and one other reviewer). The screening process will be piloted on the first 50 articles to ensure compatibility and functionality without major discrepancies between independent screeners. Full-text articles will be retrieved and assessed for inclusion by two independent reviewers (HK and one other reviewer). Disagreements will be resolved

through discussion or consensus with a third reviewer (WM, KY, or CV). Disagreements specific to the *definition of AI or ML* will be discussed with an author with an AI and ML content expert (WM).

Data Collection Process and Data Items

A standardized data extraction form will be developed and piloted on the first 3 full text articles to ensure functionality and ability to capture all relevant items. Data will be extracted independently by a single reviewer (HK). We will extract elements of the TRIPOD-AI and DECIDE-AI guidelines,[14, 15] including study characteristics (author, title, publication year, country, journal), study design (observational versus interventional, randomized versus non- or pseudo-randomized, cohort versus case-control versus cross-sectional, prospective versus retrospective data collection), phase of development (in silico, silent, early live clinical, comparative prospective, or vigilance), patient identifiers (age, gender, sex, presenting complaint, other specific selection criteria), intended timing and use of CDS tool, data input, AI or ML technique used and modifications done, main results, safety and error concerns, and key barriers and facilitators regarding development and implementation (e.g. patient and provider education or incentives, human factor and user experience considerations).

Data Analysis

We will perform a narrative synthesis of findings. In order to provide a clear summary of the current evidence for of AI-CDS in the ED, we will categorize included

studies by research type (published studies, conference abstract only, registered trial and stage of recruitment) and map according to the development framework proposed by the DECIDE-AI guidelines.[14] We will provide simple frequency counts of these concepts and the other extracted data.

Presentation of the Results

The anticipated tables and figures will include a flow diagram of the study selection process according to PRISMA-ScR,[24] Table 1 describing characteristics of the included studies, a diagram of the mapping exercise according to phase of development, and a table of the electronic search strategy for MEDLINE.

Risk of Bias

Given the scoping nature of the review, a risk of bias assessment will not be performed.

Ethics

This scoping review does not require ethical approval as it involves the analysis of publicly available literature.

References

1. Kelen GD, Wolfe R, D'Onofrio G, Mills AM, Diercks D, Stern SA, et al. Emergency Department Crowding: The Canary in the Health Care System. Catalyst non-issue content. 2021;2(5).
2. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. npj Digital Medicine. 2020;3(1):17.
3. Wasylewicz ATM, Scheepers-Hoeks A. Clinical Decision Support Systems. In: Kubben P, Dumontier M, Dekker A, editors. Fundamentals of Clinical Data Science. Cham (CH): Springer Copyright 2019, The Author(s). 2019. p. 153-69.
4. Sim I, Gorman P, Greenes RA, Haynes RB, Kaplan B, Lehmann H, et al. Clinical Decision Support Systems for the Practice of Evidence-based Medicine. Journal of the American Medical Informatics Association. 2001;8(6):527-34.
5. Deo RC. Machine Learning in Medicine. Circulation. 2015;132(20):1920-30.
6. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. Lancet Digit Health. 2019;1(6):e271-e97.
7. Kareemi H, Vaillancourt C, Rosenberg H, Fournier K, Yadav K. Machine Learning Versus Usual Care for Diagnostic and Prognostic Prediction in the Emergency Department: A Systematic Review. Acad Emerg Med. 2021;28(2):184-96.
8. Yin J, Ngiam KY, Teo HH. Role of Artificial Intelligence Applications in Real-Life Clinical Practice: Systematic Review. J Med Internet Res. 2021;23(4):e25759.

9. Plana D, Shung DL, Grimshaw AA, Saraf A, Sung JJY, Kann BH. Randomized Clinical Trials of Machine Learning Interventions in Health Care: A Systematic Review. JAMA Netw Open. 2022;5(9):e2233946.
10. Lam TYT, Cheung MFK, Munro YL, Lim KM, Shung D, Sung JJY. Randomized Controlled Trials of Artificial Intelligence in Clinical Practice: Systematic Review. J Med Internet Res. 2022;24(8):e37188.
11. Keane PA, Topol EJ. With an eye to AI and autonomous diagnosis. npj Digital Medicine. 2018;1(1):40.
12. Adam H, Balagopalan A, Alsentzer E, Christia F, Ghassemi M. Mitigating the impact of biased artificial intelligence in emergency decision-making. Communications Medicine. 2022;2(1):149.
13. Cadario R, Longoni C, Morewedge CK. Understanding, explaining, and utilizing medical artificial intelligence. Nature Human Behaviour. 2021;5(12):1636-42.
14. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med. 2022;28(5):924-33.
15. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. Ann Intern Med. 2015;162(1):55-63.
16. Sounderajah V, Ashrafian H, Golub RM, Shetty S, Fauw JD, Hooft L, et al. Developing a reporting guideline for artificial intelligence-centred diagnostic test accuracy studies: the STARD-AI protocol. BMJ Open. 2021;11(6):e047709.

17. Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ, Spirit AI, et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. Lancet Digit Health. 2020;2(10):e549-e60.
18. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, Spirit AI, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med. 2020;26(9):1364-74.
19. van der Vegt AH, Scott IA, Dermawan K, Schnetler RJ, Kalke VR, Lane PJ. Implementation frameworks for end-to-end clinical AI: derivation of the SALIENT framework. J Am Med Inform Assoc. 2023;30(9):1503-15.
20. Kirubarajan A, Taher A, Khan S, Masood S. Artificial intelligence in emergency medicine: A scoping review. J Am Coll Emerg Physicians Open. 2020;1(6):1691-702.
21. Masoumian Hosseini M, Masoumian Hosseini ST, Qayumi K, Ahmady S, Koohestani HR. The Aspects of Running Artificial Intelligence in Emergency Care; a Scoping Review. Arch Acad Emerg Med. 2023;11(1):e38.
22. van der Vegt AH, Scott IA, Dermawan K, Schnetler RJ, Kalke VR, Lane PJ. Deployment of machine learning algorithms to predict sepsis: systematic review and application of the SALIENT clinical AI implementation framework. J Am Med Inform Assoc. 2023;30(7):1349-61.
23. Peters MDJ, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth. 2020;18(10):2119-26.

24. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med. 2018;169(7):467-73.

25. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev. 2015;4(1):1.

26. Gama F, Tyskbo D, Nygren J, Barlow J, Reed J, Svedberg P. Implementation Frameworks for Artificial Intelligence Translation Into Health Care Practice: Scoping Review. J Med Internet Res. 2022;24(1):e32215.

27. Zidaru T, Morrow EM, Stockley R. Ensuring patient and public involvement in the transition to AI-assisted mental health care: A systematic scoping review and agenda for design justice. Health Expect. 2021;24(4):1072-124.

28. Cartolovni A, Tomicic A, Lazic Mosler E. Ethical, legal, and social considerations of AI-based medical decision-support tools: A scoping review. Int J Med Inform. 2022;161:104738.

29. Murphy K, Di Ruggiero E, Upshur R, Willison DJ, Malhotra N, Cai JC, et al. Artificial intelligence for good health: a scoping review of the ethics literature. BMC Med Ethics. 2021;22(1):14.

30. Vasey B, Ursprung S, Beddoe B, Taylor EH, Marlow N, Bilbro N, et al. Association of Clinician Diagnostic Performance With Machine Learning-Based Decision Support Systems: A Systematic Review. JAMA Netw Open. 2021;4(3):e211276.

31. Campbell S, Kung J. Filter to Retrieve Studies Related to Artificial Intelligence from the OVID MEDLINE Database. Geoffrey & Robyn Sperber Health Sciences Library, University of Alberta 2023.
32. Campbell S. A Filter to Retrieve Studies related to Emergency Departments from the OVID MEDLINE Database. Geoffrey & Robyn Sperber Health Sciences Library, University of Alberta. 2021.
33. Chojecki D, Kung J. A Filter to Retrieve Diagnostic Studies from the OVID MEDLINE Database. . Geoffrey & Robyn Sperber Health Sciences Library, University of Alberta 2023.
34. Wilczynski NL, Haynes RB. Developing optimal search strategies for detecting clinically sound prognostic studies in MEDLINE: an analytic survey. BMC Med. 2004;2:23.
35. Veritas Health Information. Covidence systematic review software. Melbourne, Australia2023.

Appendix

Appendix 3.1 – Table 1. Search Strategy

Database(s): Ovid MEDLINE(R) ALL 1946 to December 07, 2023

#	Searches	Results
1	exp Artificial Intelligence/ or (((machin* or artific* or comput* or robot* or automat*) adj3 (intelligen* or reasoning)) or ((assist* or augment* or autonomous) adj1 intelligen*) or ((machin* or deep or transfer or hierarchical) adj2 learning) or "classification algorithm*" or "computer heuristic*" or "feature detection" or "generative pre-trained transformer" or "language learning model*" or "large language model*" or "learning algorithm*" or (Markov adj3 model*) or "natural language process*" or "nearest neighbor*" or "neural network*" or "random forest" or "representation learning" or "support vector machine*" or "transfer learning" or "Bing chat" or ChatGPT* or "Chat GPT" or "Google* Bard" or "IBM Watson" or "Microsoft* Bing" or OpenAI or "Open AI" or PathAI or "Path AI").ti,ab,kf.	370663
2	Emergency Treatment/ or Emergency Medical Services/ or Emergencies/ or Emergency Service, Hospital/ or Evidence-Based Emergency Medicine/ or Trauma Centers/ or Triage/ or (((emergenc* or ED) adj1 (room* or accident or ward or wards or unit or units or department* or physician* or doctor* or nurs* or treatment* or visit*)) or (triage adj2 (centre or centres or center or centers or department? or unit or units)) or (urgent adj2 (care or healthcare or health care)) or critical care or (trauma adj2 (centre or centres or center or centers or department? or unit or units)) or ("accident and emergency" or "accident & emergency") or ((emergency or non-emergency or nonemergency or urgent or non-urgent or nonurgent or semi-urgent or semiurgent) adj2 patient?)).ti,ab,kf.	359822
3	Decision Making, Computer-Assisted/ or Diagnosis, Computer-Assisted/ or Therapy, Computer-Assisted/ or Drug Therapy, Computer-Assisted/ or exp Decision Support Systems, Clinical/ or Clinical Decision Rules/ or Clinical Decision-Making/ or (CDSS* or CCDSS* or ((decision* or decid*) adj2 (making* or make* or support)) or ((computer* or digit*) adj1 (assist* or aid or support* or help*))).ti,ab,kf.	345451
4	Delayed Diagnosis/ or Diagnosis, Differential/ or Diagnostic Errors/ or "Diagnostic Techniques and Procedures"/ or Early Diagnosis/ or Overdiagnosis/ or "Sensitivity and Specificity"/ or (diagnos* or sensitivit* or specificit* or receiver operat* or roc).ti,ab,kf.	4665529
5	Prognosis/ or Treatment Outcome/ or Incidence/ or Patient Outcome Assessment/ or (prognos* or predict*).ti,ab,kf.	4124035
6	or/3-5	7974024
7	and/1-2,6	2526
8	limit 7 to yr="2010 -Current"	2321

General approach to building search strategy:

- Look for a search filter
- Interrogate associated MeSH terms
- For Terms 1, 2 and 3a: terms were corroborated with search of reference systematic or scoping revieW
- For Terms 3b-3d: terms were corroborated with titles and abstracts of target articles

Term 1: Artificial Intelligence

exp Artificial Intelligence/ or (((machin* or artific* or comput* or robot* or automat*) adj3 (intelligen* or reasoning)) or ((assist* or augment* or autonomous) adj1 intelligen*) or ((machin* or deep or transfer or hierarchical) adj2 learning) or "classification algorithm*" or "computer heuristic*" or "feature detection" or "generative pre-trained transformer" or "language learning model*" or "large language model*" or "learning algorithm*" or (Markov adj3 model*) or "natural language process*" or "nearest neighbor*" or "neural network*" or "random forest" or "representation learning" or "support vector machine*" or "transfer learning" or "Bing chat" or ChatGPT* or "Chat GPT" or "Google* Bard" or "IBM Watson" or "Microsoft* Bing" or OpenAI or "Open AI" or PathAI or "Path AI").ti,ab,kf.

Notes:

- Changed .mp → .ti,ab,kf for relevance
- Added:
    - (((machin* or artific* or comput* or robot* or automat*) adj3 (intelligen* or reasoning))
    - ((assist* or augment* or autonomous) adj1 intelligen*)
    - ((machin* or deep or transfer or hierarchical) adj2 learning)
- Removed:
    - AI to mitigate irrelevant terms (e.g. adrenal insufficiency, artificial insemination)
    - ((multifactor* or multicriteria) adj3 ("decision analysis" or "decision making"))
    - "decision tree"
    - "outlier detection"
    - "pattern recognition"
    - Terms captured by above added terms

Search filter: Campbell 2023[1]
Reference study: Vasey 2021[2]

Term 2: Emergency Department

Emergency Treatment/ or Emergency Medicine/ or emergency medical services/ or emergency service, hospital/ or trauma centers/ or triage/ or exp Evidence-Based Emergency Medicine/ or exp Emergency Nursing/ or Emergencies/ or casualty department* or ((emergenc* or ED) adj1 (room* or accident or ward or wards or unit or units or

department* or physician* or doctor* or nurs* or treatment* or visit*)).mp. or (triage or critical care or (trauma adj1 (cent* or care))).mp.

Notes:

- Removed emergicent*
- Changed .mp to .ti,ab,kf for relevance

Search filter: Campbell 2021[3]
Reference study: Kirubarajan 2020[4]

Term 3a: Clinical Decision Support

*Decision Making, Computer-Assisted/ or Diagnosis, Computer-Assisted/ or *Therapy, Computer-Assisted/ or Drug Therapy, Computer-Assisted/ or exp Decision Support Systems, Clinical/ or (CDSS* or CCDSS* or "decision support" or "decision making" or algorithm* or ((computer* or digit*) adj1 (assist* or aid or support* or help*))).ti,ab,kw.

Notes:

- Added:
    - Clinical Decision-Making/
    - Clinical Decision Rules/
    - ((computer* or digit*) adj1 (assist* or aid or support* or help*) to provide more thorough coverage of these terms
    - ((decision* or decid*) adj2 (making* or make* or support)) to provide more thorough coverage of these terms
- Removed:
    - *Algorithms/ and algorithm*
    - Terms covered by above added broad terms
    - “diagnos* support” (covered by diagnos* below)

Search filter: (none identified)
Reference study: Vasey 2021[2]

Term 3b: Diagnostic Studies
Delayed Diagnosis/ or Diagnosis, Differential/ or Diagnostic Errors/ or "Diagnostic Techniques and Procedures"/ or Early Diagnosis/ or Overdiagnosis/ or "Sensitivity and Specificity"/ or (diagnos* or sensitivit* or specificit* or receiver operat* or roc).ti,ab,kf. exp "Sensitivity and Specificity"/ or exp Diagnosis/ or (accurac* or probabilit* or "detection rate*").tw,kw. or (predictive adj3 value*).tw. or (diagnos* or performance).tw,kw.

Notes:

- Removed for relevance:
    - Exp Reference Values/

    - Likelihood Functions/
    - False positive*.tw,kw
    - False negative*.tw,kw
    - False rate*.tw,kw
    - Screen positive.tw,kw
    - Reference value*.tw.
    - Likelihood.tw,kw
    - accura*.tw
    - reliab*.tw
    - observer adj variation*.tw
    - "multiples of the median".tw
    - MoM.tw,kw
- Changed:
    - diagnosis.tw,kw → diagnos*.tw,kw for comprehensive search
- exp Diagnosis/ had several irrelevant subheadings, so replaced with following MeSH subheadings:
    - (several covered under Term 3a)
    - Delayed Diagnosis
    - Diagnosis, Differential
    - Diagnostic Errors
    - Diagnostic Techniques and Procedures
    - Early Diagnosis
    - Overdiagnosis

Search filter: Chojecki 2023[5]
Target studies and identified terms:

- Taylor 2018[6]: *Diagnosis, diagnostic, diagnosed, performance, tests, prediction tools, predicting, train, validate, retrospective cohort analysis, area under the curve, performing, specificity, sensitivity, false positive, true negative, false negative, true positive*
- Baxt 2002[7]: *Identify, diagnosis, trained, accurately, train, test, recognize, evaluation, accuracy, sensitivity, specificity*
- Delahanty 2019[8]: *Development, screening, surveillance, derive, evaluate, alert rate, area under the receiver operator characteristic curve, sensitivity, specificity, precision, performance, precise, discriminant*
- Molaei 2016[9]: *Identifying, evaluate, determine, sensitive, diagnostic, accuracy, accurate, sensitivity, area under the ROC curve, specificity, accuracy*

Term 3c: Prognostic Studies
Treatment Outcome/ or Incidence/ or Patient Outcome Assessment/ or (prognos* or predict*).ti,ab,kf.

Notes:

- Added:

    - Treatment Outcome/
    - Incidence/
    - Patient Outcome Assessment/
- Removed for relevance:
    - incidence.sh
    - exp mortality.sh
    - follow-up studies.sh
    - course:.tw

Search filter: Wilczynski 2004[10]
Target studies and identified terms:

- Levin 2018[11]: *Triage, risk-stratify, differentiate, predicts, likelihood, outcomes, predictions, predicted, patient outcomes, predictive, accurately, prospective*
- Taylor 2016[12]: *Prediction, predictive, scoring systems, models, rules, clinical decision rules, predict, diagnosis, training, validation, area under receiver operating characteristic curve (AUC), outperformed, prospectively, clinical outcomes, predictions*
- Liu 2014[13]: *Prediction, risk, outcome, predicting, area under the curve (AUC), sensitivity, specificity, predictors, outperformed*

Term 3d: Treatment Studies
~~exp Contraindications/ or Medication Errors/ or Drug Administration Routes/ or Drug Administration Schedule/ or Drug Dosage Calculations/ or Drug Prescriptions/ or Anticoagulation Reversal/ or exp Airway Management/ or Intubation/ or Thrombolytic Therapy/ or Fluid Therapy/ or Manipulation, Orthopedic/ or Rewarming/~~

Notes:

- MeSH terms selected based on clinical expertise of HK
- Challenging to find relevant target studies (even the target below is not ED specific); most involved diagnosis or prognosis. Therefore, limited the search terms here to MeSH headings.
- Only added 17 irrelevant articles to the search
- Therapeutic considerations are already covered MeSH terms and associated key terms above:
    - Emergency Treatment/
    - *Therapy, Computer-Assisted/
    - Drug Therapy, Computer-Assisted/
- Therefore, the line for treatment studies was removed (November 27 updated)

Search filter: none found
Target studies and identified terms:

- Corny 2020[14]: *patient safety, clinical outcomes, prescribing errors, accuracy, prescription, pharmacist, predicted, area under the receiving-operating characteristic curve, tools, validation, precision-recall, prescription errors*

Appendix 3.1 – Table 2. Eligibility Criteria

Inclusion Criteria

1. Studies involving AI- or ML-based CDS tools applied in the ED.
    - *Definition of AI or ML:* authors should refer to the CDS tool as an example of "artificial intelligence", "AI", "machine learning", "deep learning", "neural network", or "natural language processing". We have encountered variable definitions of these terms in the literature based on our previous work,[15] and we will therefore accept a broad definition in keeping with our objective and other reviews.[16] Any disputes about inclusion related to the *definition of AI or ML* will be resolved in consensus with an author with expertise in AI and ML methodology (WM).
    - *Definition of CDS*: by "clinical decision support", we are referring to any model, algorithm or system that provides recommendations to a clinician regarding diagnostic, prognostic, or therapeutic decisions related to an individual ED patient's care.
    - *Timing of Output*: the recommendation of the CDS tool must be applicable to an individual ED patient's care during their index ED visit (i.e. from registration or triage to discharge or admission to hospital). This may include outcomes related to future return to ED or future hospitalization if the output of the recommendations affects decisions and are (or conceivably could be) available at the time of index ED visit (e.g. decision to discharge, extended duration of therapy or diagnostic testing after discharge). The input data may include prehospital data if it is used to inform a clinical decision at the time of the ED visit.
    - *Phase of Development*: The CDS tool may be in any phase of development as defined by the DECIDE-AI guidelines,[17] including preclinical/"in silico", silent/"shadow" evaluation, small-scale "live" clinical evaluation, large-scale comparative prospective evaluation, or post-market vigilance.
2. Full-text articles, conference abstracts and trial protocols of primary research including observational, interventional or implementation studies.
    - In keeping with objectives and scoping nature of this review, we want to capture any relevant CDS tools that are in development or deployed. In our analysis, we will distinguish articles with full text publications from those with only conference abstracts and trial protocols.
3. Publications from the year 2010 to date of search.
    - AI technologies are advancing rapidly. Other scoping reviews of AI technologies have used a variety of year cut-offs including 2000,[18] 2010,[19] 2014,[20, 21] and 2015.[22]
    - A systematic review of AI-CDS systems that subsequently informed the derivation of the DECIDE-AI guidelines used a cut-off date of 2010.[2]

    - For feasibility and relevance, we therefore selected 2010 as the cut-off for our search.
4. English language studies only.
    - For feasibility, we will only include English studies.

Exclusion Criteria:

- Studies not focusing on AI- or ML-based CDS.
    - Non-clinical tasks would include operational and departmental tasks (e.g. work scheduling, surge prediction, database construction) not related to the care of an individual patient.
- Studies not directly applicable to ED care, including those focusing on patients in prehospital and hospitalized settings.
    - If the study involves a mixed cohort (e.g. ED and intensive care unit), then the inclusion criteria regarding Timing of Output must be met to warrant inclusion.
- Studies published before the year 2010.
- Non-English studies.

Appendix 3.1 – Table 3. Post-Screening Protocol Updates

Post-Screening Updates – February 3, 2024

Inclusion Criteria

- Definition of CDS – 1: We will further refine "clinical decision support" to those that involve multiple clinical variables including history of presenting event, physical exam findings, vital signs, investigations in the emergency department, and past medical history, including that accessible through the electronic medical record. We will exclude studies of models that interpret specific diagnostic tests or image analysis only, such as radiographs (e.g. x-ray or CT or MRI), electrocardiogram, specific laboratory investigations, retinal fundoscopic images, and photographs of rashes, unless the AI model additionally consider multiple other clinical variables.
    - Rationale: While some of these theoretically could be useful during the ED assessment (e.g. identification of abnormalities on chest radiograph), the focus of this scoping review is on supporting complex clinical decisions that are in the purview of the emergency clinician, as opposed to expertise in interpreting specific diagnostic tests that are within the scope of respective consultants such as radiologists, ophthalmologists, dermatologists, etc.
    - N.B. We "tagged" these studies in Covidence and will separate them from the remainder of the full text review. We may consider a separate review or analysis of these studies as part of an "AI for emergency diagnostics" study.
- Definition of CDS – 2: Outcomes related to operations and logistics, specifically ED length of stay or wait times, were excluded.
    - Rationale: The planning authors (HK, CV, KY, WM) did not feel that prediction of these outcomes would necessarily influence clinical decision making for an individual patient (e.g. knowing that a patient will wait 2 hours versus 10 hours would not make a clinician order different tests for that patient; alternatively, knowing a patient has high likelihood of clinical decompensation or ICU admission may affect diagnostic testing or expediting consultation).
- Timing of Output – 1: In cases where the timing of the primary outcome is of uncertain relevance to the emergency clinician (e.g. risk of subsequent suicide attempt within 1 year after index ED visit), a discussion and decision was made between the co-reviewers, all of whom are emergency medicine physicians working at large urban centres in Canada. If there remained uncertainty after discussion, the case would have corroborated with one of the emergency clinicians from the planning authors (CV or KY), however there were no cases in which this was required.
- Timing of Output – 2: In cases where the input data used by the AI models was unclear (e.g. predicting in-hospital mortality using "admission data", which in some cases referred to data from the ED assessment and in other cases referred to data including repeat vitals/bloodwork upon admission to an inpatient ward), studies

were only included if they explicitly stated the timing of the input data or specific variables considered were from the ED assessment, or that the model output would inform ED management.
    - In keeping with this, studies using large trauma or ICU databases were excluded unless the latter criteria were met.
    - If the co-reviewers felt there was a high likelihood of relevance to ED studies, studies were included for further clarification on full text review.
    - Similarly, studies using natural language processing (NLP) to mine unstructured data retrospectively from emergency physician charts for the purpose of epidemiological surveillance or improving diagnostic coding accuracy were excluded.
- Timing of Output – 3: Studies that used animal models or simulated cases were excluded.
- Definition of AI or ML: In some instances, relevant studies did not refer to their models as "AI", "artificial intelligence", "ML", or "machine learning". The primary author (HK), who has experience with AI research and co-screened every study, included these studies for full text review, which he will co-review with a second ED clinician with AI research experience.

Exclusion Criteria:

- Based on the above, we have refined the reasons for exclusion for the full text review, including a new category referred to as "ED Input":
    - Not CDS – 1: Specific Diagnostic Test/Finding Only
    - Not CDS – 2: Operations/Logistic Only (e.g. LOS, wait time)
    - Not ED Input – 1*: Considers Post-ED Data (e.g. Inpatient, Post-surgical)
        - *If the study involves a mixed cohort (e.g. ED and ICU), then the inclusion criteria regarding Timing of Output (i.e. outcome can be made at time of ED assessment) must be met to warrant inclusion.
    - Not ED Input – 2: Animal Model, Simulated Cases/Vignettes, etc.
    - Not ED Output: Outcome Not Relevant to ED Assessment
    - Not AI/ML
    - Not Primary Study (e.g. Review, Editorial, Commentary)
    - Not English

Appendix 3.1 – References

1. Campbell S, Kung J. Filter to Retrieve Studies Related to Artificial Intelligence from the OVID MEDLINE Database. Geoffrey & Robyn Sperber Health Sciences Library, University of Alberta 2023.
2. Vasey B, Ursprung S, Beddoe B, Taylor EH, Marlow N, Bilbro N, et al. Association of Clinician Diagnostic Performance With Machine Learning-Based Decision Support Systems: A Systematic Review. JAMA Netw Open. 2021;4(3):e211276.
3. Campbell S. A Filter to Retrieve Studies related to Emergency Departments from the OVID MEDLINE Database. Geoffrey & Robyn Sperber Health Sciences Library, University of Alberta. 2021.
4. Kirubarajan A, Taher A, Khan S, Masood S. Artificial intelligence in emergency medicine: A scoping review. J Am Coll Emerg Physicians Open. 2020;1(6):1691-702.
5. Chojecki D, Kung J. A Filter to Retrieve Diagnostic Studies from the OVID MEDLINE Database. . Geoffrey & Robyn Sperber Health Sciences Library, University of Alberta 2023.
6. Taylor RA, Moore CL, Cheung KH, Brandt C. Predicting urinary tract infections in the emergency department with machine learning. PLoS One. 2018;13(3):e0194085.
7. Baxt WG, Shofer FS, Sites FD, Hollander JE. A neural network aid for the early diagnosis of cardiac ischemia in patients presenting to the emergency department with chest pain. Ann Emerg Med. 2002;40(6):575-83.

8. Delahanty RJ, Alvarez J, Flynn LM, Sherwin RL, Jones SS. Development and Evaluation of a Machine Learning Model for the Early Identification of Patients at Risk for Sepsis. Ann Emerg Med. 2019;73(4):334-44.
9. Molaei S, Korley FK, Reza Soroushmehr SM, Falk H, Sair H, Ward K, et al. A machine learning based approach for identifying traumatic brain injury patients for whom a head CT scan can be avoided. Annu Int Conf IEEE Eng Med Biol Soc. 2016;2016:2258-61.
10. Wilczynski NL, Haynes RB. Developing optimal search strategies for detecting clinically sound prognostic studies in MEDLINE: an analytic survey. BMC Med. 2004;2:23.
11. Levin S, Toerper M, Hamrock E, Hinson JS, Barnes S, Gardner H, et al. Machine-Learning-Based Electronic Triage More Accurately Differentiates Patients With Respect to Clinical Outcomes Compared With the Emergency Severity Index. Ann Emerg Med. 2018;71(5):565-74 e2.
12. Taylor RA, Pare JR, Venkatesh AK, Mowafi H, Melnick ER, Fleischman W, et al. Prediction of In-hospital Mortality in Emergency Department Patients With Sepsis: A Local Big Data-Driven, Machine Learning Approach. Acad Emerg Med. 2016;23(3):269-78.
13. Liu N, Koh ZX, Goh J, Lin Z, Haaland B, Ting BP, et al. Prediction of adverse cardiac events in emergency department patients with chest pain using machine learning for variable selection. BMC Med Inform Decis Mak. 2014;14:75.
14. Corny J, Rajkumar A, Martin O, Dode X, Lajonchère J-P, Billuart O, et al. A machine learning–based clinical decision support system to identify prescriptions

with a high risk of medication error. Journal of the American Medical Informatics Association. 2020;27(11):1688-94.

15. Kareemi H, Vaillancourt C, Rosenberg H, Fournier K, Yadav K. Machine Learning Versus Usual Care for Diagnostic and Prognostic Prediction in the Emergency Department: A Systematic Review. Acad Emerg Med. 2021;28(2):184-96.

16. Lam TYT, Cheung MFK, Munro YL, Lim KM, Shung D, Sung JJY. Randomized Controlled Trials of Artificial Intelligence in Clinical Practice: Systematic Review. J Med Internet Res. 2022;24(8):e37188.

17. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med. 2022;28(5):924-33.

18. Gama F, Tyskbo D, Nygren J, Barlow J, Reed J, Svedberg P. Implementation Frameworks for Artificial Intelligence Translation Into Health Care Practice: Scoping Review. J Med Internet Res. 2022;24(1):e32215.

19. Zidaru T, Morrow EM, Stockley R. Ensuring patient and public involvement in the transition to AI-assisted mental health care: A systematic scoping review and agenda for design justice. Health Expect. 2021;24(4):1072-124.

20. Cartolovni A, Tomicic A, Lazic Mosler E. Ethical, legal, and social considerations of AI-based medical decision-support tools: A scoping review. Int J Med Inform. 2022;161:104738.

21. Masoumian Hosseini M, Masoumian Hosseini ST, Qayumi K, Ahmady S, Koohestani HR. The Aspects of Running Artificial Intelligence in Emergency Care; a Scoping Review. Arch Acad Emerg Med. 2023;11(1):e38.

22. Murphy K, Di Ruggiero E, Upshur R, Willison DJ, Malhotra N, Cai JC, et al. Artificial intelligence for good health: a scoping review of the ethics literature. BMC Med Ethics. 2021;22(1):14.

3.2. Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Scoping Review

This section presents the manuscript for a scoping review of AI-CDS tools for individual patient care in the ED. It includes the original manuscript that was published in *Academic Emergency Medicine* on February 4, 2025, including materials that were included as online supplementary material (see section 3.2.j. Appendix). For brevity and space limitations, Table S5 (15 pages) and Table S6 (78 pages) have been excluded from this report but can be found referred to at https://doi.org/10.1111/acem.15099.

3.2.1. Preface

Contributions of Co-Authors: HK, KY, WM, and CV conceptualized and planned the study. HK, CP, NB, AM, HL, and GG performed the title and abstract screening. HK, SM, LG, and MB performed the full text screening. HK wrote the manuscript. All authors edited, reviewed, and approved the final manuscript.

Ethics Approval: Not required.

Thesis-Related Appendices: None.

Citation: Kareemi H, Yadav K, Price C, et al. Artificial intelligence-based clinical decision support in the emergency department: A scoping review. *Acad Emerg Med*. 2025;32(4):386-395. doi:10.1111/acem.15099

3.2.2. Title Page

Title
Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Scoping Review

Running Title AI-CDS in the ED: A Scoping Review

Authors
Hashim Kareemi, MD[1,2]
Krishan Yadav, MD MSc[2,3,4]
Courtney Price, MD[3]
Niklas Bobrovitz, MD[5,6]
Andrew Meehan, MD[3]
Henry Li, MD[7,8]
Gautam Goel, MD[9,10]
Sameer Masood, MD[11,12]
Lars Grant, MD[13,14]
Maxim Ben-Yakov, MDCM, MHI[11,12]
Wojtek Michalowski, ME PhD[15]
Christian Vaillancourt, MD[2,3,4]

Affiliations
1. Department of Emergency Medicine, University of British Columbia, Vancouver, BC, Canada
2. School of Epidemiology and Public Health, University of Ottawa, Ottawa, ON, Canada
3. Department of Emergency Medicine, University of Ottawa, Ottawa, ON, Canada
4. Ottawa Hospital Research Institute, University of Ottawa, ON, Canada
5. Department of Emergency Medicine, University of Calgary, Calgary, AB, Canada
6. Centre for Health Informatics, University of Calgary, Calgary, AB, Canada
7. Department of Emergency Medicine, Faculty of Medicine & Dentistry, University of Alberta, Edmonton, AB, Canada
8. Division of Pediatrics, Department of Pediatrics, Faculty of Medicine & Dentistry, University of Alberta, Edmonton, AB, Canada
9. Department of Emergency Medicine, Queensway Carleton Hospital, Ottawa, ON
10. Whiting School of Engineering, Johns Hopkins University, Baltimore, MD, USA
11. Department of Medicine, University of Toronto, Toronto, ON, Canada
12. Toronto General Hospital Research Institute, University Health Network, Toronto, ON, Canada
13. Department of Emergency Medicine, McGill University, Montreal, QC, Canada
14. Lady Davis Institute for Medical Research, Montreal, QC, Canada
15. Telfer School of Management, University of Ottawa, Ottawa, ON, Canada

Corresponding Author
Hashim Kareemi

Address: Emergency Medicine, 910 W 10th Ave, Vancouver, BC, V5Z 1M9

Financial Support
Hashim Kareemi's Master's of Science thesis work, of which this study is a component, is supported by a Canadian Graduate Scholarship - Master's.

Conflicts of Interest

- Hashim Kareemi:
  - Received Canadian Institutes of Health Research's Canada Graduate Scholarship – Master's to fund this work.
- Krishan Yadav: None
- Courtney Price: None
- Niklas Bobrovitz: None
- Andrew Meehan: None
- Henry Li:
  - Received grants from Canadian Medical Association and Canadian Association of Emergency Physicians to fund separate work.
- Gautam Goel:
  - Chief Medical Information Officer and owns shares in Hero AI, a clinical automation platform company.
- Sameer Masood:
  - Received Academic Medical Organization Grant to fund separate work.
  - Received honorarium for plenary presentation, "Artificial Intelligence in Emergency Medicine"
  - Co-chair of Clinical Decision Support Committee at the University Health Network, Toronto, ON.
- Lars Grant:
  - Medical consultant in the development of a Connected Health Record for CIUSSS Centre-Ouest-de-L'Ile, Montreal, QC.
  - Received honorarium for McGill Continuing Professional Education on Artificial Intelligence in Medicine, Montreal, QC.
  - Member of Data Governance Committee of CIUSSS Centre-Ouest-de-L'Ile, Montreal, QC.
- Maxim Ben-Yakov: None
- Wojtek Michalowski: None
- Christian Vaillancourt:
  - Received grants from Canadian Institutes of Health Research's Planning and Dissemination Grant, Canadian Institutes of Health Research's Operating Grant, and The Ottawa Hospital Academic Medical Organization for separate work.

Data Sharing Statement
All data will be made available upon request to Dr. Hashim Kareemi to investigators who provide an IRB letter of approval.

Author Contributions
HK, KY, WM, and CV conceptualized and planned the study.
HK, CP, NB, AM, HL, and GG performed the title and abstract screening.
HK, SM, LG, and MB performed the full text screening.
HK wrote the manuscript.
All authors edited, reviewed, and approved this manuscript.

Acknowledgments
Lindsey Sikora and Valentina Ly, health sciences librarians at the University of Ottawa, for their assistance in developing and reviewing the electronic search strategy.

Word Count
Abstract: 252/300
Manuscript: 4057/5000

3.2.3. Abstract

Objective

Artificial intelligence (AI) based clinical decision support (CDS) has the potential to augment high-stakes clinical decisions in the emergency department (ED). However, its current usage and translation to implementation remains poorly understood. We asked:

1) What is the current landscape of AI-CDS for individual patient care in the ED?

2) What phases of development have AI-CDS tools achieved?

Methods

We performed a scoping review of AI for prognostic, diagnostic, and treatment decisions regarding individual ED patient care. We searched five databases (MEDLINE, EMBASE, Cochrane Central, Scopus, Web of Science) and grey literature sources from January 1, 2010 to December 11, 2023. We adhered to guidelines from the Joanna Briggs Institute and PRISMA Extension for Scoping Reviews. We published our protocol on Open Science Framework (DOI 10.17605/OSF.IO/FDZ3Y).

Results

Of 5,168 unique records identified, we selected 605 studies for inclusion. The majority (369, 61%) were published in 2021-2023. The studies ranged over a variety of clinical applications, patient populations, and AI model types. Prognostic outcomes were most commonly assessed (270, 44.6%), followed by diagnostic (193, 31.9%) and disposition

(115, 19%). Most studies remained in the earliest phase of preclinical development (572, 94.5%) with few advancing to later phases (33, 5.5%).

Conclusions

By thoroughly mapping the landscape of AI-CDS in the ED, we demonstrate a rapidly increasing volume of studies covering a breadth of clinical applications, yet few have achieved advanced phases of testing or implementation. A more granular understanding of the barriers and facilitators to implementing AI-CDS in the ED is needed.

3.2.4. Introduction

Background

Emergency medical care involves high-stakes decisions that are often limited by information and time. The added challenges of emergency department (ED) crowding, staffing issues and wait times further complicate these decisions.[1] Strategies to optimize clinician decision-making in these circumstances are needed. One such proposed strategy is the use of Clinical Decision Support (CDS), referring to tools that synthesize information to improve clinician decision-making, benefit patient care, and optimize resource utilization.[2, 3] CDS may include non-computerized tools such as clinical guidelines and clinical decision rules, as well as "CDS systems", which refers to software that integrates individual patient information from an electronic health record.[2] In general, these tools provide support related to different components of the clinical decision making process - namely diagnosis, prognosis, and treatment - in the form of recommendations based on

different thresholds. These thresholds may be defined as “knowledge-based” if they utilize rules based on literature- or expert-based recommendations.[4] With the rapid advancement of artificial intelligence (AI) technologies and the digitization of massive amounts of previously uncaptured data (referred to as “big data”), there is increasing interest in “non-knowledge based” thresholds, which are based on advanced statistical and non-statistical modeling techniques known as machine learning (ML).[4, 5]

Previous studies suggest that AI-CDS tools may match or even outperform clinicians and knowledge-based CDS,[6] including a systematic review of ED-based models published by our team.[7] However, the majority of these tools remained in preclinical “in silico” phases of development and only demonstrated improved performance using observational data; there remains little evidence for their effect on patient outcomes when implemented in clinical practice, with relatively few having been tested in clinical trials.[8-10] Several explanations for the disparity between the number of AI-CDS tools in development and implementation, the so-called “AI chasm”,[11] have been raised. These include practical concerns such as integration with electronic health records and regulatory approvals, as well as ethical concerns regarding bias against marginalized populations and the “black box” of model explainability.[12,13] A reporting pathway for AI-CDS tools has recently been proposed by the authors of the DECIDE-AI guidelines, which mirrors those of pharmaceutical and surgical innovations.[14] It includes the use of specific guidelines for each stage of development, including TRIPOD-AI and STARD-AI for “offline” validation of predictive (prognosis or diagnosis) and diagnostic test accuracy,[15,

16] respectively, DECIDE-AI for early “live” evaluations,[14] and SPIRIT-AI (for protocols) and CONSORT-AI (for reporting findings) for large-scale effectiveness trials.[17, 18]

Importance

The usage of AI-CDS tools and their translation to implementation in the ED setting has not been investigated adequately. Understanding the current landscape of these tools is a critical first step in exploring the barriers to their effective and appropriate use. A preliminary search of PubMed was conducted on November 1, 2023 for existing systematic and scoping reviews on this topic. Two scoping reviews of AI technologies in the ED setting were found; however, neither focused specifically on CDS and instead included studies related to non-CDS such as administrative task automation.[19, 20] Furthermore, there were methodological limitations in both studies that we feel preclude a comprehensive mapping of the current AI-CDS literature, including the requirement of terms related to “ethics” in the electronic search,[20] and excluding conference abstracts and grey literature.[19, 20] Finally, to our knowledge, no emergency medicine-specific reviews have incorporated the DECIDE-AI pathway, and thus there also remains an incomplete understanding of the status of emergency AI research in regards to these phases of development.

Goals of This Investigation

Our objective is to review and synthesize the current application of AI-CDS in the ED setting. We intend to answer the following primary and secondary research questions: 1)

What is the current landscape of AI-CDS tools for prognostic, diagnostic and treatment decisions for individual patients in the ED? and 2) What phase of development have these AI-CDS tools achieved?

3.2.5. Methods

Protocol and Registration

We performed a scoping review according to the Joanna Briggs Institute methodology and have reported our findings following the PRISMA Extension for Scoping Reviews (see Table S1).[21, 22] We published a protocol according to the PRISMA Extension for Protocols on the Open Science Framework on December 11, 2023 and updated it on February 3, 2024 prior to full text screening (DOI 10.17605/OSF.IO/FDZ3Y). [23, 24]

Search Strategy and Information Sources

With the assistance of two health sciences librarians, we developed an electronic search for MEDLINE (OVID) and Embase (OVID). After refining the search terms for appropriateness, we translated the search to three additional databases (Scopus, CINAHL, and Cochrane Central) and had a health sciences librarian review the final search for accuracy. We ran this final search on all five databases from January 1, 2010 to December 11, 2023. Given the speed with which AI technologies are advancing, similar reviews have used a variety of start date cut-offs to ensure clinical relevance including 2000,[25] 2010,[26] 2014,[20, 27] and 2015.[28] A systematic review of AI-CDS systems that

subsequently informed the derivation of the DECIDE-AI guidelines used a cut-off date of 2010.[29] We therefore selected 2010 as the onset search date cut-off for our search. As we included abstracts and protocols in addition to full text articles, we did not contact authors for additional resources. For feasibility, we only included English articles and non-English studies with English abstracts.

For the search strategy, we used an additive search using three themes: *artificial intelligence or machine learning*, *emergency department*, and *clinical decision support*. We utilized search filters for studies related to artificial intelligence and EDs.[30, 31] For the CDS theme, we incorporated terms used in a systematic review of AI-CDS by Vasey et al.[29] We added search filters and terms related to components of the clinical decision making process (i.e. diagnosis, prognosis, treatment) to capture studies of AI-CDS tools that may not yet be defined as "clinical decision support" or a "clinical decision support system". The electronic search strategy for MEDLINE is provided in Table S2.[32, 33]

We explored grey literature using a hand search of Google Scholar (first 50 results) and two trial registries (ClinicalTrials.gov and World Health Organization International Clinical Trials Registry Platform), using the following search: *(artificial intelligence* or *machine learning)* and *emergency*, using the same cut-off dates as the database searches. We planned to hand search reference lists of included articles for additional studies but ultimately forewent this due to feasibility given the large number of studies included.

Selection of Sources of Evidence

We entered all titles into COVIDENCE for deduplication and screening.[34] Two independent reviewers (HK and one of CP, NB, AM, HL, or GG) performed title and abstract screening. Each of the title and abstract screeners is a Canadian emergency attending or resident physician. We piloted the title and abstract screening process on a random selection of 50 articles for each reviewer to ensure compatibility and functionality without major discrepancies between independent screeners. Based on this screening process, the planning authors (HK, CV, KY, and WM) clarified the inclusion criteria and published these updates on the Open Science Framework on February 3, 2024.[24] Two independent reviewers (HK and one of SM, LG, or MB) then performed full text screening. Each of the full text reviewers is a Canadian emergency attending physician with research experience in AI and ML applications for the ED. We piloted the title and abstract screening process on a random selection of 10 articles for each reviewer. Disagreements were resolved through consensus. Although it was planned a priori, no studies required arbitration by a third reviewer.

Eligibility Criteria

We included studies involving AI- or ML-based CDS tools applied to the care of individual patients in the ED. Given the anticipated variability of these terms in the literature, we created a priori definitions based on our collective experience and study objectives for the following: *definition of AI and ML, definition of CDS* and *timing of output*. Complete inclusion criteria, terminology and rationale can be found in Table S3. In brief: *AI* or *ML* was defined as a specific reference by the authors to the CDS tool as an example of

"artificial intelligence", "AI", "machine learning", "deep learning", "neural network", or "natural language processing." This was updated to include models that inferred machine learning methodology, based on the expertise of one or more reviewers (see Tables S3 and S4).

The term *clinical decision support* was defined as any model, algorithm or system that provided recommendations to a clinician regarding diagnostic, prognostic, or therapeutic decisions related to an individual ED patient's care. Studies of models that interpreted a specific diagnostic test (e.g. CT head) without consideration of other clinical variables (vital signs, physical exam findings, etc.) were excluded to ensure a focus on decisions that mirror an emergency physician's clinical decision-making process. We similarly excluded studies investigating administrative or operations outcomes such as automated documentation or predicting patient census.

The *timing of output* required the defined primary outcome of the study to be available and relevant to an emergency clinician's care. Studies incorporating irrelevant input data (e.g. animal models, simulated vignettes, or data only available at the end of or following the emergency assessment) were excluded, as were studies with outcomes deemed by the screeners not to be relevant to emergency care. The planning authors (HK, CV, KY, and WM) decided that predicting an individual patient's ED length of stay is overly dependent on non-clinical factors and would not necessarily influence clinical decision-making, therefore studies using this as a primary outcome were excluded.

In keeping with the objectives and scoping nature of this review, we included full-text articles, conference abstracts and trial protocols of primary research including

observational, interventional or implementation studies. Review articles, editorials, and commentaries were excluded.

Data Items and Charting Process

We created a data extraction table based on elements of the TRIPOD and DECIDE-AI guidelines,[14, 15] and ultimately included study characteristics (author, title, publication year, country, journal), stage of publication (full article, abstract only, or protocol), data collection (prospective versus retrospective), phase of development (in silico, silent, early live clinical, comparative prospective, or vigilance), patient identifiers (including presenting complaint and other specifiers such as geriatric or pediatric populations), AI model specifiers (including model type, use of supervised versus unsupervised prediction, use of natural language processing or large language models, or analysis of specific diagnostic tests), and primary outcome description and category (e.g. prognosis, diagnosis, etc.) Given the scoping nature of this review, we did not extract the results of the studies. We piloted the extraction table on 10 included studies and refined the extraction elements based on discussions between the planning authors (HK, CV, KY, and WM) before proceeding with full data extraction.

Synthesis of Results

In accordance with our a priori analysis plan, we have included a narrative synthesis of the findings, frequency counts of the extracted data items, a flow diagram according to PRISMA-ScR (see Figure 1),[22] a table of the electronic search strategies (see Table S2), a

description of the included studies (see Table S3), a figure mapping studies to the development framework proposed by the DECIDE-AI guidelines (see Figure 4).[14] Given the scoping nature of this review, we did not perform a critical appraisal of individual sources of evidence or risk of bias across studies.

3.2.6. Results

Search Results

We identified a total of 5,168 unique records through the electronic search, of which 721 were selected for full text screening. Three of these studies were found through review of included later-phase studies (i.e. these three studies were relevant preclinical development studies that were not found through our electronic search or initial grey literature searches). A total of 605 studies were included in our final analysis. See Figure 1 for further screening details. A full list of references is provided in Table S5.

Of the 116 excluded full texts, 58 (50.0%) used data that would not be available at the time of the ED clinician's assessment, including repeat vital signs when admitted to a hospital ward or the completed ED physician note. Thirteen (22.4%) studies either did not use AI or ML models or only used AI or ML to identify risk factors (i.e. not used as an independent AI-based CDS). Ten (8.5%) studies assessed outcomes that were deemed by reviewers not to be relevant to the ED clinician's care. Several studies did not meet the a priori definition of "clinical decision support"; eight (6.9%) of these used AI or ML to analyze a specific diagnostic test only (such as a chest x-ray or CT head), and five (4.3%) used AI or ML for administrative tasks.

Characteristics of Included Studies

A full description of the 605 included studies is provided in Table S6. We marked any unavailable data items as "unknown". In several studies, multiple primary outcomes and multiple development phases were included by the authors.

Overall, there has been a steadily increasing number of studies published in this field over time, with the publication rate increasing substantially since 2019 (see Figure 2). In the nine years from 2010 to 2018, only 102 studies were published, which was less than the annual publication rate in the three years from 2021 to 2023, where 119, 113, and 137 studies were published, respectively.

The largest proportion of single-nation studies were from the United States (223, 36.9%). Overall, there were 244 (40.3%) North American studies, 156 (25.8%) Asian studies, 118 (19.5%) European studies, 14 (2.3%) international studies (i.e. from multiple continents), 12 (2.0%) South American studies, nine (1.5%) Oceanian studies, and four (0.7%) African studies.

Only 124 (20.5%) studies used prospectively collected data, while 475 (78.5%) used retrospectively collected data.

Outcomes

The outcomes assessed by the studies varied greatly and were categorized as seen in Figure 3. A description and count of the most common applications within each of these categories is provided in Table 1. Overall, 270 (44.6%) studies assessed prognostic

outcomes including mortality, survival with favorable neurologic outcome, or critical care outcomes such as need for intensive care unit admission or mechanical ventilation. The outcome timings varied but were generally within one day to three months to be deemed as a relevant ED outcome. There were 193 (31.9%) studies assessing diagnostic outcomes such as stroke, sepsis, or acute coronary syndrome. Another 115 (19.0%) studies involved predicting patient disposition, including hospital admission or return to ED within several days.

Model Types

The vast majority of studies investigated supervised ML models (589, 97.4%), with three studies of unsupervised ML models and six including both supervised and unsupervised ML. "Supervised" ML refers to models that are trained on data labeled with the output of interested and subsequently observed to assess if they can accurately determine the output in a new unlabeled dataset.[5] "Unsupervised" ML involves grouping or "clustering" unlabeled data into similar groups that can then be investigated for relevance or utility in subsequent analyses.[5]

In total, 61 (10.1%) studies used natural language processing as part of the CDS, and ten (1.7%) used large language models. Another 31 (5.1%) studies included AI analysis of a specific diagnostic test, usually chest x-ray, electrocardiogram, or heart rate variability, as a component of the CDS in addition to other clinical variables such as patient demographics or vital signs.

Categorization by specific ML model type was challenging given the variety of names of models and lack of clear reporting of the methods by authors. We observed a range of supervised model types including variants of neural networks (including artificial neural networks, convolutional neural networks, and multilayer perceptron), random forests, *k*-nearest neighbour, naïve Bayes, and decision trees. In addition, there were commercial or established research models such as Lucia,[35] Ada,[36] and TREWS.[37]

Phases of Development

We categorized studies according to phases of development according to the pathway established by the DECIDE-AI guidelines, as seen in Figure 4.[14] This pathway begins with "preclinical development", which first involves using data to develop a model and then evaluating its performance "in silico" (i.e. retrospective computational assessment) followed by "offline" settings (i.e. deployed and running live in a clinical setting but the output is not made available for clinicians to use). Most of the included studies involved initial "in silico" preclinical evaluation (584, 96.5%), with fewer reaching offline validation (11, 1.8%). The third phase is "small-scale safety/utility" evaluation, where the tool's real-time output directly influences patient care, albeit with a focus on clinical utility, safety, and human factors, whereas the fourth phase of "large-scale safety/effectiveness" shifts the live evaluation focus to effectiveness in a randomized controlled trial. Few studies in our analysis reached these third (14, 2.3%) or fourth (5, 0.8%) phases of development. The final phase of development of "post-market surveillance" involves iterative evaluation of a deployed tool for the emergence of biases

that jeopardize its clinical utility. Only 6 studies (1.0%) correlated to this phase of development. Excluding full text publications, only two abstracts and nine protocols were beyond the initial development phase.

3.2.7. Discussion

Through this scoping review, we found a large volume of primary studies investigating AI-based clinical decision support tools for a variety of applications in the care of individual patients in the ED. We demonstrated that the rate of publication in the emergency medicine field has accelerated rapidly over the last three years, which has similarly been shown across other medical specialties.[9, 10] Furthermore, most of these studies remained at early phases of development with very few having been tested in live clinical settings. Therefore, the ability of AI-CDS tools to improve outcomes that are relevant to patients, providers, or healthcare systems remains limited.

Our study builds upon prior reviews of AI in emergency medicine, while providing more detailed analyses to better map this rapidly expanding and changing field. Hosseini et al's scoping review suggests an increasing publication rate, although does not provide clear values of this rate or years of publication.[20] It was also limited by the inclusion of terms related to "ethics" in the electronic search, which likely resulted in missing a significant proportion of the available evidence. The scoping review by Kirubarajan et al found a prevalence of retrospective datasets and a lack of interventional studies but, in addition to using limited search specifiers for the themes of AI and ML, was published in 2020 and

does not capture the recent proliferation of studies that we identified.[19] Boonstra et al's systematic review found that most AI applications were related to clinical decision support and to relieve overcrowding, but was similarly limited by narrow electronic search and inclusion criteria.[38] These three reviews included only 47, 150, and 37 studies,[19, 20, 38] respectively, and therefore provide only a limited scope of this expansive field. Furthermore, none categorized studies by phase of development or incorporated the DECIDE-AI pathway into their analysis. We are therefore able to contribute several novel and important findings.

Alignment with Desired Applications

We found the AI-CDS tools in our study align well with the desired applications for emergency clinicians established in prior studies. An international Delphi study of emergency physicians established consensus for several clinical applications of AI models as important,[39] including guidance for antibiotic choice,[40] analysis of cardiac monitors,[41] generating differential diagnoses,[42] and estimating risk of poor outcomes if discharged,[43] all of which were covered by the constituent studies in our review. Notably, there was a paucity of studies assessing medication guidance for pregnant or lactating patients, which was the third "most important" application from that Delphi study.[39] A separate survey of priorities for AI tools by Canadian emergency physicians found triage as the third highest priority when asking for a general category of "work type", and clinical prediction rules and monitoring vitals as the second and third highest priorities when using specific published examples, respectively.[44] Again, these applications were covered well

by the constituent studies in our review. Notably, both studies found that administrative applications of AI, such as assistance with documentation, computer tasks, and language translation, ranked as the highest priorities, above CDS tasks. Overall, it appears the CDS topics being studied in EM aligns well with identified physician priorities, but given the principal desire for administrative solutions, the latter may be a more pertinent initial undertaking for future work in this field.

Lack of Progression into Later Stages of Development

We found that most ED-based AI-CDS studies remain in the earliest phase of development, "preclinical development", with very few progressing into real-time testing in a clinical setting. Evidently, the "AI chasm" in emergency medicine remains wide. Our findings corroborate those of two systematic reviews of randomized controlled trials of AI in healthcare, both of which identified only a single prehospital study and no ED-based studies.[9, 10] Concerningly, one of these reviews found an increasing number of AI-CDS tools had achieved FDA approval without a randomized trial to show efficacy or potential generalizability.[9] In keeping with these findings, our study yielded only three protocols for randomized controlled trials,[45-47] and one protocol for a quasi-experimental study.[48] None of the six studies investigating efficacy of a deployed model in a live clinical setting, corresponding to the "post-market surveillance" phase of development, appeared to have had adequate prospective randomized studies prior to their actual or proposed implementation, as determined by the absence of such studies in our study or mentioned by their respective authors. This paucity of prospective trials demonstrating superiority or

non-inferiority compared with non-AI interventions or usual care should bring scrutiny to any claims that AI tools provide better alternatives to CDS tasks.

Establishing a Development Pathway

Demonstrating adequate performance of an AI-CDS tool in a live clinical setting should be necessary prior to its implementation. The DECIDE-AI guidelines provide an evidence-based development pathway to help facilitate the process through which these tools can be evaluated.[14] This is the first scoping review to map the emergency medicine literature to the DECIDE-AI phases of development, thereby providing a well-defined overview of the field. While it remains only a snapshot in time, our review further stratified the development phase outcomes by abstracts and protocols, suggesting that there remains a lack of anticipated later-phase studies in the near-future. We therefore recommend that emergency researchers initiate AI-CDS development with implementation in mind, following the DECIDE-AI guidelines and anticipating issues they may face during offline validation, trial and surveillance phases. During the completion of this study, the TRIPOD+AI guidelines were published, providing guidance for the reporting of initial development and preclinical evaluation of AI-CDS models.[49] We recommend these guidelines to any emergency researchers interested in undertaking an AI-CDS study.

Limitations

Our study has important limitations that should be considered when interpreting its results. First, we did not assess the performance or efficacy of the AI-CDS tools. In keeping

with scoping review methodological standards,[21, 22] our objective was to provide an overview of the types of applications in the emergency medicine field. Investigating model performance for a particular application would best be served by separate context-specific systematic reviews. Furthermore, we have already demonstrated that preclinical development studies of prognostic and diagnostic ML models tend to outperform usual care.[7] A cross-specialty review also found limited evidence showing improved diagnostic performance with ML-based CDS.[29]

Second, we did not assess the quality or risk of bias of the constituent studies. This assessment is best served by systematic, as opposed to scoping, reviews,[21, 22] and we have also previously commented on limitations of preclinical development of AI-CDS models.[7] Given reporting guidelines corresponding to particular phases of development were only published in the last three years, we felt assessing adherence to these standards would be of limited utility, but we suggest this as an important consideration for future quality assessments.

Finally, we did not investigate the reasons why the AI chasm between the preclinical development phase and later phases exists. Given the large number of studies included, we found it was unfeasible to extract this data and draw useful conclusions to this question within the confines of a scoping review. In many cases, the constituent studies did not provide adequate explanations for this question. We therefore felt this would best be answered through a separate qualitative study to identify challenges with conducting later stage development studies, which we intend to conduct.

3.2.8. Conclusion

In conclusion, we have thoroughly mapped the current landscape of AI-CDS tools in the ED, identifying a large number of studies involving a variety of clinical applications, patient populations, and AI models. Despite an increased rate of publication in recent years, few studies have advanced from preclinical development to later phases of clinical evaluation and implementation. The advent of an AI-specific reporting pathway provides a framework for researchers to bridge this gap, but a more granular understanding of the barriers and facilitators to implementing AI-CDS in the clinical context of the ED is still needed.

References

1. Kelen GD, Wolfe R, D'Onofrio G, et al. Emergency Department Crowding: The Canary in the Health Care System. *Catalyst non-issue content.* 2021;2(5).
2. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. *npj Digital Medicine.* 2020;3(1):17.
3. Wasylewicz ATM, Scheepers-Hoeks A. Clinical Decision Support Systems. In: Kubben P, Dumontier M, Dekker A, eds. *Fundamentals of Clinical Data Science.* Cham (CH): Springer Copyright 2019, The Author(s). 2019:153-169.
4. Sim I, Gorman P, Greenes RA, et al. Clinical Decision Support Systems for the Practice of Evidence-based Medicine. *Journal of the American Medical Informatics Association.* 2001;8(6):527-534.
5. Deo RC. Machine Learning in Medicine. *Circulation.* 2015;132(20):1920-1930.
6. Liu X, Faes L, Kale AU, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. *Lancet Digit Health.* 2019;1(6):e271-e297.
7. Kareemi H, Vaillancourt C, Rosenberg H, Fournier K, Yadav K. Machine Learning Versus Usual Care for Diagnostic and Prognostic Prediction in the Emergency Department: A Systematic Review. *Acad Emerg Med.* 2021;28(2):184-196.
8. Yin J, Ngiam KY, Teo HH. Role of Artificial Intelligence Applications in Real-Life Clinical Practice: Systematic Review. *J Med Internet Res.* 2021;23(4):e25759.

9. Plana D, Shung DL, Grimshaw AA, Saraf A, Sung JJY, Kann BH. Randomized Clinical Trials of Machine Learning Interventions in Health Care: A Systematic Review. *JAMA Netw Open.* 2022;5(9):e2233946.
10. Lam TYT, Cheung MFK, Munro YL, Lim KM, Shung D, Sung JJY. Randomized Controlled Trials of Artificial Intelligence in Clinical Practice: Systematic Review. *J Med Internet Res.* 2022;24(8):e37188.
11. Keane PA, Topol EJ. With an eye to AI and autonomous diagnosis. *npj Digital Medicine.* 2018;1(1):40.
12. Adam H, Balagopalan A, Alsentzer E, Christia F, Ghassemi M. Mitigating the impact of biased artificial intelligence in emergency decision-making. *Communications Medicine.* 2022;2(1):149.
13. Cadario R, Longoni C, Morewedge CK. Understanding, explaining, and utilizing medical artificial intelligence. *Nature Human Behaviour.* 2021;5(12):1636-1642.
14. Vasey B, Nagendran M, Campbell B, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *Nat Med.* 2022;28(5):924-933.
15. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med.* 2015;162(1):55-63.
16. Sounderajah V, Ashrafian H, Golub RM, et al. Developing a reporting guideline for artificial intelligence-centred diagnostic test accuracy studies: the STARD-AI protocol. *BMJ Open.* 2021;11(6):e047709.

17. Cruz Rivera S, Liu X, Chan AW, et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Lancet Digit Health.* 2020;2(10):e549-e560.
18. Liu X, Cruz Rivera S, Moher D, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nat Med.* 2020;26(9):1364-1374.
19. Kirubarajan A, Taher A, Khan S, Masood S. Artificial intelligence in emergency medicine: A scoping review. *J Am Coll Emerg Physicians Open.* 2020;1(6):1691-1702.
20. Masoumian Hosseini M, Masoumian Hosseini ST, Qayumi K, Ahmady S, Koohestani HR. The Aspects of Running Artificial Intelligence in Emergency Care; a Scoping Review. *Arch Acad Emerg Med.* 2023;11(1):e38.
21. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. *JBI Evid Synth.* 2020;18(10):2119-2126.
22. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann Intern Med.* 2018;169(7):467-473.
23. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev.* 2015;4(1):1.
24. Kareemi H, Yadav K, Michalowski W, Vaillancourt C. Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Scoping Review

Protocol. Open Science Framework, Nov. 2023; pages 1-8. https://doi.org/10.17605/OSF.IO/FDZ3Y

25. Gama F, Tyskbo D, Nygren J, Barlow J, Reed J, Svedberg P. Implementation Frameworks for Artificial Intelligence Translation Into Health Care Practice: Scoping Review. *J Med Internet Res.* 2022;24(1):e32215.

26. Zidaru T, Morrow EM, Stockley R. Ensuring patient and public involvement in the transition to AI-assisted mental health care: A systematic scoping review and agenda for design justice. *Health Expect.* 2021;24(4):1072-1124.

27. Cartolovni A, Tomicic A, Lazic Mosler E. Ethical, legal, and social considerations of AI-based medical decision-support tools: A scoping review. *Int J Med Inform.* 2022;161:104738.

28. Murphy K, Di Ruggiero E, Upshur R, et al. Artificial intelligence for good health: a scoping review of the ethics literature. *BMC Med Ethics.* 2021;22(1):14.

29. Vasey B, Ursprung S, Beddoe B, et al. Association of Clinician Diagnostic Performance With Machine Learning-Based Decision Support Systems: A Systematic Review. *JAMA Netw Open.* 2021;4(3):e211276.

30. Campbell S, Kung J. Filter to Retrieve Studies Related to Artificial Intelligence from the OVID MEDLINE Database. *Geoffrey & Robyn Sperber Health Sciences Library, University of Alberta* 2023.

31. Campbell S. A Filter to Retrieve Studies related to Emergency Departments from the OVID MEDLINE Database. *Geoffrey & Robyn Sperber Health Sciences Library, University of Alberta.* 2021.

32. Chojecki D, Kung J. A Filter to Retrieve Diagnostic Studies from the OVID MEDLINE Database. . *Geoffrey & Robyn Sperber Health Sciences Library, University of Alberta* 2023.

33. Wilczynski NL, Haynes RB. Developing optimal search strategies for detecting clinically sound prognostic studies in MEDLINE: an analytic survey. *BMC Med.* 2004;2:23.

34. *Covidence systematic review software* [computer program]. Melbourne, Australia2023.

35. Schwab K, Nguyen D, Ungab G, et al. Artificial intelligence MacHIne learning for the detection and treatment of atrial fibrillation guidelines in the emergency department setting (AIM HIGHER): Assessing a machine learning clinical decision support tool to detect and treat non-valvular atrial fibrillation in the emergency department. *J Am Coll Emerg Physicians Open.* 2021;2(4):e12534.

36. Faqar-Uz-Zaman SF, Anantharajah L, Baumartz P, et al. The Diagnostic Efficacy of an App-based Diagnostic Health Care Application in the Emergency Room: eRadaR-Trial. A prospective, Double-blinded, Observational Study. *Ann Surg.* 2022;276(5):935-942.

37. Saria S, Henry K, Soleimani H, et al. 1429: LEAD TIME AND ACCURACY OF TREWS, A MACHINE LEARNING-BASED SEPSIS ALERT. *Critical Care Medicine.* 2022;50(1):717.

38. Boonstra A, Laven M. Influence of artificial intelligence on the work design of emergency department clinicians a systematic literature review. *BMC Health*

Serv Res. 2022;22(1):669. Published 2022 May 18. doi:10.1186/s12913-022-08070-7

39. Li H, Hayward J, Aguilar LS, Franc JM. Desired clinical applications of artificial intelligence in emergency medicine: A Delphi study. *Am J Emerg Med.* 2024;79:217-220.

40. Corbin CK, Sung L, Chattopadhyay A, et al. Personalized antibiograms for machine learning driven antibiotic selection. *Commun Med (Lond).* 2022;2:38.

41. Sundrani S, Chen J, Jin BT, Abad ZSH, Rajpurkar P, Kim D. Predicting patient decompensation from continuous physiologic monitoring in the emergency department. *NPJ Digital Medicine.* 2023;6(1):60.

42. Fraser H, Crossland D, Bacher I, Ranney M, Madsen T, Hilliard R. Comparison of Diagnostic and Triage Accuracy of Ada Health and WebMD Symptom Checkers, ChatGPT, and Physicians for Patients in an Emergency Department: Clinical Data Analysis Study. *JMIR Mhealth Uhealth.* 2023;11:e49995.

43. Grant L, Joo P, Nemnom MJ, Thiruganasambandamoorthy V. Machine learning versus traditional methods for the development of risk stratification scores: a case study using original Canadian Syncope Risk Score data. *Intern Emerg Med.* 2022;17(4):1145-1153.

44. Eastwood KW, May R, Andreou P, Abidi S, Abidi SSR, Loubani OM. Needs and expectations for artificial intelligence in emergency medicine according to Canadian physicians. *BMC Health Serv Res.* 2023;23(1):798.

45. van der Zaag AY, Bhagirath SC, Boerman AW, et al. Appropriate use of blood cultures in the emergency department through machine learning (ABC): study

protocol for a randomised controlled non-inferiority trial. *BMJ Open.* 2024;14(5):e084053.

46. Kijpaisalratana N. Evaluating a sepsis prediction machine learning algorithm. *Evaluating a sepsis prediction machine learning algorithm in Thai population.* 2023.

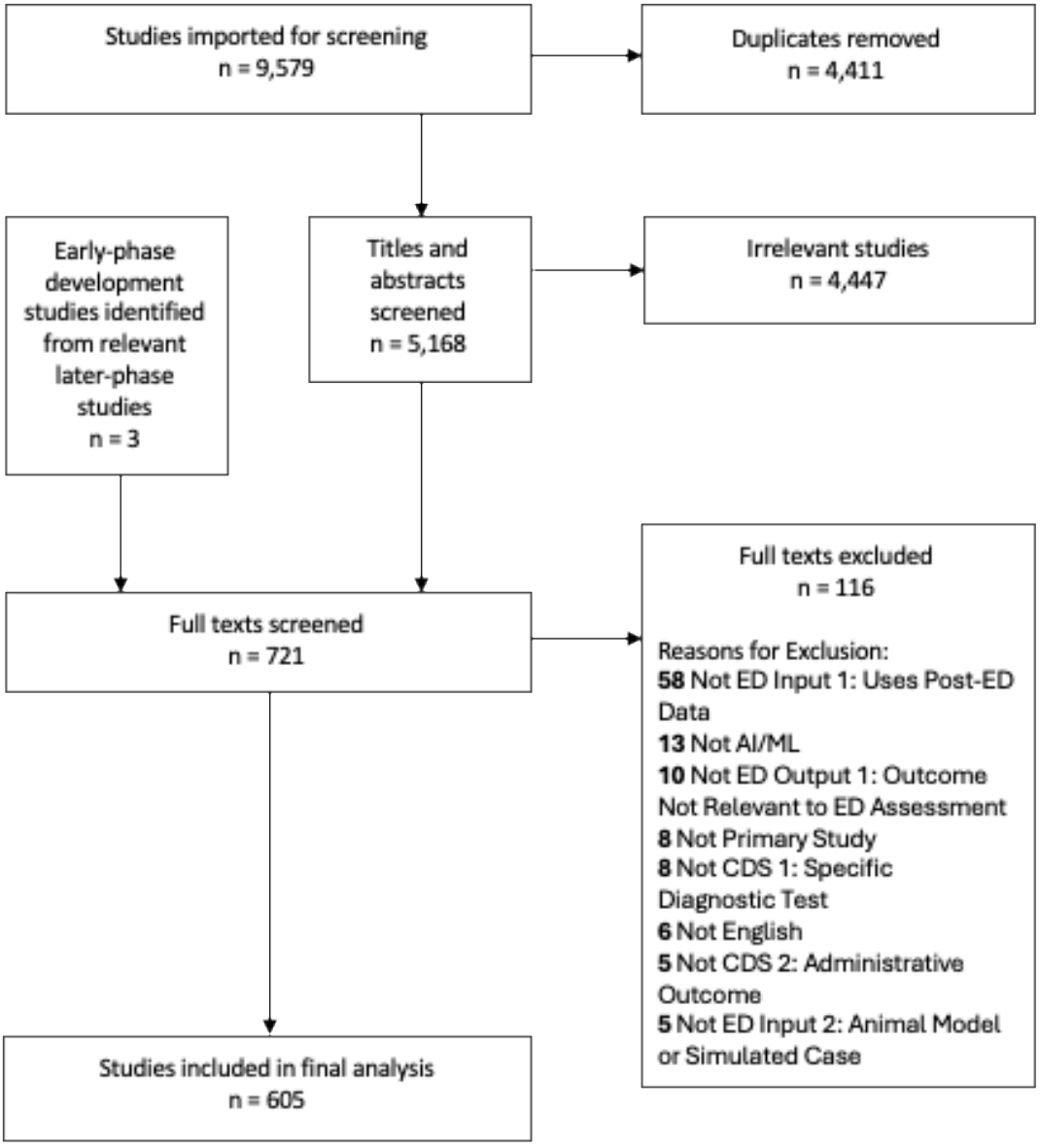
47. van Dam PM. Machine Learning for Risk Stratification in the Emergency Department (MARS-ED). *Machine Learning for Risk Stratification in the Emergency Department: a Pilot Clinical Trial.* 2022.

48. Hekman DJ, Cochran AL, Maru AP, et al. Effectiveness of an Emergency Department-Based Machine Learning Clinical Decision Support Tool to Prevent Outpatient Falls Among Older Adults: Protocol for a Quasi-Experimental Study. *JMIR research protocols.* 2023;12(101599504):e48128.

49. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *Bmj.* 2024;385:e078378.

Tables and Figures

Figure 3.2-1. PRISMA-ScR Flow Diagram



AI: artificial intelligence; CDS: clinical decision support; ED: emergency department; ML: machine learning; PRISMA-ScR: Preferred Reporting Items for Systematic reviews and Meta-Analyses for Scoping Reviews

Figure 3.2-2. Number of Included Studies by Publication Year since 2010

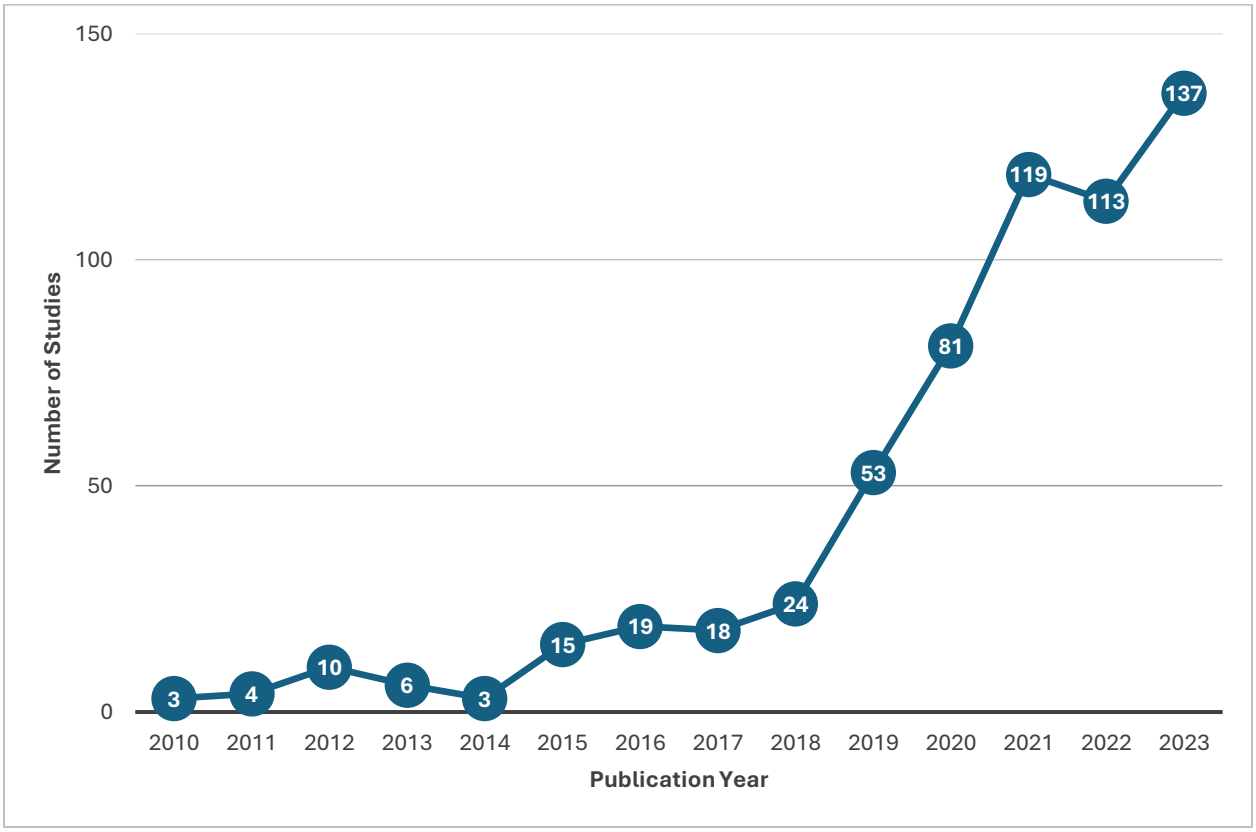
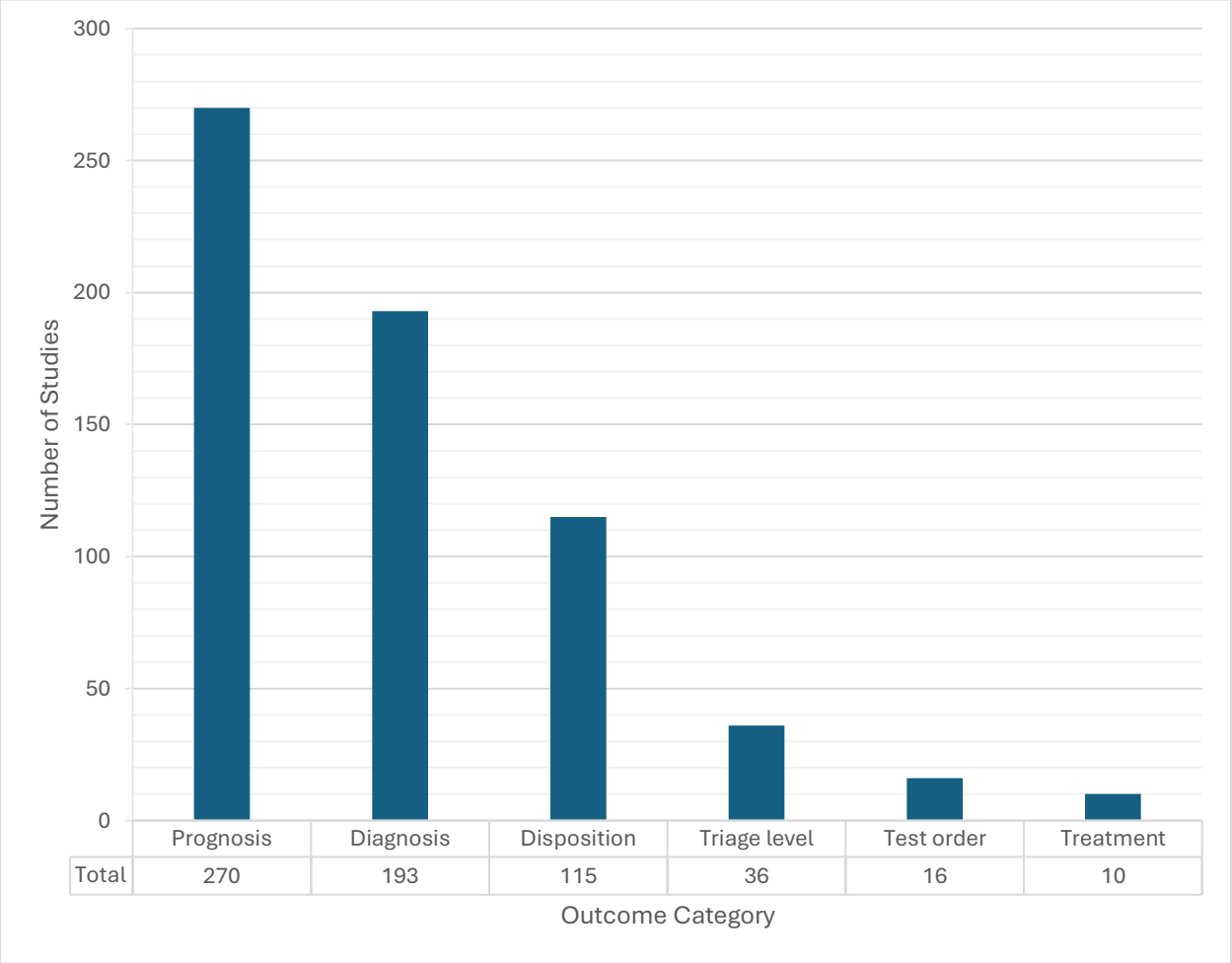
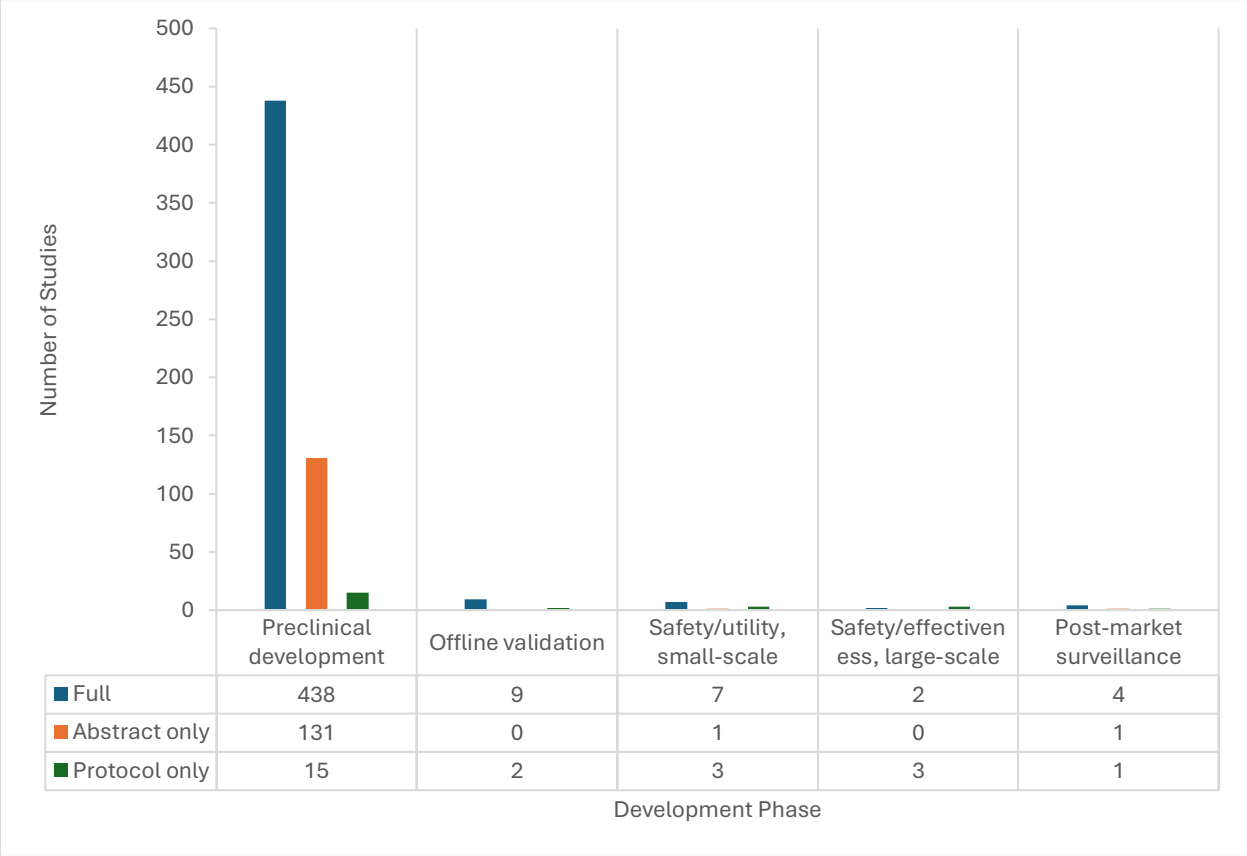


Figure 3.2-3. Categorization of Outcomes in the Included Studies



N.b. Several studies did not specify a single primary outcome or included multiple primary outcomes that could be categorized into multiple categories; therefore, there were more total categories (n = 640) than total included studies (n = 606).

Figure 3.2-4. Phases of Development of the Included Studies, as Defined by the DECIDE-AI Guidelines, Stratified by Publication Stage



	Preclinical development	Offline validation	Safety/utility, small-scale	Safety/effectiveness, large-scale	Post-market surveillance
Full	438	9	7	2	4
Abstract only	131	0	1	0	1
Protocol only	15	2	3	3	1

N.b. Several studies included data collection and analyses from multiple phases of development; therefore, there were more total phases of development (n = 617) than total included studies (n = 606).

Table 3.2-1. Outcome Categories and Most Commonly Assessed Subcategories (Greater than Ten Studies) in the Included Studies

Category	Number of Studies Assessing (n, % of all included studies)
Prognosis	270 (44.6%)
Mortality	112 (18.5%)
Intensive care unit admission	50 (8.3%)
Critical care outcome*	50 (8.3%)
Cardiac arrest	29 (4.8%)
Survival with favorable neurologic outcome	14 (2.3%)
Diagnosis	193 (31.9%)
Sepsis	27 (4.5%)
Myocardial Infarction or Acute Coronary Syndrome	26 (4.3%)
COVID-19 infection	21 (3.5%)
Stroke or Transient Ischemic Attack	18 (3.0%)
Disposition	115 (19.0%)
Hospital admission	97 (16.0%)
Return to ED	18 (3.0%)
Triage Level	36 (6.0%)
Test Order	16 (2.6%)
Treatment	10 (1.7%)

*"Critical care outcome" was a heterogeneously defined composite outcome that included a combination of ICU admission, in-hospital mortality, need for clinical intervention, etc.

Broad outcome categories are designated by highlighted rows. Subcategories are designated by unhighlighted rows.

Appendix

Appendix 3.2 – Table 1. Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) Checklist

SECTION	ITEM	PRISMA-ScR CHECKLIST ITEM	REPORTED ON PAGE #
TITLE			
Title	1	Identify the report as a scoping review.	Title Page
ABSTRACT			
Structured summary	2	Provide a structured summary that includes (as applicable): background, objectives, eligibility criteria, sources of evidence, charting methods, results, and conclusions that relate to the review questions and objectives.	1
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known. Explain why the review questions/objectives lend themselves to a scoping review approach.	3-4
Objectives	4	Provide an explicit statement of the questions and objectives being addressed with reference to their key elements (e.g., population or participants, concepts, and context) or other relevant key elements used to conceptualize the review questions and/or objectives.	5
METHODS			
Protocol and registration	5	Indicate whether a review protocol exists; state if and where it can be accessed (e.g., a Web address); and if available, provide registration information, including the registration number.	6
Eligibility criteria	6	Specify characteristics of the sources of evidence used as eligibility criteria (e.g., years considered, language, and publication status), and provide a rationale.	8-9
Information sources*	7	Describe all information sources in the search (e.g., databases with dates of coverage and contact with authors to identify additional sources), as well as the date the most recent search was executed.	6-8
Search	8	Present the full electronic search strategy for at least 1 database, including any limits used, such that it could be repeated.	Appendix Table II
Selection of sources of evidence†	9	State the process for selecting sources of evidence (i.e., screening and eligibility) included in the scoping review.	7-8
Data charting process‡	10	Describe the methods of charting data from the included sources of evidence (e.g., calibrated forms or forms that have been tested by the team before their use, and whether data charting was done independently or in duplicate) and any processes for obtaining and confirming data from investigators.	9-10

SECTION	ITEM	PRISMA-ScR CHECKLIST ITEM	REPORTED ON PAGE #
Data items	11	List and define all variables for which data were sought and any assumptions and simplifications made.	9-10
Critical appraisal of individual sources of evidence§	12	If done, provide a rationale for conducting a critical appraisal of included sources of evidence; describe the methods used and how this information was used in any data synthesis (if appropriate).	N/a
Synthesis of results	13	Describe the methods of handling and summarizing the data that were charted.	10
RESULTS			
Selection of sources of evidence	14	Give numbers of sources of evidence screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally using a flow diagram.	10-11, Figure 1
Characteristics of sources of evidence	15	For each source of evidence, present characteristics for which data were charted and provide the citations.	11-12, Appendix Table VI
Critical appraisal within sources of evidence	16	If done, present data on critical appraisal of included sources of evidence (see item 12).	N/a
Results of individual sources of evidence	17	For each included source of evidence, present the relevant data that were charted that relate to the review questions and objectives.	12-14
Synthesis of results	18	Summarize and/or present the charting results as they relate to the review questions and objectives.	Figures 2-4, Table 1
DISCUSSION			
Summary of evidence	19	Summarize the main results (including an overview of concepts, themes, and types of evidence available), link to the review questions and objectives, and consider the relevance to key groups.	14-18
Limitations	20	Discuss the limitations of the scoping review process.	18-19
Conclusions	21	Provide a general interpretation of the results with respect to the review questions and objectives, as well as potential implications and/or next steps.	19
FUNDING			
Funding	22	Describe sources of funding for the included sources of evidence, as well as sources of funding for the scoping review. Describe the role of the funders of the scoping review.	Title Page

JBI = Joanna Briggs Institute; PRISMA-ScR = Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews.

* Where *sources of evidence* (see second footnote) are compiled from, such as bibliographic databases, social media platforms, and Web sites.

† A more inclusive/heterogeneous term used to account for the different types of evidence or data sources (e.g., quantitative and/or qualitative research, expert opinion, and policy documents) that may be eligible in a scoping review as opposed to only studies. This is not to be confused with *information sources* (see first footnote).

‡ The frameworks by Arksey and O'Malley (6) and Levac and colleagues (7) and the JBI guidance (4, 5) refer to the process of data extraction in a scoping review as data charting.

§ The process of systematically examining research evidence to assess its validity, results, and relevance before using it to inform a decision. This term is used for items 12 and 19 instead of "risk of bias" (which is more applicable to systematic reviews of interventions) to include and acknowledge the various sources of evidence that may be used in a scoping review (e.g., quantitative and/or qualitative research, expert opinion, and policy document).

From: Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMAScR): Checklist and Explanation. Ann Intern Med. 2018;169:467–473. doi: 10.7326/M18-0850.

Appendix 3.2 – Table 2. Electronic Search Strategy for Ovid MEDLINE

#	Searches	Results
1	exp Artificial Intelligence/ or (((machin* or artific* or comput* or robot* or automat*) adj3 (intelligen* or reasoning)) or ((assist* or augment* or autonomous) adj1 intelligen*) or ((machin* or deep or transfer or hierarchical) adj2 learning) or "classification algorithm*" or "computer heuristic*" or "feature detection" or "generative pre-trained transformer" or "language learning model*" or "large language model*" or "learning algorithm*" or (Markov adj3 model*) or "natural language process*" or "nearest neighbor*" or "neural network*" or "random forest" or "representation learning" or "support vector machine*" or "transfer learning" or "Bing chat" or ChatGPT* or "Chat GPT" or "Google* Bard" or "IBM Watson" or "Microsoft* Bing" or OpenAI or "Open AI" or PathAI or "Path AI").ti,ab,kf.	370663
2	Emergency Treatment/ or Emergency Medical Services/ or Emergencies/ or Emergency Service, Hospital/ or Evidence-Based Emergency Medicine/ or Trauma Centers/ or Triage/ or (((emergenc* or ED) adj1 (room* or accident or ward or wards or unit or units or department* or physician* or doctor* or nurs* or treatment* or visit*)) or (triage adj2 (centre or centres or center or centers or department? or unit or units)) or (urgent adj2 (care or healthcare or health care)) or critical care or (trauma adj2 (centre or centres or center or centers or department? or unit or units)) or ("accident and emergency" or "accident & emergency") or ((emergency or non-emergency or nonemergency or urgent or non-urgent or nonurgent or semi-urgent or semiurgent) adj2 patient?)).ti,ab,kf.	359822
3	Decision Making, Computer-Assisted/ or Diagnosis, Computer-Assisted/ or Therapy, Computer-Assisted/ or Drug Therapy, Computer-Assisted/ or exp Decision Support Systems, Clinical/ or Clinical Decision Rules/ or Clinical Decision-Making/ or (CDSS* or CCDSS* or ((decision* or decid*) adj2 (making* or make* or support)) or ((computer* or digit*) adj1 (assist* or aid or support* or help*))).ti,ab,kf.	345451
4	Delayed Diagnosis/ or Diagnosis, Differential/ or Diagnostic Errors/ or "Diagnostic Techniques and Procedures"/ or Early Diagnosis/ or Overdiagnosis/ or "Sensitivity and Specificity"/ or (diagnos* or sensitivit* or specificit* or receiver operat* or roc).ti,ab,kf.	4665529
5	Prognosis/ or Treatment Outcome/ or Incidence/ or Patient Outcome Assessment/ or (prognos* or predict*).ti,ab,kf.	4124035
6	or/3-5	7974024
7	and/1-2,6	2526
8	limit 7 to yr="2010 -Current"	2321

Appendix 3.2 – Table 3. Complete Inclusion Criteria and Relevant Terminology with Subsequent Post-Screening Updates

Inclusion Criterion	Initial – December 11, 2023	Update – February 3, 2024
1. Study focus	Studies involving AI- or ML-based CDS tools applied in the ED.	See definitions below.
Definition of AI or ML	Authors should refer to the CDS tool as an example of "artificial intelligence", "AI", "machine learning", "deep learning", "neural network", or "natural language processing". We have encountered variable definitions of these terms in the literature based our previous work,[1] and we will therefore accept a broad definition in keeping with our objective and other reviews.[2] Any disputes about inclusion related to the *definition of AI or ML* will be resolved in consensus with an author with expertise in AI and ML methodology (WM).	Definition of AI or ML: In some instances, relevant studies did not refer to their models as "AI", "artificial intelligence", "ML", or "machine learning". The primary author (HK), who has experience with AI research and co-screened every study, included these studies for full text review, which he will co-review with a second ED clinician with AI research experience.
Definition of CDS	By "clinical decision support", we are referring to any model, algorithm or system that provides recommendations to a clinician regarding diagnostic, prognostic, or therapeutic decisions related to an individual ED patient's care.	Definition of CDS – 1: We will further refine "clinical decision support" to those that involve multiple clinical variables including history of presenting event, physical exam findings, vital signs, investigations in the emergency department, and past medical history, including that accessible through the electronic medical record. We will exclude studies of models that interpret specific diagnostic tests or image analysis only, such as radiographs (e.g. x-ray or CT or MRI), electrocardiogram, specific laboratory investigations, retinal fundoscopic images, and photographs of rashes, unless the AI model additionally consider multiple other clinical variables. • Rationale: While some of these theoretically could be useful during the ED assessment (e.g. identification of abnormalities on chest radiograph), the focus of this scoping review is on supporting complex clinical decisions that are in the purview of the emergency clinician, as opposed to expertise in interpreting specific diagnostic tests that are within the scope of respective consultants such as radiologists, ophthalmologists, dermatologists, etc.

		Definition of CDS – 2: Outcomes related to operations and logistics, specifically ED length of stay or wait times, were excluded. • Rationale: The planning authors (HK, CV, KY, WM) did not feel that prediction of these outcomes would necessarily influence clinical decision making for an individual patient (e.g. knowing that a patient will wait 2 hours versus 10 hours would not make a clinician order different tests for that patient; alternatively, knowing a patient has high likelihood of clinical decompensation or ICU admission may affect diagnostic testing or expediting consultation).
Timing of Output	The recommendation of the CDS tool must be applicable to an individual ED patient's care during their index ED visit (i.e. from registration or triage to discharge or admission to hospital). This may include outcomes related to future return to ED or future hospitalization if the output of the recommendations affects decisions and are (or conceivably could be) available at the time of index ED visit (e.g. decision to discharge, extended duration of therapy or diagnostic testing after discharge). The input data may include prehospital data if it is used to inform a clinical decision at the time of the ED visit. If the study involves a mixed cohort (e.g. ED and intensive care unit), then the inclusion criteria regarding *Timing of Output* must be met to warrant inclusion.	Timing of Output – 1: In cases where the timing of the primary outcome is of uncertain relevance to the emergency clinician (e.g. risk of subsequent suicide attempt within 1 year after index ED visit), a discussion and decision was made between the co-reviewers, all of whom are emergency medicine physicians working at large urban centres in Canada. If there remained uncertainty after discussion, the case would have corroborated with one of the emergency clinicians from the planning authors (CV or KY), however there were no cases in which this was required. Timing of Output – 2: In cases where the input data used by the AI models was unclear (e.g. predicting in-hospital mortality using "admission data", which in some cases referred to data from the ED assessment and in other cases referred to data including repeat vitals/bloodwork upon admission to an inpatient ward), studies were only included if they explicitly stated the timing of the input data or specific variables considered were from the ED assessment, or that the model output would inform ED management. • In keeping with this, studies using large trauma or ICU databases were excluded unless the latter criteria were met. • If the co-reviewers felt there was a high likelihood of relevance to ED studies, studies were included for further clarification on full text review. • Similarly, studies using natural language processing (NLP) to mine unstructured data retrospectively from

		emergency physician charts for the purpose of epidemiological surveillance or improving diagnostic coding accuracy were excluded. Timing of Output – 3: Studies that used animal models or simulated cases were excluded.
Phase of Development	The CDS tool may be in any phase of development as defined by the DECIDE-AI guidelines,[3] including preclinical/“in silico”, silent/“shadow” evaluation, small-scale “live” clinical evaluation, large-scale comparative prospective evaluation, or post-market vigilance.	No change.
2. Study type	Full-text articles, conference abstracts and trial protocols of primary research including observational, interventional or implementation studies. In keeping with objectives and scoping nature of this review, we want to capture any relevant CDS tools that are in development or deployed. In our analysis, we will distinguish articles with full text publications from those with only conference abstracts and trial protocols.	No change.
3. Publication date	AI technologies are advancing rapidly. Other scoping reviews of AI technologies have used a variety of year cut-offs including 2000,[4] 2010,[5] 2014,[6, 7] and 2015.[8] A systematic review of AI-CDS systems that subsequently informed the derivation of the DECIDE-AI guidelines used a cut-off date of 2010.[9] For feasibility and relevance, we therefore selected 2010 as the cut-off for our search.	No change.
4. Study Language	For feasibility, we will only include English studies.	No change.

Appendix 3.2 – Table 4. Final Prioritized List of Exclusion Criteria Used for Excluded Studies.

- Not CDS – 1: Specific Diagnostic Test/Finding Only
- Not CDS – 2: Operations/Logistic Only (e.g. LOS, wait time)
- Not ED Input – 1*: Considers Post-ED Data (e.g. Inpatient, Post-surgical)
    - *If the study involves a mixed cohort (e.g. ED and ICU), then the inclusion criteria regarding Timing of Output (i.e. outcome can be made at time of ED assessment) must be met to warrant inclusion.
- Not ED Input – 2: Animal Model, Simulated Cases/Vignettes, etc.
- Not ED Output: Outcome Not Relevant to ED Assessment
- Not AI/ML
- Not Primary Study (e.g. Review, Editorial, Commentary)
- Not English

For brevity and space limitations, Table S5 (15 pages) and Table S6 (78 pages) have been excluded from this thesis report but can be found at https://doi.org/10.1111/acem.15099.

Appendix 3.2 – References

1. Kareemi H, Vaillancourt C, Rosenberg H, Fournier K, Yadav K. Machine Learning Versus Usual Care for Diagnostic and Prognostic Prediction in the Emergency Department: A Systematic Review. Acad Emerg Med. 2021;28(2):184-96.
2. Lam TYT, Cheung MFK, Munro YL, Lim KM, Shung D, Sung JJY. Randomized Controlled Trials of Artificial Intelligence in Clinical Practice: Systematic Review. J Med Internet Res. 2022;24(8):e37188.
3. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med. 2022;28(5):924-33.
4. Gama F, Tyskbo D, Nygren J, Barlow J, Reed J, Svedberg P. Implementation Frameworks for Artificial Intelligence Translation Into Health Care Practice: Scoping Review. J Med Internet Res. 2022;24(1):e32215.
5. Zidaru T, Morrow EM, Stockley R. Ensuring patient and public involvement in the transition to AI-assisted mental health care: A systematic scoping review and agenda for design justice. Health Expect. 2021;24(4):1072-124.
6. Cartolovni A, Tomicic A, Lazic Mosler E. Ethical, legal, and social considerations of AI-based medical decision-support tools: A scoping review. Int J Med Inform. 2022;161:104738.
7. Masoumian Hosseini M, Masoumian Hosseini ST, Qayumi K, Ahmady S, Koohestani HR. The Aspects of Running Artificial Intelligence in Emergency Care; a Scoping Review. Arch Acad Emerg Med. 2023;11(1):e38.

8. Murphy K, Di Ruggiero E, Upshur R, Willison DJ, Malhotra N, Cai JC, et al. Artificial intelligence for good health: a scoping review of the ethics literature. BMC Med Ethics. 2021;22(1):14.
9. Vasey B, Ursprung S, Beddoe B, Taylor EH, Marlow N, Bilbro N, et al. Association of Clinician Diagnostic Performance With Machine Learning-Based Decision Support Systems: A Systematic Review. JAMA Netw Open. 2021;4(3):e211276.

Chapter 4. Qualitative Analysis of Expert Interviews

This chapter describes a qualitative study involving interviews with researchers who have experience developing and implementing AI-CDS tools in the ED setting. We conducted an analysis based on grounded theory methodology to understand why the "AI chasm" identified in our scoping review exists. In so doing, we address our second and third thesis objectives: "Construct a theory-driven explanation for the limited clinical translation of AI-CDS tools in the ED" and "Describe the barriers to development and implementation of AI-CDS tools in the ED and identify facilitators that may help surmount them."

Herein, we include the manuscript, "Barriers and Facilitators to Developing Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Qualitative Study" that has been submitted for publication and is currently under review by *Canadian Journal of Emergency Medicine*.

4.1. Barriers and Facilitators to Developing Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Qualitative Study

4.1.1. Preface

Contributions of Co-Authors: HK, KY, WM, and CV conceptualized and planned the study. HK conducted the interviews. HK and AC performed coding and analysis of the interview transcripts. HK created the tables and figures and wrote the manuscript. All authors edited, reviewed, and approved the final manuscript.

Ethics Approval: This study was approved by the Ottawa Health Science Network-Research Ethics Board on May 29, 2024 (protocol ID: 20240338-01H).

Thesis-Related Appendices: Appendices 1-6.

Citation: Kareemi H, Colak A, Yadav K, Michalowski W, Vaillancourt C. Barriers and Facilitators to Developing and Implementing Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Qualitative Study. Submitted to CJEM. 2025 June (Under Review).

4.1.2. Title Page

Title
Barriers and Facilitators to Developing and Implementing Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Qualitative Study

Running Title
AI-CDS in ED: Qualitative Study

Authors
Hashim Kareemi, MD[1,2]
Alex Colak[3]
Krishan Yadav, MD MSc[2,4,5]
Wojtek Michalowski, ME PhD[6]
Christian Vaillancourt, MD MSc[2,4,5]

Affiliations
1. Department of Emergency Medicine, University of British Columbia, Vancouver, BC, Canada
2. School of Epidemiology and Public Health, University of Ottawa, Ottawa, ON, Canada
3. Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada
4. Department of Emergency Medicine, University of Ottawa, Ottawa, ON, Canada
5. Acute Care Research Program, Ottawa Hospital Research Institute, ON, Canada
6. Telfer School of Management, University of Ottawa, Ottawa, ON, Canada

Corresponding Author
Hashim Kareemi

Address: Emergency Medicine, 910 W 10th Ave, Vancouver, BC, V5Z 1M9

Financial Support
Hashim Kareemi's Master's of Science thesis work, of which this study is a component, is supported by a Canadian Graduate Scholarship - Master's.

Conflicts of Interest
- HK:
    - Received Canadian Institutes of Health Research's Canada Graduate Scholarship – Master's to fund this work.
    - Received consulting fees with University of British Columbia Digital Emergency Medicine Unit unrelated to this work.
    - Received research award from Vancouver Acute Emergency Physicians Association unrelated to this work.

- AC: None
- KY:
    - Received grants from Canadian Institutes of Health Research (CIHR), TOHAMO (The Ottawa Hospital Academic Medical Organization) - both unrelated to this work.
    - Received salary support via Physician Services Incorporated (PSI) Graham Farquharson Early Career Knowledge Translation Fellowship – unrelated to this work.
- WM: None
- CV:
    - Received grants from Canadian Institutes of Health Research's Planning and Dissemination Grant, Canadian Institutes of Health Research's Operating Grant, and The Ottawa Hospital Academic Medical Organization for separate work.

Author Contributions

HK, KY, WM, and CV conceptualized and planned the study.
HK conducted the interviews.
HK and AC performed coding and analysis of the interview transcripts.
HK wrote the manuscript.
All authors edited, reviewed, and approved this manuscript.

4.1.3. Abstract

Purpose

We sought to understand how barriers at the development phase of (AI)-based clinical decision support tools may lead to challenges with implementation, and what facilitators may help surmount them.

Methods

We conducted a qualitative study involving semi-structured interviews with researchers who have expertise developing and implementing AI-clinical decision support tools for use in the emergency department (ED). We used purposive and snowball sampling to identify participants. We used platform-based AI transcription and anonymized transcripts manually. Using grounded theory framework, two coders iteratively analyzed transcripts in three stages (initial, focused, and theoretical) to identify barriers, facilitators, and themes. We adhered to SRQR and COREQ guidelines.

Results

We achieved data saturation after ten interviews conducted between October 15, 2024 and March 22, 2025. Participants ranged across a variety of medical and academic professions. We identified eight themes pertaining to developing and implementing AI-clinical decision support in the ED, in descending frequency: team capacity; data infrastructure; defining the clinical problem and solution; research, ethics, and regulatory approval; legal and liability; model building and performance; time; and cost. We identified “engaging multiple

healthcare end-users" and "sharing resources with other departments" as the highest yield facilitators.

Conclusion

Successful implementation of AI-clinical decision support tools in the ED necessitates a clinician- and patient-defined problem, robust data infrastructure, and a diversely experienced research team that can navigate challenges with regulatory approval processes, legal and liability concerns, and significant cost and time requirements. Anticipating known barriers and enacting facilitators early in the development process will increase the likelihood of successful implementation.

4.1.4. Introduction

There is rapidly growing interest in using artificial intelligence (AI) and machine learning models to support high-stakes, time-limited clinical decision making in the emergency department (ED), particularly as crowding, staffing issues, and prolonged wait times further complicate these decisions.[1] Clinical decision support, referring to tools that augment decisions made by clinicians related to the diagnosis, prognosis, or treatment of patients, have been identified as a priority AI application by emergency physicians.[2, 3] Our team conducted a scoping review that found that although a large number of AI-clinical decision support tools are being developed for use in the ED, very few have been implemented into clinical workflows or tested in interventional trials.[4] This so-called "AI chasm"[5] has been identified in other medical fields beyond emergency

medicine,[6-8] but a thorough understanding as to why it exists remain elusive, as do solutions to bridge over it. One possible factor is the prevalence of methodological flaws at the model development stage, which may preclude adoption into clinical practice by limiting model functionality, introducing unacceptable biases, or worsening performance in subsequent validation or clinical trials.[9-11]

The development pathway for AI-clinical decision support models proposed by the Developmental and Exploratory Clinical Investigations of Decision Support Systems Driven by AI (DECIDE-AI) authors would ideally address and mitigate known methodological issues.[12] Mirroring those of pharmaceutical and surgical innovations, this pathway combines guidelines for specific stages of model development including "offline" validation of models for prediction (TRIPOD+AI) and diagnostic test accuracy (STARD-AI),[13, 14] early "live" evaluations (DECIDE-AI),[12] production of trial protocols (SPIRIT-AI),[15] and reporting findings for large-scale effectiveness trials (CONSORT-AI).[16] However, given the component guidelines were only published in the last 5 years and are limited to reporting standards, expert insight to supplement understanding and provide practical context to their use is still critically needed. The experience of those who have developed and implemented AI-clinical decision support models, particularly in the unique clinical environment of the ED, may provide insight into the benefits of well-designed models and the shortcomings of poorly designed ones. Previous qualitative research in this area has mostly been limited to interviews and surveys of end-users regarding their needs, wants, and apprehensions regarding AI in emergency medicine in general,[17] or towards

specific applications.[18-20] None have focused on expert insight into how these models were derived and deployed.

Primarily, we sought to construct a theory-driven explanation for the limited clinical translation of AI-CDS tools in the ED. Secondarily, we sought to describe the barriers to development and implementation of AI-CDS tools in the ED and identify facilitators that may help surmount them. We achieved these objectives by interviewing researchers who have studied AI-clinical decision support tools in the ED, with the intent of providing guidance to other researchers and improving the rate of successful implementation of AI-clinical decision support tools for the benefit of ED patients, providers, and health systems as a whole.

4.1.5. Methods

Qualitative Approach

We employed a qualitative design using one-on-one semi-structured interviews with experienced researchers in the field of AI applied to emergency medicine. We have reported our findings according to the SRQR (Standards for Reporting Qualitative Research)[21] and COREQ (Consolidated Criteria for Reporting Qualitative Research)[22] checklists to ensure methodological rigor and transparency (see Online Resource Tables 1 and 2). We applied the grounded theory framework, which uses inductive analysis to develop a “theory grounded in data”.[23] We felt this would best suit our objective due to its coding flexibility, invocation of higher level themes, and building of a unified theory,[23] as opposed to more rigid and deductive frameworks like qualitative content analysis.[24]

Interviewer Characteristics

The interviews were conducted solely by the primary author (HK), who identifies as a male of South Asian heritage. He is a Canadian emergency physician in an urban academic centre. This work is part of his thesis for a Master's of Science in Epidemiology. HK had experience with qualitative methodology through medical training, graduate and undergraduate coursework, and prior publications.

HK had prior professional relationships with two of the participants as co-members of the Canadian Association of Emergency Physicians' Artificial Intelligence Special Interest Group, a community of emergency AI practitioners. The participants were aware HK was conducting this work for his thesis and that he had previously published studies on AI-clinical decision support in the ED. None of the participants had any shared financial interests with HK or the other authors. None of the participants were directly involved in the formation of the research questions or methodological approach of this work.

Participant Selection

We included participants that were clinician (e.g. physician, nurse) or non-clinician (e.g. data scientist, epidemiologist, engineer, statistician) researchers that were named authors on at least two completed or ongoing projects involving AI-based tools used for clinical decision support. These projects, whether individually or in total, needed to satisfy the following three criteria: A) Model development (corresponding to either the "preclinical development" or "offline validation" phases of the DECIDE-AI development pathway);[12]

B) Model implementation (corresponding to either the “safety/utility, small scale”, “Safety/effectiveness, large-scale”, or “Post-market surveillance” phases);[12] or C) Clinical decision support related to the individual care of ED patients. We judged these criteria based on web-based review of relevant publications and projects online prior to approaching participants. The rationale for our criteria was to identify researchers who could speak to barriers and facilitators regarding both development and implementation of AI-clinical decision support tools within the ED context.

We used purposive sampling, identifying participants from relevant research publications found in our previous or ongoing work in this field, as well as through the Canadian Association of Emergency Physician’s Artificial Intelligence Special Interest Group. We also used snowball recruitment, by which participants were asked to recommend other potential participants. We approached eligible persons via email for voluntary participation in the study. Follow-up and final requests were sent at 2- and 6-week intervals from the initial email.

As recommended by Francis et al regarding data saturation,[25] we targeted our initial analysis sample as 10, based on comparable studies.[18-20, 26] We defined saturation *a priori* as, at the point of completing 10 interviews, encountering 2 consecutive participants that did not contribute additional material to advancing the grounded theory, as judged by consensus of the authors of this study.

Ethical Considerations

This study was approved by the [blinded] Research Ethics Board on May 29, 2024 (protocol ID: 20240338-01H). We attained verbal consent from participants at the start of each interview. We maintained confidentiality of participants throughout the study.

Setting

HK conducted the interviews via video call using the Microsoft Teams platform (versions 2024-2025).[27] Only a single participant and the interviewer were present to ensure a personalized and consistent interview process.

Data Collection and Processing

We developed the interview guide based on our research question and prior work in the field. The guide is provided in Online Resource Table 3. The interviewer used additional probing questions to allow elaboration on pertinent points. We piloted the interview on one of the authors, WM, who met the inclusion criteria, to ensure logical progression and successful coding processes. The data from the pilot interview were excluded from the main analysis.

We recorded the interviews, transcribed them using the built-in Microsoft Teams transcription function, and stored the audio files and transcriptions in an encrypted password-protected file. Prior to storage, HK read each line of the transcripts and removed any identifying information. Participants were not provided copies of their transcripts and did not provide feedback on the results.

HK and AC performed parallel coding using NVivo 14 (QSR International, Doncaster, Australia)[28] on de-identified transcripts in three iterative stages: initial, focused, and theoretical.[23] The initial stage compared "data with data", using concise codes to efficiently parse through the data.[23] After the first five interviews were completed, HK and AC performed focused coding, first in isolation and then collaboratively, by comparing and categorizing their codes from the initial stage into "more directed, selective, and conceptual" codes.[29] After ten interviews were completed, HK and AC collaboratively performed "theoretical" coding, which involved integrating focused codes into an underlying grounded theory, as well as to assess for data saturation.[23]

Techniques to Enhance Trustworthiness

We ensured trustworthiness of results by having coders perform the initial and focused coding stages independently and then comparing their codes collaboratively. All authors reviewed the coding tree and agreed to the higher-level codes that were derived. Three of the five authors ([blinded]) are clinician-investigators and practicing emergency physicians, and have experience with and confirmed the veracity of the identified codes, particularly those pertaining to clinical and research processes. One of the authors ([blinded]) has extensive research experience with clinical AI tools and confirmed the veracity of codes pertaining to data and model building.[30, 31]

4.1.6. Results

Participants

Of the 43 persons meeting inclusion criteria who were contacted for this study, ten (23.3%) agreed to be interviewed. We conducted interviews from October 15, 2024 to March 22, 2025. Interviews lasted 30-45 minutes. As described in Table 1, the majority (9/10; 90%) of participants identified as male, with only one non-binary and no female participants. Participants had a variety of professional backgrounds and time in their respective fields, with seven (70%) being practicing physicians. Three had direct private industry affiliation, in that they were officers in or part-owners of a technology corporation, whereas six had collaborated with private industry in their projects but did not have direct affiliation, and one had no affiliation nor collaboration with private industry. All participants were based in North America (seven from Canada, three from USA).

Coding Process

The coders categorized codes into “barriers”, “facilitators”, and higher-level “themes”, which served as the basis for subsequent analyses. These are provided in the coding tree in Online Resource Table 4 along with representative quotations.

Data Saturation

Once the grounded theory was established, coders discussed whether data from the last two interviews had contributed significantly to the development of the grounded theory. Both coders agreed that the grounded theory had been established after the

seventh interview, and that subsequent interviews provided further context to the theory but did not change its content or the established themes. Thus, data saturation was achieved after ten interviews.

Data Synthesis: Barriers

Each code represents a "barrier" or "facilitator" to developing an AI-clinical decision support tool, implementing it into an ED clinical workflow, or both. For the purposes of the grounded theory and our primary objective, we first focused on the "barriers", which are summarized along with frequency counts in Table 2.

The most commonly mentioned barriers pertained to "research team capacity" and the need for team members with a vast array of expertise (including clinical, technical, and design aspects), as well as leadership and collaborative abilities. This theme represented 36/119 (30.3%) of all barriers mentioned. There was also frequent discussion of "data infrastructure", representing 33/119 (27.7%) of mentioned barriers. This included the need for a clinical data warehouse and effective, efficient, and exhaustive data collection and analysis procedures. Several participants discussed challenges with "defining the clinical problem and solution", specifically the need for emphasizing clinician and patient priorities and de-emphasizing research questions secondary to profit or availability of data.

To a lesser extent, participants raised issues pertaining to "research, ethics, and regulatory approvals" and "legal and liability", with reference to unclear policies and potential liability related to data sharing, privacy, and medical error. "Model building and performance" issues involving technical aspects of the data context or models themselves

were also raised. Finally, resource concerns regarding “time” and “cost” were also identified by several participants.

Data Synthesis: Facilitators

The identified facilitators are also included in the coding tree in Online Resource Table 4. These were often raised by participants in direct resolution to a particular barrier. We performed a post-hoc analysis to identify other potential barriers each facilitator could address. This is summarized in the concept map shown in Figure 1, in which the central squares represent the themes (with size corresponding to relative frequency) connected to individual barriers represented by the small rings. The blue circles along the outside represent facilitators, with arrows signifying connected barriers.

In Figure 2, we display frequency counts of the facilitators according to the number of barriers they could remove, thereby identifying which facilitators provide the highest potential benefit. Overall, being able to “engage multiple healthcare end-users as part of your research team” and “share resources with other departments within your institution” were found to address the most barriers.

Grounded Theory

In order to develop the grounded theory, we considered the content and relative frequencies of the identified themes and codes in context of our research question, “How do barriers at the development stage of AI-clinical decision support tools cause

subsequent challenges with implementation into clinical workflows and practice in the ED?" We summarize our theory as follows:

The limited clinical translation of AI-based clinical decision support tools in emergency medicine stems from insufficient consideration of implementation during preclinical development. This may reflect either a lack of awareness of the complexity of real-world barriers or a deliberate focus on theoretical performance at the expense of practical feasibility. Proactively identifying barriers and leveraging facilitators early in the development process is essential to enhance the likelihood of successful clinical adoption.

4.1.7. Discussion

Interpretation of Findings

Our qualitative analysis of interviews with experienced researchers suggests that it is critical for implementation to be considered from the moment the idea for an AI-clinical decision support tool is conceived. The inability to anticipate the clinical, technical, and regulatory context of the tool or to adequately outfit a research team to achieve this may result in a protracted timeline or preclude implementation altogether. Our findings help explain why only 3.5% of AI-clinical decision support tools are tested or deployed in a live ED setting.[4] Below, we discuss some of the reasons why this gap persists in clinical AI research and how it can be bridged, in light of the findings of our study and context of the current literature.

Comparison to Previous Studies

Much of the qualitative research done in this field has focused on barriers to implementation of AI-clinical decision support tools. Leenen similarly identified such barriers as the lack of clinician-informed solutions, multidisciplinary collaboration, technical expertise, and clear guidance in the climate of rapidly evolving regulatory requirements.[32] Hummelsberger categorized major implementation challenges as regulatory (data security and regulatory approval processes), logistical (mainly related to funding and a lack of information technology professionals), and technical (lack of data quantity and quality).[33] Notably, these studies were broad in scope and lacked insight into the nuances of the ED setting. Other qualitative studies have focused on the perspectives of emergency clinicians. One study of a heart failure risk stratification tool found that providers primarily wanted support with disposition decisions and required the tool to be streamlined into existing clinical workflows.[19] Similarly, a study of a diagnostic tool intended to identify red flag conditions found a major barrier to adoption was the lack of integration with their electronic medical record, instead requiring a paper printout that was handed to physicians prior to seeing the patient.[34] Other studies also emphasized the need to focus on local clinical workflows and building a user-friendly interface to enhance successful implementation.[18, 26]

The barriers outlined by these studies and our own emphasize the need to promote implementation science as an essential and early component of AI-clinical decision support research. Implementation science aims to improve the adoption and reach of interventions in healthcare settings, but is usually only considered after an intervention has demonstrated effectiveness.[35] There are increasing calls for creating “implementation

science centers" that provide a pathway for evaluating the clinical effectiveness of AI-clinical decision support tools, ensuring ethical development, integration into local workflows, and continuous evaluation after deployment.[36] Efforts are underway to create a national network of AI "assurance labs" in the USA that can carry out standardized tests for AI-clinical decision support tools.[37]

Strengths and Limitations

Our study's main strength lies in our participants' ability to speak to both development and implementation phases of AI-clinical decision support research, allowing for a nuanced understanding of the barriers to translating these tools. Additionally, the participants have experience with the clinical context of the ED, which provides specific perspective that has been missing from prior studies. We also followed the highest methodological standards for qualitative research by adhering to both the SRQR and COREQ guidelines.

There are limitations to consider with our work. First, as with all grounded theory and inductive reasoning, there is the possibility of the authors' own biases and experiences influencing the results. We mitigated this by pairing a coder (HK) with experience in the field of AI-clinical decision support tools with another coder (AC) who does not, in addition to perspective from other experienced emergency medicine and AI researchers to achieve a balanced inductive analysis. Second, we did not record audio or video of the interviews for feasibility reasons, so only the interviewer (HK) could assess the interviewees' body language and tone. HK could implicitly incorporate some of this non-verbal

communication while coding, but AC could not, therefore this may have introduced bias between coders in the initial analysis. Third, our study suffered from lack of diversity, given we had nine males, one non-binary person, and no female participants. Additionally, all participants were North American and primarily English-speaking. Despite contacting several researchers in Europe and Asia and proposing interview times that we deemed would be amenable to people in those time zones, none were available.

Clinical Implications

Our work outlines the many barriers clinicians need to consider when implementing an AI-clinical decision support tool. Beyond developing a tool themselves, clinicians will need to assemble a team including information technology experts and clinical end-users, in addition to leveraging or improving available data infrastructure, and monitoring for data drift to ensure long term efficacy. By utilizing the solutions we identified, particularly the sharing of resources with other departments in their institution, it is our hope that clinicians will be able to successfully deploy AI-clinical decision support tools.

Research Implications

Compared with "traditional" medical research, the nascency of AI-clinical decision support research presents challenges with research ethics and regulatory approval processes. Several of our interviewed experts expressed difficulties with attaining institutional approval from research ethics boards that have limited experience with AI-clinical decision support. Fortunately, institutions are moving towards separate,

streamlined approval processes for AI research. [38, 39] On a government regulatory level, both Health Canada and the U.S. Food and Drug Administration have recently created guidance in the forms of their respective "Pre-Market Guidance for Machine Learning-Enabled Medical Devices" and "Artificial Intelligence and Machine Learning Software as a Medical Device Action Plan".[40, 41] Whether the advent of these guidelines represents a new facilitator by clarifying the approval process or a new barrier by complicating and adding to the previous approval requirements remains unknown.

4.1.8. Conclusion

Through interviews with experienced researchers, we concluded that successful implementation of AI-clinical decision support tools in the ED necessitates a clear clinician- and patient-defined problem, robust data infrastructure, and a diversely experienced research team that can navigate challenges with regulatory approval processes, legal and liability concerns, and significant cost and time requirements. Anticipating known barriers and enacting facilitators early in the development process will increase the likelihood of successful implementation.

References

1. Kelen GD, Wolfe R, D'Onofrio G, Mills AM, Diercks D, Stern SA, et al. Emergency Department Crowding: The Canary in the Health Care System. Catalyst non-issue content. 2021;2(5).

2. Eastwood KW, May R, Andreou P, Abidi S, Abidi SSR, Loubani OM. Needs and expectations for artificial intelligence in emergency medicine according to Canadian physicians. BMC Health Serv Res. 2023;23(1):798.
3. Li H, Hayward J, Aguilar LS, Franc JM. Desired clinical applications of artificial intelligence in emergency medicine: A Delphi study. Am J Emerg Med. 2024;79:217-20.
4. Kareemi H, Yadav K, Price C, Bobrovitz N, Meehan A, Li H, et al. Artificial intelligence-based clinical decision support in the emergency department: A scoping review. Acad Emerg Med. 2025.
5. Keane PA, Topol EJ. With an eye to AI and autonomous diagnosis. npj Digital Medicine. 2018;1(1):40.
6. Yin J, Ngiam KY, Teo HH. Role of Artificial Intelligence Applications in Real-Life Clinical Practice: Systematic Review. J Med Internet Res. 2021;23(4):e25759.
7. Plana D, Shung DL, Grimshaw AA, Saraf A, Sung JJY, Kann BH. Randomized Clinical Trials of Machine Learning Interventions in Health Care: A Systematic Review. JAMA Netw Open. 2022;5(9):e2233946.
8. Lam TYT, Cheung MFK, Munro YL, Lim KM, Shung D, Sung JJY. Randomized Controlled Trials of Artificial Intelligence in Clinical Practice: Systematic Review. J Med Internet Res. 2022;24(8):e37188.
9. Kareemi H, Vaillancourt C, Rosenberg H, Fournier K, Yadav K. Machine Learning Versus Usual Care for Diagnostic and Prognostic Prediction in the Emergency Department: A Systematic Review. Acad Emerg Med. 2021;28(2):184-96.

10. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. Lancet Digit Health. 2019;1(6):e271-e97.
11. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. J Clin Epidemiol. 2019;110:12-22.
12. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med. 2022;28(5):924-33.
13. Sounderajah V, Ashrafian H, Golub RM, Shetty S, Fauw JD, Hooft L, et al. Developing a reporting guideline for artificial intelligence-centred diagnostic test accuracy studies: the STARD-AI protocol. BMJ Open. 2021;11(6):e047709.
14. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. Bmj. 2024;385:e078378.
15. Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ, Spirit AI, et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. Lancet Digit Health. 2020;2(10):e549-e60.
16. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, Spirit AI, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med. 2020;26(9):1364-74.

17. Stewart J, Freeman S, Eroglu E, Dumitrascu N, Lu J, Goudie A, et al. Attitudes towards artificial intelligence in emergency medicine. Emerg Med Australas. 2023.
18. Fujimori R, Liu K, Soeno S, Naraba H, Ogura K, Hara K, et al. Acceptance, Barriers, and Facilitators to Implementing Artificial Intelligence-Based Decision Support Systems in Emergency Departments: Quantitative and Qualitative Evaluation. JMIR Form Res. 2022;6(6):e36501.
19. Sax DR, Sturmer LR, Mark DG, Rana JS, Reed ME. Barriers and Opportunities Regarding Implementation of a Machine Learning-Based Acute Heart Failure Risk Stratification Tool in the Emergency Department. Diagnostics (Basel). 2022;12(10).
20. Jordan M, Hauser J, Cota S, Li H, Wolf L. The Impact of Cultural Embeddedness on the Implementation of an Artificial Intelligence Program at Triage: A Qualitative Study. J Transcult Nurs. 2023;34(1):32-9.
21. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med. 2014;89(9):1245-51.
22. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care. 2007;19(6):349-57.
23. Thornberg R, Charmaz K. Grounded theory and theoretical coding. The SAGE handbook of qualitative data analysis. 2014;5(2014):153-69.

24. Schreier M. Qualitative Content Analysis. The SAGE handbook of qualitative data analysis. 2014;5:170-83.
25. Francis JJ, Johnston M, Robertson C, Glidewell L, Entwistle V, Eccles MP, et al. What is an adequate sample size? Operationalising data saturation for theory-based interview studies. Psychol Health. 2010;25(10):1229-45.
26. Sandhu S, Lin AL, Brajer N, Sperling J, Ratliff W, Bedoya AD, et al. Integrating a Machine Learning System Into Clinical Workflows: Qualitative Study. J Med Internet Res. 2020;22(11):e22421.
27. Microsoft Teams. Microsoft Corporation. Version 25093.2105.3614.8220 ed. New York City: Microsoft Corporation.
28. NVivo pro qualitative analysis software. 12.6.0.959 ed: QSR International Pty Ltd; 2018.
29. Charmaz K. Constructing grounded theory: A practical guide through qualitative analysis: sage; 2006.
30. Farion KJ, Wilk S, Michalowski W, O'Sullivan D, Sayyad-Shirabad J. Comparing predictions made by a prediction model, clinical score, and physicians: pediatric asthma exacerbations in the emergency department. Applied clinical informatics. 2013;4(3):376-91.
31. Klement W, Wilk S, Michalowski W, Farion KJ, Osmond MH, Verter V. Predicting the need for CT imaging in children with minor head injury using an ensemble of Naive Bayes classifiers. Artificial Intelligence in Medicine. 2012;54(3):163-70.

32. Zullig LL, Drake C, Check DK, Brunkert T, Deschodt M, Olson M, et al. Embedding implementation science in the research pipeline. Translational Behavioral Medicine. 2023;14(2):73-9.

33. Leenen JPL, Hiemstra P, Ten Hoeve MM, Jansen ACJ, van Dijk JD, Vendel B, et al. Exploring the complex nature of implementation of Artificial intelligence in clinical practice: an interview study with healthcare professionals, researchers and Policy and Governance Experts. PLOS Digital Health. 2025;4(5):e0000847.

34. Hummelsberger P, Koch TK, Rauh S, Dorn J, Lermer E, Raue M, et al. Insights on the Current State and Future Outlook of AI in Health Care: Expert Interview Study. Jmir ai. 2023;2:e47353.

35. Petitgand C, Motulsky A, Denis JL, Régis C. Investigating the Barriers to Physician Adoption of an Artificial Intelligence- Based Decision Support System in Emergency Care: An Interpretative Qualitative Study. Stud Health Technol Inform. 2020;270:1001-5.

36. Longhurst CA, Singh K, Chopra A, Atreja A, Brownstein JS. A Call for Artificial Intelligence Implementation Science Centers to Evaluate Clinical Effectiveness. NEJM AI. 2024;1(8):AIp2400223.

37. Shah NH, Halamka JD, Saria S, Pencina M, Tazbaz T, Tripathi M, et al. A Nationwide Network of Health AI Assurance Laboratories. JAMA. 2024;331(3):245-9.

38. The Hospital for Sick Children Research Ethics B. SickKids Research Ethics Handbook. Toronto, ON: The Hospital for Sick Children; 2022.

39. Board UCaWsRE. AI Research Ethics Application Checklist v1.2. Vancouver, BC: Provincial Health Services Authority; 2022.
40. Health Canada. Pre-market Guidance for Machine Learning-Enabled Medical Devices. Ottawa, ON: Health Canada; 2025.
41. U.S. Food and Drug Administration. Artificial Intelligence and Machine Learning in Software as a Medical Device. Silver Spring, MD; 2025.

Tables and Figures

Table 4.1-1. Characteristics of the Interviewed Experts

Participant Number	Date of Interview	Profession	Gender	Country	Time in Field (years)
1	October 15, 2024	Pharmacist; Data Scientist	Male	Canada	>20
2	October 21, 2024	Emergency Physician	Male	Canada	5
3	November 13, 2024	Pediatric Emergency Physician; Computer Scientist	Male	Canada	8
4	December 12, 2024	Emergency Physician	Male	Canada	8
5	December 18, 2024	Critical Care Physician; Thoracic Surgeon	Male	Canada	>20
6	February 2, 2025	Emergency Physician	Male	USA	12
7	February 14, 2025	Emergency Physician; Informatician	Male	Canada	30
8	February 14, 2025	Registered Nurse; Informatician	Male	USA	16
9	March 11, 2025	Human Factors Engineer	Non-binary	USA	6
10	March 22, 2025	Emergency Physician	Male	Canada	11

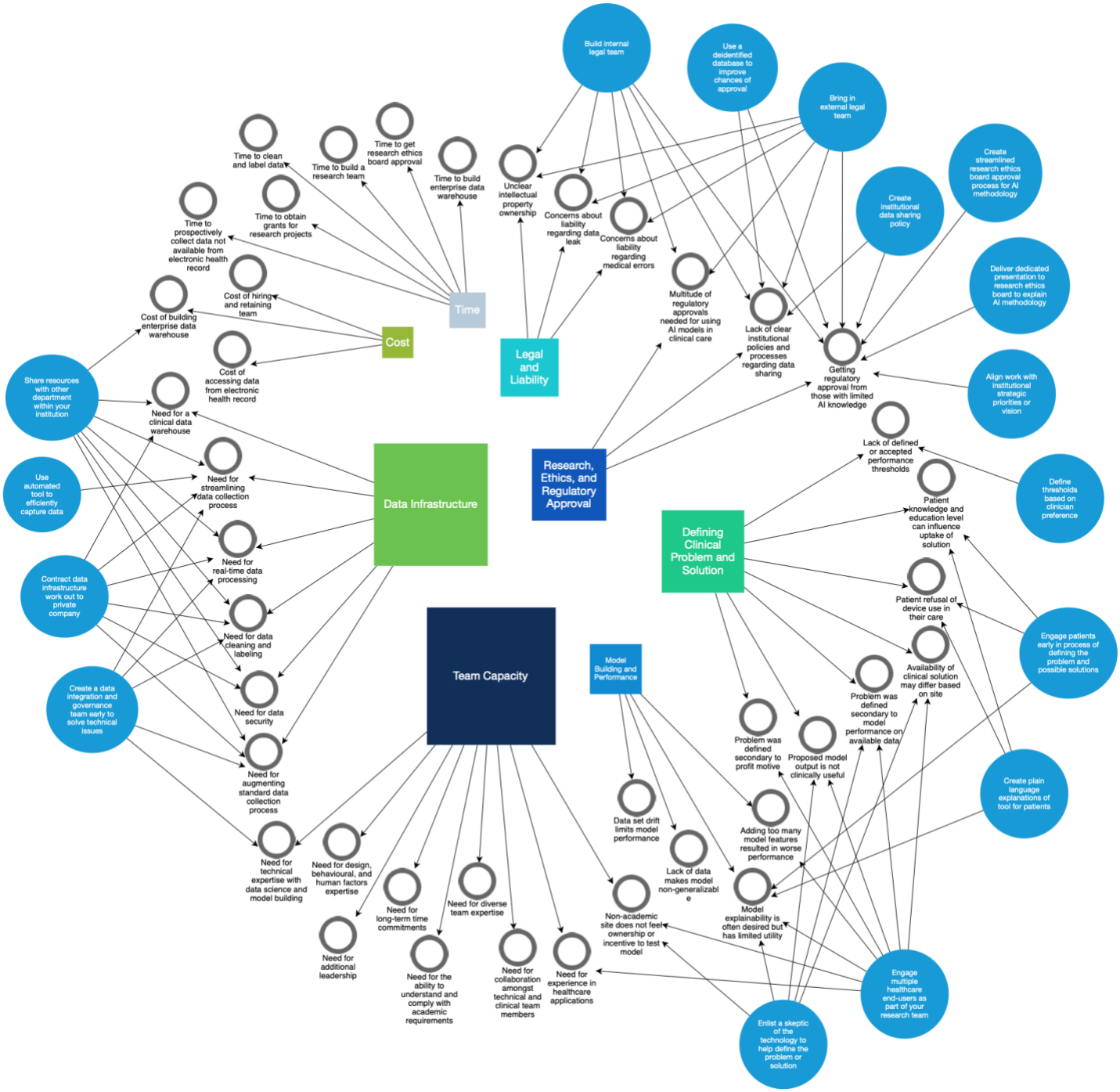
Table 4.1-2. Themes (shaded rows) and Codes of Barriers in Descending Frequency of Mentions by the Interviewed Experts

Code	Phase	Count
Team Capacity		36
Need for diverse team expertise	Both	9
Need for technical expertise with data science and model building	Both	6
Need for design, behavioural, and human factors expertise	Both	5
Need for long-term time commitments	Both	5
Need for experience in healthcare applications	Both	4
Need for collaboration amongst technical and clinical team members	Both	3
Non-academic site does not feel ownership or incentive to test model	Implementation	2
Need for additional leadership	Both	1
Need for the ability to understand and comply with academic requirements	Both	1
Data Infrastructure		33
Need for a clinical data warehouse	Both	8
Need for augmenting standard data collection process	Both	7
Need for data security	Both	6
Need for data cleaning and labeling	Both	5
Need for real-time data processing	Both	4
Need for streamlining data collection process	Both	3
Defining Clinical Problem and Solution		17
Proposed model output is not clinically useful	Both	6
Problem was defined secondary to model performance on available data	Development	4
Lack of defined or accepted performance thresholds	Development	2
Problem was defined secondary to profit motive	Development	2
Availability of clinical solution may differ based on site	Implementation	1
Patient knowledge and education level can influence uptake of solution	Implementation	1
Patient refusal of device use in their care	Implementation	1
Research, Ethics, and Regulatory Approval		10
Getting regulatory approval from those with limited AI knowledge	Both	7
Lack of clear institutional policies and processes regarding data sharing	Both	2
Multitude of regulatory approvals needed for using AI models in clinical care	Implementation	1
Legal and Liability		7
Concerns about liability regarding medical errors	Implementation	4
Concerns about liability regarding data leak	Both	2
Unclear intellectual property ownership	Both	1
Model Building and Performance		6
Data set drift limits model performance	Both	3

Adding too many model features resulted in worse performance	Development	1
Lack of data makes model non-generalizable	Both	1
Model explainability is often desired but has limited utility	Both	1
Time		6
Time to build a research team	Both	1
Time to build enterprise data warehouse	Both	1
Time to clean and label data	Development	1
Time to get research ethics board approval	Both	1
Time to obtain grants for research projects	Both	1
Time to prospectively collect data that is not available from electronic health record	Development	1
Cost		4
Cost of hiring and retaining team	Both	2
Cost of accessing data from electronic health record	Both	1
Cost of building enterprise data warehouse	Both	1

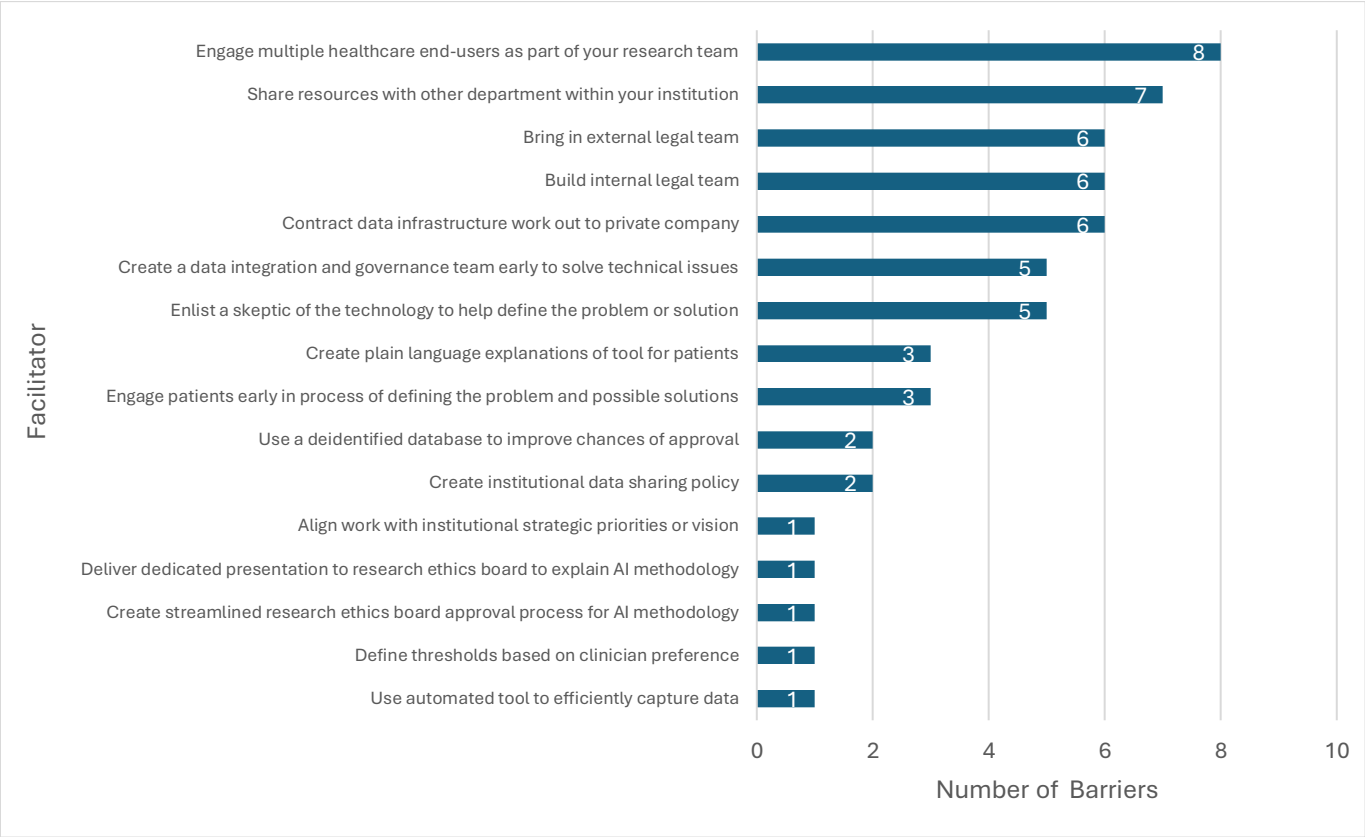
AI: artificial intelligence

Figure 4.1-1. Concept Map of Themes, Barriers, and Facilitators Identified from Interview Analysis



Squares: themes (size proportional to frequency count)
Small rings: barriers
Blue circles: facilitators

Figure 4.1-2. Frequency Counts of Facilitators by the Number of Barriers they can Potentially Remove



AI: artificial intelligence

Appendix

Appendix 4.1 – Table 1. Standards for Reporting Qualitative Research (SRQR) Checklist

	Page no(s).
Title and abstract	
Title - Concise description of the nature and topic of the study Identifying the study as qualitative or indicating the approach (e.g., ethnography, grounded theory) or data collection methods (e.g., interview, focus group) is recommended	Title Page
Abstract - Summary of key elements of the study using the abstract format of the intended publication; typically includes background, purpose, methods, results, and conclusions	1-2
Introduction	
Problem formulation - Description and significance of the problem/phenomenon studied; review of relevant theory and empirical work; problem statement	3-4
Purpose or research question - Purpose of the study and specific objectives or questions	4-5
Methods	
Qualitative approach and research paradigm - Qualitative approach (e.g., ethnography, grounded theory, case study, phenomenology, narrative research) and guiding theory if appropriate; identifying the research paradigm (e.g., postpositivist, constructivist/ interpretivist) is also recommended; rationale**	5
Researcher characteristics and reflexivity - Researchers' characteristics that may influence the research, including personal attributes, qualifications/experience, relationship with participants, assumptions, and/or presuppositions; potential or actual interaction between researchers' characteristics and the research questions, approach, methods, results, and/or transferability	5-6
Context - Setting/site and salient contextual factors; rationale**	7
Sampling strategy - How and why research participants, documents, or events were selected; criteria for deciding when no further sampling was necessary (e.g., sampling saturation); rationale**	6-7
Ethical issues pertaining to human subjects - Documentation of approval by an appropriate ethics review board and participant consent, or explanation for lack thereof; other confidentiality and data security issues	7
Data collection methods - Types of data collected; details of data collection procedures including (as appropriate) start and stop dates of data collection and analysis, iterative process, triangulation of sources/methods, and modification of procedures in response to evolving study findings; rationale**	8
Data collection instruments and technologies - Description of instruments (e.g., interview guides, questionnaires) and devices (e.g., audio recorders) used for data collection; if/how the instrument(s) changed over the course of the study	Supplement Table S3
Units of study - Number and relevant characteristics of participants, documents, or events included in the study; level of participation (could be reported in results)	9
Data processing - Methods for processing data prior to and during analysis, including transcription, data entry, data management and security, verification of data integrity, data coding, and anonymization/de-identification of excerpts	8

Data analysis - Process by which inferences, themes, etc., were identified and developed, including the researchers involved in data analysis; usually references a specific paradigm or approach; rationale**	5,8
Techniques to enhance trustworthiness - Techniques to enhance trustworthiness and credibility of data analysis (e.g., member checking, audit trail, triangulation); rationale**	9

Results/findings

Synthesis and interpretation - Main findings (e.g., interpretations, inferences, and themes); might include development of a theory or model, or integration with prior research or theory	10-12
Links to empirical data - Evidence (e.g., quotes, field notes, text excerpts, photographs) to substantiate analytic findings	Supplement Table S4

Discussion

Integration with prior work, implications, transferability, and contribution(s) to the field - Short summary of main findings; explanation of how findings and conclusions connect to, support, elaborate on, or challenge conclusions of earlier scholarship; discussion of scope of application/generalizability; identification of unique contribution(s) to scholarship in a discipline or field	13-17
Limitations - Trustworthiness and limitations of findings	16-17

Other

Conflicts of interest - Potential sources of influence or perceived influence on study conduct and conclusions; how these were managed	Title Page
Funding - Sources of funding and other support; role of funders in data collection, interpretation, and reporting	Title Page

*The authors created the SRQR by searching the literature to identify guidelines, reporting standards, and critical appraisal criteria for qualitative research; reviewing the reference lists of retrieved sources; and contacting experts to gain feedback. The SRQR aims to improve the transparency of all aspects of qualitative research by providing clear standards for reporting qualitative research.

**The rationale should briefly discuss the justification for choosing that theory, approach, method, or technique rather than other options available, the assumptions and limitations implicit in those choices, and how those choices influence study conclusions and transferability. As appropriate, the rationale for several items might be discussed together.

Reference:

O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. **Standards for reporting qualitative research: a synthesis of recommendations.** *Academic Medicine*, Vol. 89, No. 9 / Sept 2014. DOI: 10.1097/ACM.0000000000000388

Appendix 4.1 – Table 2. Consolidated Criteria for Reporting Qualitative Studies (COREQ) Checklist

Item No	Guide Questions/Description	Reported on Page #
Domain 1: Research team and reflexivity		
Personal Characteristics		
1. Interviewer/ facilitator	Which author/s conducted the interview or focus group?	5
2. Credentials	What were the researcher's credentials? E.g., PhD, MD	5
3. Occupation	What was their occupation at the time of the study?	5
4. Gender	Was the researcher male or female?	5
5. Experience and training	What experience or training did the researcher have?	5-6
Relationship with participants		
6. Relationship established	Was a relationship established prior to study commencement?	6
7. Participant knowledge of the interviewer	What did the participants know about the researcher? e.g. personal goals, reasons for doing the research?	6
8. Interviewer characteristics	What characteristics were reported about the interviewer/facilitator? e.g. Bias, assumptions, reasons and interests in the research topic	5-6
Domain 2: study design		
Theoretical framework		
9. Methodological orientation and Theory	What methodological orientation was stated to underpin the study? e.g. grounded theory, discourse analysis, ethnography, phenomenology, content analysis	5
Participant selection		
10. Sampling	How were participants selected? e.g., purposive, convenience, consecutive, snowball	7
11. Method of approach	How were participants approached? e.g., face-to-face, telephone, mail, email	7
12. Sample size	How many participants were in the study?	9
13. Non-participation Setting	How many people refused to participate or dropped out? Reasons?	9
14. Setting of data collection	Where was the data collected? e.g., home, clinic, workplace	7
15. Presence of nonparticipants	Was anyone else present besides the participants and researchers?	7

Item No	Guide Questions/Description	Reported on Page #
16. Description of sample	What are the important characteristics of the sample? e.g. demographic data, date	Table 1
Data collection		
17. Interview guide	Were questions, prompts, and guides provided by the authors? Was it pilot tested?	8, Table S1
18. Repeat interviews	Were repeat interviews carried out? If yes, how many?	N/a
19. Audio/visual recording	Did the research use audio or visual recording to collect the data?	8
20. Field notes	Were field notes made during and/or after the interview or focus group?	N/a
21. Duration	What was the duration of the interviews or focus group?	9
22. Data saturation	Was data saturation discussed?	7,10
23. Transcripts returned	Were transcripts returned to participants for comment and/or correction?	8
Domain 3: analysis and findings		
Data analysis		
24. Number of data coders	How many data coders coded the data?	8
25. Description of the coding tree	Did the authors provide a description of the coding tree?	Table S4
26. Derivation of themes	Were themes identified in advance or derived from the data?	10
27. Software	What software, if applicable, was used to manage the data?	8
28. Participant checking	Did participants provide feedback on the findings?	8
Reporting		
29. Quotations presented	Were participant quotations presented to illustrate the themes/findings? Was each quotation identified? e.g., participant number	Table S4
30. Data and findings consistent	Was there consistency between the data presented and the findings?	10-12
31. Clarity of major themes	Were major themes clearly presented in the findings?	10-12, Table 2
32. Clarity of minor themes	Is there a description of diverse cases or a discussion of minor themes?	10-12, Table 2

Developed from:

Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. International Journal for Quality in Health Care. 2007. Volume 19, Number 6: pp. 349 – 357

Appendix 4.1 – Table 3. Interview Guide

Identification	• Today is [insert date] and this is my interview with [Insert ID#]
Preamble: Purpose	• During this interview, we will be talking about your experiences with developing and implementing artificial intelligence-based clinical decision support for the care of patients in the emergency department. • By clinical decision support, I am referring to decisions involving diagnosis, treatment, and predicting patient outcomes. • In our previous work, our team demonstrated that AI models may perform as well or better than clinicians for these kinds of decisions. However, most of this research is limited to theoretical performance in datasets. • We want to understand why, despite hundreds of AI models in development, most don't make it to the stage of being tested in live clinical scenarios. • And we want to learn strategies for how to build these models more effectively in the early development stages to mitigate issues when it comes to implementation. • We believe your perspective will be invaluable for helping us answer these questions.
Demographics	To begin, I would like to gather some demographic information to provide context to our discussion today. • Gender and preferred pronouns • Profession (e.g. physician, nurse, statistician, engineer) • Years in profession • Post-secondary educational degrees (e.g. Master's, PhD) • Academic experience with AI-based CDS (e.g. research, quality improvement, etc.) • Private industry experience with AI-based CDS (e.g. technology company, consulting, etc.)
Experience	First, I would like to know more about your experience with developing and implementing AI models for clinical decision support in the ED. 1. What AI models have you been involved in developing? 2. What AI models have you been involved in implementing? Were these the same that you had been involved in developing?
Challenges	3. Regarding model development, what were the major challenges that you faced? 4. In your opinion, could any of these challenges have mitigated later issues with model implementation? How?
Strategies and Solutions	5. Can you share successful strategies you have used or believe should be used to ensure effective AI model development? • Probe: Are there specific research or methodological guidelines that have proven effective in your work? • Probe: Are there specific experts or stakeholders that you feel are important to have on the team (e.g. data scientists, software engineers, patients, nurses, hospital administrators)? 6. Do you have any other recommendations for researchers aiming to translate a successful AI model from the development stage to implementation?
Closing	Is there anything else you'd like to add about your experiences in this field?
Recruitment	We are hoping to speak to other researchers who have experience with implementing AI/ML models into emergency department workflows. Do you have any individuals who you could recommend? If so, kindly pass along my e-mail address to them:
Conclusion	Thank you. That concludes our interview. I will end the recording now.

Appendix 4.1 – Table 4. Coding Tree with Identified Barriers and Solutions

Theme	Code	Type	Representative Quotation
Team Capacity	n/a	n/a	n/a
Team Capacity	Need for diverse team expertise	Barrier	*"To me, that's where innovation lies: it is in multidisciplinary science."*
Team Capacity	Need for technical expertise with data science and model building	Barrier	*"Find technical staff that that you really get along with and enjoy working with and that you can totally trust their work because as clinicians we're not going to be the ones that are able to do the predictive modelling or at least I am not."*
Team Capacity	Need for design, behavioural, and human factors expertise	Barrier	*"In their group they have software developers, human design experts, design human factors experts, design experts; people who actually really spend time understanding, "how do we create a solution around the models that people actually want to engage with?""*
Team Capacity	Need for long-term time commitments	Barrier	*"When you hire staff, hire to retain... You want to hire someone who's like, yeah, they're interested in a career of being a research coordinator and they're bright and they're committed to that kind of role."*
Team Capacity	Need for experience in healthcare applications	Barrier	*"Even when we started building models, we would do focus groups to, you know, our approach is always from the clinical side first to really ask clinicians... Not only do you capture some of the nuances in the clinical workflow from the very beginning, you also really get to know the domain expertise."*
Team Capacity	Need for collaboration amongst technical and clinical team members	Barrier	*"There's a dance there that happens that is actually bringing together very intimately the technical person who is generating those outcome metrics and fine tuning the model, and the clinical experts on the ground who understand what is an acceptable workflow."*
Team Capacity	Non-academic site does not feel ownership or incentive to test model	Barrier	*"It's a little tricky in that third setting, again, because like, they're not particularly incentivized to work with us... They're not like researchers. They're not like, oh, I'm gonna get out of publication or whatever else."*
Team Capacity	Need for additional leadership	Barrier	*"I find having another Co-PI is really a helpful and wonderful asset to the whole project, and I have pretty much uniformly done that for many, many years."*
Team Capacity	Need for the ability to understand and comply with academic requirements	Barrier	*"They had one individual who had I'll say minimal background in sort of research or understanding of the research process... I think the ability to understand the academic rigor that would be required for a peer reviewed process was lacking within their own team."*

Team Capacity	Create a data integration and governance team early to solve technical issues	Facilitator	*"First and foremost, the data integration and governance team. That's where we have our data modelers, architects, data engineer types. They do all the pipelines. It basically points at any source system at the production level and it grabs that data in real time without any disruption."*
Data Infrastructure	n/a	n/a	n/a
Data Infrastructure	Need for a clinical data warehouse	Barrier	*"Our ecosystem is terrible. It's like, really, really fractured across different databases and it's not integrated. So that data warehousing work is very, very important."*
Data Infrastructure	Need for augmenting standard data collection process	Barrier	*"You're often limited by whatever fields the hospital decides goes into the chart for the patients, and so if they're not asking about ethnicity or language or identified race, then it you can't really track any of that stuff. Which then makes it hard to get grants because all the big granting agencies wanna know, like, well, what's your gender based analysis plan?"*
Data Infrastructure	Need for data security	Barrier	*"There's also the governance piece like where is the data going? Like, do you have an infrastructure? Is the infrastructure approved from a privacy perspective? Is it approved from a cyber security perspective? Do you physically have the infrastructure? Are you putting it on premise? Are you putting it in a cloud environment?"*
Data Infrastructure	Need for data cleaning and labeling	Barrier	*"It's important to think about your inputs. You know, garbage in, garbage out. That's, I think, the most important thing. So spending the time upfront to figure out: where is your data going to come from? what's it going to look like? What kind of clean up you're going to have to do? What kind of labeling do you have to do to train the models?"*
Data Infrastructure	Need for real-time data processing	Barrier	*"You've got over 50 different source systems. It's a train wreck in here. We need to be able to point to servers in real time. Get data. Take it to a central place. Clean it up in real time. Feed it through AI algorithms, spit it back out into clinicians, and all in real time."*
Data Infrastructure	Need for streamlining data collection process	Barrier	*"So those wave forms are not collected routinely in any EHR... They show up on the monitor and they get displayed and then they get thrown away. And it's absurd how much data we actually throw away.* *Like it's incredibly challenging"*
Data Infrastructure	Use automated tool to efficiently capture data	Facilitator	*"...we're gonna use large language models to scan through the charts and try to extract all this stuff automatically, because it's just gonna be too onerous to, like, read through everyone's mental health chart."*
Data Infrastructure	Contract data infrastructure work out to private company	Facilitator	*"So it does sometimes help to have these private partnerships because, you know, they have lots of computer scientists and data scientists and people with expertise that are working with our is department to get those technological things working."*

Data Infrastructure	Share resources with other department within your institution	Facilitator	*"So we had these two guys and so we sort of used those two resources and shared those two resources with our digital health department on other projects."*
Defining Clinical Problem and Solution	n/a	n/a	n/a
Defining Clinical Problem and Solution	Proposed model output is not clinically useful	Barrier	*"The data scientists are like, "I got this big sexy model" and [the clinicians] were like, "well, I don't need that output," right?"*
Defining Clinical Problem and Solution	Problem was defined secondary to model performance on available data	Barrier	*"I think there's a tendency in this kind of work to pick a question that's easy to develop because you have the data... You're working under the assumption that you're going to be able to build something useful with this data."*
Defining Clinical Problem and Solution	Lack of defined or accepted performance thresholds	Barrier	*"We set a 5% miss rate. I'm not sure I will say if that's the right number or not. I know when we think about emergency medicine for certain conditions you know we might accept a 1 to 2% miss rate and that's reasonable, but there actually was no gold standard that I could appreciate at the time when we were doing this."*
Defining Clinical Problem and Solution	Problem was defined secondary to profit motive	Barrier	*"Their perspective was really about trying to make money and get this product out to market. My perspective was, "I'm trying to make sure that this is a safe product to do what we were hoping to do." And so sometimes our goals don't always align that way."*
Defining Clinical Problem and Solution	Availability of clinical solution may differ based on site	Barrier	*"When we went to implement it in this third emergency department, it was clear that patients weren't going to drive to an outpatient falls clinic, you know, 90 miles away."*
Defining Clinical Problem and Solution	Patient knowledge and education level can influence uptake of solution	Barrier	*"There's a real significant class and education difference between patients who want to know everything and understand exactly what happened... If we design to them, we completely overwhelm all the other patients who need like the bare minimum."*
Defining Clinical Problem and Solution	Patient refusal of device use in their care	Barrier	*"First patient we brought in said, "how dare you violate my privacy and tell people about an IV drug use issue I have when I don't want anyone to know, including the physicians treating me because I'm not here for that - I'm here for something else.""*

Defining Clinical Problem and Solution	Engage multiple healthcare end-users as part of your research team	Facilitator	*"...Design methodologies call upon stakeholder engagement, to understand a problem from multiple lenses of stakeholders. So, you know, the nurse, the physician, the patient, the ward clerk, whoever it needs to be..."*
Defining Clinical Problem and Solution	Engage patients early in process of defining the problem and possible solutions	Facilitator	*"So first we talked with patients and that's where it was very clear: these patients are not going to drive 90 miles, and so we're going to need a different intervention."*
Defining Clinical Problem and Solution	Define thresholds based on clinician preference	Facilitator	*"The clinicians kept coming back, saying again, "precision has to be X, you need to pick up this many cases, then I'm gonna trust the algorithm. Otherwise, I'm not gonna use it.""*
Defining Clinical Problem and Solution	Create plain language explanations of tool for patients	Facilitator	*"Even if it's a non medical jargon, explainability from a patient lens - that goes a long way, and if you don't have that explainability or you can't really explain some of the methodology, I think it's a very steep uphill to be able to move these forward into that implementation phase."*
Defining Clinical Problem and Solution	Enlist a skeptic of the technology to help define the problem or solution	Facilitator	*"Find somebody who's skeptical in your department, another emergency physician and nurse, somebody who you think is not going to believe in it... They're going to tell you why it's not going to work. They're going to tell you right away before you spend 2 years developing it..."*
Research, Ethics, and Regulatory Approval	n/a	n/a	n/a
Research, Ethics, and Regulatory Approval	Getting regulatory approval from those with limited AI knowledge	Barrier	*"One of the first things, for example, that came up at the REB was a statement back to me saying, "This is a black box. We don't understand AI. What is this methodology of machine learning?" Even though I will say it was the longest REB I've ever submitted, it felt like they still didn't understand, both clinicians and non-clinicians who sat on the REB."*
Research, Ethics, and Regulatory Approval	Lack of clear institutional policies and processes regarding data sharing	Barrier	*"So first of all getting access to data was very, very challenging for us at our institution because it was the first large scale like kind of big data... from a healthcare perspective, this was big data to, let's say, the ethics board. Nobody had ever asked them for 600,000 records before."*
Research, Ethics, and Regulatory Approval	Multitude of regulatory approvals needed for using AI models in clinical care	Barrier	*"There's like different factors and considerations that start to kick in, such as does Health Canada regulatory approvals apply to this? ...there's a new ADA act that's coming out that might apply actually to high risk models. There's maybe provincial legislation that comes through in the upcoming years."*

Research, Ethics, and Regulatory Approval	Create streamlined research ethics board approval process for AI methodology	Facilitator	*"What we recognized was that the traditional approach to REB isn't necessarily the best one for the first phase of retrospective AI model building. And so we reinvented the REB to accommodate AI models."*
Research, Ethics, and Regulatory Approval	Deliver dedicated presentation to research ethics board to explain AI methodology	Facilitator	*"And then I actually responded with an offer and said, "Would you like me to present to the REB?" This was actually the recommendation of one of our external methodologists and said, "Maybe you just need to offer to talk to them because they're clearly not understanding this.""*
Research, Ethics, and Regulatory Approval	Align work with institutional strategic priorities or vision	Facilitator	*"...we were very much aligned, I think, in terms of some of the key metrics that we were looking at for this, but this kind of takes that to a whole new level now. And so, you know, grassroots has become now this organizational process."*
Research, Ethics, and Regulatory Approval	Create institutional data sharing policy	Facilitator	*"Now there is, very recently, we actually have a confirmed policy on AI development and AI use. It's brief, but it does go through at least the steps of who needs to be engaged and at what level. And it's also now becoming part of our strategy within our research institute at the hospital."*
Research, Ethics, and Regulatory Approval	Use a deidentified database to improve chances of approval	Facilitator	*"...we developed a deidentification process between the two [servers], which is set up and then managed by the data management team. And so that means the number of people who had access to completely identified data is very, very small."*
Legal and Liability	n/a	n/a	n/a
Legal and Liability	Concerns about liability regarding medical errors	Barrier	*"Ultimately the legal liability will lie with you and me when we are trying to use that tool for clinical decision making, and not the vendor or the hospital... And that's where I think there's been some problems with like automatic deploying of consults or automatic ordering of tests."*
Legal and Liability	Concerns about liability regarding data leak	Barrier	*"It took about six months of me having multiple meetings with lawyers asking me things like, "Okay, we're going to give you access to this data. If the data gets leaked somehow, are you going to be legally liable? Like, can we sue you afterwards?""*
Legal and Liability	Unclear intellectual property ownership	Barrier	*"I know that other hospitals in the city do have more stronger IP protection then it's a big issue... There's like constant years of bickering over like, does the university own it? Does the hospital own it?"*
Legal and Liability	Build internal legal team	Facilitator	*"My recommendation would be having a legal representative. In our case, we had an information security record representative and an information technology representative - slightly different - and then someone from a research arm or a data sharing perspective... I think having those folks on board, internally having an aligned plan, definitely made things easier, externally..."*

Legal and Liability	Bring in external legal team	Facilitator	*"What we heard from our internal folks, saying, "You can't do this." Within a few months, we brought in an external legal firm, and they went through the legislation. They went through everything and said, "Absolutely you can do this. Here's the framework.""*
Model Building and Performance	n/a	n/a	n/a
Model Building and Performance	Data set drift limits model performance	Barrier	*"What you might start to capture is a bit of drift in the environment. Maybe there's seasonality, right? Maybe there's a new disaster that occurs COVID 2.0 or 3.0, and then all of a sudden, the way we behave clinically is changing. Your model might fall apart."*
Model Building and Performance	Adding too many model features resulted in worse performance	Barrier	*"We had [four to six hundred features] and quickly realized there's a marginal rate of return on information, and there comes a plateau where you just don't need more to get a certain level of performance. In fact, when you keep adding more, your performance decreases."*
Model Building and Performance	Lack of data makes model non-generalizable	Barrier	*"Maybe the algorithm works in our population that I see in my hospital, but I can't generalize it to another hospital because I have no idea [if those patients are similar to ours]."*
Model Building and Performance	Model explainability is often desired but has limited utility	Barrier	*"People were wondering what's driving these predictions. And so we have Shapley values and the typical things that you do to see what's driving the probability estimates. Well, we quickly realized while it's interesting from a research perspective, it's not very useful clinically… because the beauty with machine learning is it actually gets at all the complexities in the data that we have a hard time understanding and processing."*
Time	n/a	n/a	n/a
Time	Time to build a research team	Barrier	*"But to actually evolve to that point where we're building the teams, I think it took between three to four years."*
Time	Time to build enterprise data warehouse	Barrier	*"Spent a lot of time, three years, $5,000,000 building our enterprise data warehouse, which I think we could have done in a third of the time…"*
Time	Time to clean and label data	Barrier	*"Getting the data and labeling the data is definitely a big problem because it's very time consuming, especially in the initial training."*
Time	Time to get research ethics board approval	Barrier	*"And REB process took months, and that didn't surprise me, but I think the expectation setting there I could have probably done that a little bit better."*
Time	Time to obtain grants for research projects	Barrier	*"It's a slog to get a CIHR grant. It's not like, oh, you know, people who get CIHR funding, just submit them and they get it on the first try. That's not the case. Even the most experienced grant writers are getting grants after the second, third attempt…"*
Time	Time to prospectively collect data that is not	Barrier	*"The only real way then is to do prospective enrollment. So for the asthma study, I had a medical student who would speak to each of the participants and then ask them those questions. But*

	available from electronic health record		*that's time consuming and costly, whereas it would be much nicer if it could just be pulled automatically from the electronic health record."*
Cost	n/a	n/a	n/a
Cost	Cost of hiring and retaining team	Barrier	*"When you hire staff, hire to retain… It gets expensive a little bit, but it's worthwhile."*
Cost	Cost of accessing data from electronic health record	Barrier	*"A lot of the EHRs make you pay an API access fee to get your own data out of the EHR, which I think is criminal. We should negotiate better on renewals of our EHR licenses and our contracts."*
Cost	Cost of building enterprise data warehouse	Barrier	*"…$5,000,000 building our enterprise data warehouse, which I think we could have done in a third of the time and a quarter of the budget, and we ended up fixing their code here after they left."*

Themes are signified by shaded rows.

Chapter 5. Discussion

This chapter provides a summary of the main thesis findings and their relevance to the field of AI-CDS research for the ED. It highlights strengths and limitations of this work and outlines implications for future research.

5.1. Summary

This thesis assessed the current state of AI-CDS tools in the ED and to identify strategies that facilitate their integration into clinical practice. To achieve this, we examined the considerable gap between AI-CDS tools that have been developed and those that have been successfully implemented. We then incorporated insights from domain experts to explore the underlying reasons for this discrepancy and propose strategies to bridge it.

In Chapter 2, we contextualized the increasing burden on emergency care systems and the potential role of AI and ML tools, particularly AI-CDS, in alleviating this crisis. Despite widespread interest among emergency physicians in adopting such technologies, there remains a general lack of awareness regarding available tools and their effectiveness in real-world clinical settings.

Chapter 3 presented the results of a comprehensive scoping review, which provided a broad overview of the rapidly expanding field of AI-CDS in EDs. While the number of publications has grown significantly since 2018, particularly in areas such as

prognostication and triage, only a small fraction of tools have progressed beyond preclinical development into live clinical validation or deployment.

In Chapter 4, we explored the reasons behind this gap – the so-called "AI chasm" – through qualitative interviews with expert researchers who have both developed and implemented AI-CDS tools in the ED. Using grounded theory methodology, we theorized that the AI chasm exists due to insufficient consideration of future implementation, and subsequently identified the key barriers and facilitators that researchers must address in order to successfully translate their tools into clinical practice.

5.2. Key Study Findings

This thesis makes two primary contributions to the field of AI-CDS in emergency medicine: it characterizes the widening gap between developed and implemented tools, and it proposes an expert-informed framework to address this issue.

The first study, a large-scale scoping review, demonstrated that the field of emergency AI-CDS has been growing rapidly, with over 100 relevant publications annually from 2021 to 2023. Prognostic outcomes, such as mortality and ICU admission, were the most frequently studied (44.6% of included articles), followed by diagnostic tools, disposition aids, triage level predictions, and therapeutic recommendations. Encouragingly, the focus areas of existing tools appear to align with the clinical needs expressed by emergency physicians. However, only 3.5% of tools had progressed to testing or validation in live clinical environments, revealing a significant disconnect between development and implementation.

To understand the factors contributing to this gap, we conducted qualitative interviews with experienced AI-CDS researchers. Analysis revealed 41 barriers, categorized into eight themes: 1) Team capacity; 2) Data infrastructure; 3) Defining the clinical problem and solution; 4) Research, ethics, and regulatory approval; 5) Legal and liability; 6) Model building and performance; 7) Time constraints; and 8) Financial costs. Concept map analysis further identified high-yield facilitators, particularly "engaging multiple healthcare end-users" and "sharing resources with other departments." Our resulting grounded theory emphasizes the necessity of incorporating implementation science frameworks from the outset of tool development and assembling agile, interdisciplinary teams equipped to manage regulatory, legal, and logistical complexities.

Based on our study findings, we recommend researchers follow these steps when conceptualizing an AI-CDS tool:

1. **Define a Clinically Salient Problem**

   Begin by identifying a clinical problem that holds demonstrable relevance for both healthcare providers and patients. Engage stakeholders from these groups early in the process to ensure the proposed research addresses genuine needs and priorities. Establish structured mechanisms for obtaining ongoing feedback from these stakeholders throughout the development of the research proposal to enhance its relevance, feasibility, and potential for impact.

2. **Assemble a Multidisciplinary Research Team**

   Create a research team comprising expertise across clinical domains, data science, software engineering, informatics, and implementation science. Incorporate patient

representatives or individuals with lived experience where appropriate to ensure alignment with patient-centered care principles. Employ an iterative feedback process with this team to refine the research proposal, methodology, and implementation plan, fostering interdisciplinary collaboration.

3. **Evaluate Institutional Data Infrastructure**

   Conduct a thorough assessment of your institution's data infrastructure to determine its capacity to support the proposed project. This includes evaluating existing data warehouses, EHR systems, and interoperability pathways. Collaborate closely with data scientists, software engineers, and informaticians to identify potential gaps and develop customized data processing pathways as needed. Where possible, leverage institutional knowledge by consulting with other research groups that have successfully navigated similar challenges.

4. **Anticipate Data Privacy, Ethics, and Legal Considerations**

   Proactively identify potential challenges related to data privacy, security, and regulatory compliance, particularly if the proposed model is intended for deployment in live clinical care. Engage medical ethicists and legal experts to evaluate ethical implications, assess adherence to jurisdictional data governance frameworks, and develop mitigation strategies for identified risks.

5. **Estimate Development Timeline and Resource Requirements**

   Develop an anticipated timeline and costs associated with each stage of the development process. This estimate should reflect the complexity of tasks such as data acquisition, model development, validation, and implementation, as well as the

contributions of involved experts. Simultaneously, explore potential funding mechanisms, including research grants, institutional support, and strategic partnerships with industry stakeholders to ensure financial sustainability.

6. **Develop and Validate a Proof-of-Concept Model**

   Initiate model development using available datasets and adhere to established methodological best practices for model design, training, and evaluation.[1] Aim to produce a robust proof-of-concept that demonstrates baseline performance and informs the design of subsequent validation and implementation studies. Ensure transparent reporting of model development processes in alignment with the TRIPOD+AI guidelines to facilitate reproducibility and critical appraisal.[2]

5.3. Importance of Findings

This work offers essential insights into the current landscape of AI-CDS tools for emergency medicine by highlighting the presence of a significant translational gap and identifying actionable strategies to address it. While previous research has suggested that AI-CDS tools can match or surpass standard care in certain ED applications, our findings underscore that such potential is moot without real-world implementation.

This thesis provides concrete guidance for researchers on integrating implementation principles early in the development process to prevent tools from becoming “research waste.” From establishing appropriate data infrastructure and fostering multidisciplinary collaboration, to navigating ambiguous regulatory pathways,

this work addresses critical considerations that differ markedly from conventional medical research.

Our findings complement and contextualize emerging guidance in the field, including the DECIDE-AI guidelines and the lead author's recently published work, *Establishing Methodological Standards for the Development of Artificial Intelligence-Based Clinical Decision Support in Emergency Medicine*.[1, 3]

5.4. Strengths and Limitations

A major strength of this thesis is its methodological rigor and clarity in synthesizing findings from a rapidly evolving and heterogeneous research field. The scoping review followed best practices, including the PRISMA Extension for Scoping Reviews and the Joanna Briggs Institute Manual. The qualitative study adhered to both the Standards for Reporting Qualitative Research (SRQR) and the Consolidated Criteria for Reporting Qualitative Research (COREQ).

Despite limited time, we screened over 5,000 unique records, reviewed 721 full texts, and synthesized data from 605 articles. Full text review was performed by emergency physicians with prior AI research experience, enhancing both the relevance and accuracy of the findings. The qualitative interviews included participants with experience in both the development and implementation of AI-CDS tools, providing comprehensive perspectives. Through a structured, iterative coding process, we produced a focused and unified explanatory theory for the observed implementation gap.

Nevertheless, several limitations must be acknowledged. The scoping review, by design, did not assess the efficacy or bias of the included studies, limiting our ability to comment on clinical effectiveness. Participant diversity in the qualitative study was limited: nine male participants, one non-binary, and no female respondents; all based in North America. This homogeneity may limit the generalizability of our findings. Additionally, one of the two qualitative coders was not present during the interviews and did not have access to audio or video recordings, potentially reducing the ability to interpret non-verbal communication cues.

5.5. Implications for Future Research and Practice

As the number of AI-CDS tools designed for emergency medicine continues to grow, this thesis provides a roadmap for improving the chances of successful translation from development to implementation. We hope our findings will reduce the incidence of research waste by encouraging developers to embed implementation science principles from the outset.

This work also supports institutional investment in robust data infrastructure and the formation of interdisciplinary teams to facilitate AI-CDS development and deployment. Future research should elaborate on specific implementation strategies, such as optimal timing of stakeholder engagement and technical guidance for building real-time data analytics platforms that integrate with electronic health records.

Finally, a crucial next step involves reforming regulatory pathways. Both academic institutions and governmental bodies must clarify approval processes and establish

realistic standards for AI-CDS evaluation. The current lack of transparency and consistency in regulatory expectations poses a significant barrier to progress in this field.

5.6. Conclusion

AI-CDS tools offer a promising solution to the mounting challenges faced by EDs worldwide. While research activity in this domain has grown rapidly, only a small proportion of tools have been tested or implemented in clinical practice. This “AI chasm” is driven by multifaceted barriers, including insufficient attention to implementation science and the high resource demands of real-world deployment.

For AI-CDS tools to deliver on their promise, researchers must recognize and plan for these challenges early in the development process. Building interdisciplinary teams capable of navigating regulatory, legal, technical, and institutional hurdles will be essential for translating research into impact and improving emergency care at a regional, national, and global scale.

References

1. Kareemi H, Li H, Rajaram A, Holodinsky JK, Hall JN, Grant L, et al. Establishing methodological standards for the development of artificial intelligence-based Clinical Decision Support in emergency medicine. Cjem. 2025;27(2):87-95.
2. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. Bmj. 2024;385:e078378.
3. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med. 2022;28(5):924-33.

Appendices

Appendix 1. OHSN-REB Ethics Approval for Qualitative Study

The Ottawa Hospital RESEARCH INSTITUTE | L'Hôpital d'Ottawa INSTITUT DE RECHERCHE

uOttawa

UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

Ottawa Health Science Network Research Ethics Board (OHSN-REB) / Conseil d'éthique de la recherche du réseau de science de la santé d'Ottawa (CÉR-RSSO)

Date:	May 29, 2024
Principal Investigator:	Dr. Christian Vaillancourt, OHRI (for TOH/OHRI Investigators)
Protocol ID:	20240338-01H
Study Title:	Strategies for Developing Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Qualitative Study
Submission Type:	Initial Application
Review Type:	Delegated
Date of Approval:	May 29, 2024
Study Approval Expiry Date:	May 29, 2025

Dear Dr. Vaillancourt,

An **Institutional approval (OHRI) letter is required prior to commencing the study** . In order to obtain the institutional approval letter, the ethics, contracts, and departmental notifications tabs of the Clinical Research Registration Form (CRRF) must be approved/marked complete by the reviewing office.

Thank you for submitting the above referenced study. The Ottawa Health Science Network Research Ethics Board (OHSN-REB) has reviewed the application and granted approval for your study. This approval is granted until the expiration date noted above. This research study is to be conducted by the investigator noted above.

The **OHSN-REB ethics approval** for this study is for The Ottawa Hospital.

Documents Approved:

Document Name	**Document Version Date**
20240338-01H English-only Verbal Consent Script, dated May 29, 2024	May 29, 2024
20240338-01H Protocol, dated May 29, 2024	May 29, 2024
20240338-01H English-only Information Sheet, dated May 29, 2024	May 29, 2024
20240338-01H English-only Initial Recruitment Email, dated May 28, 2024	May 28, 2024
20240338-01H English-only Interview Guide, dated May 09, 2024	May 9, 2024
20240338-01H English-only Reminder Email #1, dated May 28, 2024	May 28, 2024
20240338-01H English-only Reminder Email #2, dated May 28, 2024	May 28, 2024

If applicable, the pending translated documents must be uploaded into the "Translated Documents" tab of the Clinical Research Registration Form within 90 days of REB approval.

No deviations from, or changes to, the protocol should be initiated without prior written approval of an appropriate amendment from the OHSN-REB, except when necessary to eliminate immediate hazard(s) to study participants.

REB members involved in the research project do not participate in the review, discussion or decision.

If the study is to continue beyond the expiry date noted above, a Continuing Review Form must be received by the OHSN-REB on or prior to the full board submission deadline date of the meeting scheduled to occur a minimum of 30 days prior to the study expiry date. If the study has been completed by the expiry noted above, a Study Closure Form must be received by the OHSN-REB.

The OHSN-REB operates in compliance with, and is constituted in accordance with, the requirements of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2); International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; Integrated Addendum to ICH E6 (R1): Guideline for Good Clinical Practice E6 (R2); Part C, Division 5 of the Food and Drug Regulations; Part 4 of the Natural Health Products Regulations; Part 3 of the Medical Devices Regulations; and the provisions of the Ontario Personal Health Information Protection Act (PHIPA 2004) and its applicable regulations or with the definition in the Interim Order Respecting Clinical Trials for Medical Devices and Drugs Relating to COVID-19. OHSN-REB is qualified through the CTO REB Qualification Program and is registered with the U.S. Department of Health and Human Services (DHHS) Office for Human Research Protection (OHRP).

Please do not hesitate to contact us if you have any questions.

Raphael Saginur, M.D.
Chairperson
Ottawa Health Science Network Research Ethics Board

/HMc

The Ottawa Hospital | L'Hôpital d'Ottawa
RESEARCH INSTITUTE | INSTITUT DE RECHERCHE

Appendix 2. Research Protocol for Qualitative Study

Strategies for Developing Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Qualitative Study

Authors:
Hashim Kareemi
Krishan Yadav
Wojtek Michalowski
Christian Vaillancourt

1. Introduction:

There is increasing interest in using artificial intelligence (AI) and machine learning (ML) models to support high-stakes, time-limited clinical decision making in the emergency department (ED), particularly as crowding, staffing issues, and increasing wait times further complicate these decisions.[1] In a systematic review of ED studies, our team demonstrated these models may match or even outperform clinicians and non-AI clinical decision support (CDS).[2] However, all of these models remained in preclinical "in silico" phases of development, being tested retrospectively to assess performance metrics. Whether this translates into improving patient outcomes remains unknown. Early results from a scoping review our team is currently performing demonstrate a plethora of AI-based CDS models in development with very few being implemented into clinical workflows or tested in interventional trials. This disparity, termed the "AI chasm",[3] has been demonstrated in other medical fields beyond emergency medicine,[4-6] but a thorough understanding as to why it exists remain elusive, as do solutions to bridge over it.

One possible explanation for the AI chasm, and a recurrent theme demonstrated in AI-based CDS studies, is the prevalence of methodological flaws at the model development stage. Our team published a systematic review comparing AI-based CDS models to usual care for diagnostic and prognostic prediction in the ED. We found that all 23 included studies were at high risk of bias,[2] and that many of the causes of bias had similarly been raised in other reviews outside of the ED setting.[7, 8] These methodological issues may preclude adoption of these models into clinical practice, whether by limiting model functionality (e.g. inability to handle missing data), introducing unacceptable biases

(e.g. excluding or underrepresenting specific patient demographics), or worsening performance in subsequent validation or clinical trials (e.g. overfitting due to low event rates), amongst other reasons. Recently, a development pathway for AI-based CDS models has been proposed, mirroring those of pharmaceutical and surgical innovations.[9] It combines guidelines for specific stages of model development including "offline" validation of models for prediction (Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis – Artificial Intelligence; TRIPOD-AI) and diagnostic test accuracy (Standards for Reporting of Diagnostic Accuracy Studies; STARD-AI),[10] early "live" evaluations (Development and Exploratory Clinical Investigations of Decision support systems driven by Artificial Intelligence; DECIDE-AI;),[9] production of trial protocols (Standard Protocol Items: Recommendations for Interventional Trials; SPIRIT-AI),[11] and reporting findings for large-scale effectiveness trials (Consolidated Standards of Reporting Trials; CONSORT-AI).[12] Ideally, this pathway would address and mitigate many of the methodological issues that contribute to later challenges with model implementation. However, each of the component guidelines of the DECIDE-AI pathway were only published in 2020 or later, and the TRIPOD-AI guideline is yet to be published. Therefore, the ability of these guidelines to influence recent model development remains limited.

Expert recommendations are needed to supplement understanding and provide practical context to using guidelines. The experience of those who have developed and attempted to implement AI-based CDS models, particularly in the unique clinical environment of the ED, may provide insight into the benefits of well-designed models and

the shortcomings of poorly designed ones. Identifying strategies for developing effective AI-based CDS models can then be shared with other clinicians and researchers. Previous qualitative research in this area has mostly been limited to interviews and surveys of end-users regarding their needs, wants, and apprehensions regarding AI in emergency medicine in general,[13] or towards specific applications.[14-16] None have focused on the experts with insight into how these models were derived and deployed.

2. Objectives:

The objectives of this work are:

a. To identify issues at the development stage of AI-based CDS models that cause subsequent challenges with implementation into clinical workflows and practice in the ED.
b. To explore strategies for developing AI-based CDS models that can be successfully implemented into clinical workflows and practice in the ED for the purpose of improving patient care or resource utilization.

3. Methodology:

3.1 Study Design:

This research will employ a qualitative design using one-on-one semi-structured interviews with experienced researchers in the field of AI/ML applied to emergency medicine.

3.2 Theoretical Framework:

"Grounded theory" will underpin the methodological approach to this work. This approach involves inductive analysis of data to produce a "theory grounded in data".[17] The study in question will develop a theory as to how issues in the development stage of AI-based CDS

models may lead to challenges with subsequent implementation and how these issues can be mitigated. As a component of grounded theory, this study will utilize theoretical sampling to iteratively analyze and focus the data collection process to build the theory.[17]

3.3 Participants:

Participants will include clinician (e.g. physician or nurse) or non-clinician (e.g. data scientist, epidemiologist, engineer, statistician) researchers that are named authors on at least two previous or ongoing projects involving AI-based models used for clinical decision support. These projects, whether individually or in total, should satisfy the following three criteria:

A) Model development (corresponding to either the "preclinical development" or "offline validation" phases of the DECIDE-AI development pathway).[9]
B) Model implementation (corresponding to either the "safety/utility, small scale", "Safety/effectiveness, large-scale", or "Post-market surveillance" stages of the DECIDE-AI development pathway).[9]
C) Clinical decision support related to the individual care of ED patients.

These criteria will be judged by the authors prior to inclusion of the participant in the study. The rationale for the participant inclusion criteria is to identify experienced researchers who can speak to barriers and facilitators regarding model development and implementation within the ED context, in order to satisfy the research objectives.

We will use purposive sampling, identifying potential participants from relevant research publications found in our previous or ongoing work in this field. We will also identify potential participants through the Canadian Association of Emergency Physician's Artificial Intelligence Special Interest Group, which brings together 14 Canadian physicians

and scientists, including the primary author of this work (HK), with an interest in researching and using AI models in emergency medicine. We will also use snowballing recruitment in which approached participants will be asked to recommend other potential participants. Potentially eligible participants will be contacted via email for voluntary participation in the study.

As recommended by Francis et al regarding data saturation,[18] we will specify our initial analysis sample as 10, based on comparable studies.[14-16, 19] Once 10 participants are interviewed, saturation will be defined as 2 consecutive participants being interviewed without contributing additional material to advancing the ground theory, as judged by consensus of the authors of this study. For pragmatic reasons related to limited time constraints, a maximum of 20 interviews will be conducted regardless of whether saturation has been achieved, as was done in a similar study.[20]

3.4 Setting:

Interviews will be conducted by the primary author (HK) via video call using the Microsoft Teams platform. Only a single participant will be present for each interview and no other members of the research team will be present, allowing for a more personalized and in-depth exploration of participants' experiences that remains consistent between interviews.

3.5 Data Collection:

We anticipate each interview will last at least 30-45 minutes. The interview guide (see separate document) has been designed by the authors and informed by our past

systematic review and ongoing scoping review. Probing questions may be used by the interviewer to allow elaboration on pertinent points.

We will attain verbal consent to record the interview. Audio recordings of the interviews will be kept in an encrypted password-protected file and only accessible to the study team.

We will pilot the interview process with at least one participant who does not necessarily meet the inclusion criteria. We will ensure the interview questions are sufficient in meeting the research objectives, and adjust the guide as needed. We will also ensure the audio recording, transcription, and coding process is working effectively. The data from these pilot interviews will remain confidential and will not be analyzed with the subsequent interviews.

4. Analysis

4.1: Data Transcription and Analysis:

Audio recordings will be transcribed verbatim using Microsoft Teams.

Coding will be performed using NVivo 12 (QSR International, Doncaster, Australia).[21][1] The primary coder (HK) will read the transcripts line-by-line, remove any identifying information, and provide spelling and grammatical edits. The subsequent coding process will be undertaken by two independent coders (HK & TBD) on the de-identified transcripts in three iterative stages: initial, focused, and theoretical.[17] The initial stage will compare "data with data", using concise codes to efficiently parse through the data.[17] The focused stage will amalgamate the initial codes into "more directed,

selective, and conceptual" codes, as well as identifying higher-level categories of codes.[22][2] The theoretical stage will involve integrating the focused codes and categories into an underlying grounded theory.[17] The two coders will meet after the first 5, 10, and then every additional interview in order to assimilate and distinguish codes. This process will also allow for assessment of saturation, which if identified by the two coders, will be brought to the rest of the authors for consensus.

A coding tree will be produced based on the analyses of the two coders. As a component of descriptive statistics of the codes, a list of identified themes will be ranked according to frequency of the number of participants who contribute to them. A narrative synthesis of the findings will also be produced, incorporating illustrative de-identified quotes from the participants.

4.2 Reporting:

The study will adhere to the SRQR (Standards for Reporting Qualitative Research)[23][3] and COREQ (Consolidated Criteria for Reporting Qualitative Research)[24] checklists to ensure methodological rigor and transparency.

5. Ethical Considerations:

Verbal informed consent will be obtained from all participants. Confidentiality of participants will be strictly maintained throughout the study. Only the study team will know the identity of the participants. Participants will be assigned a de-identified code that will be used to refer to their responses throughout the rest of the data collection and analysis

process. Participants will have the right to withdraw from the study at any stage without consequences. We will not discuss any specific patient information during the interviews.

6. Expected Outcomes:

This research aims to contribute valuable insights into the barriers and facilitators encountered by researchers in developing and implementing AI/ML models in emergency department workflows. The outcomes will help inform future strategies and interventions to facilitate the successful integration of AI/ML technologies into clinical practice.

7. Timeline:

- REB Approval: March-April 2024
- Participant Recruitment, Data Collection and Data Analysis: May-June 2024
- Manuscript Preparation: July-August 2024

8. Impact:

By identifying strategies for developing effective AI-CDS models, this research will provide invaluable recommendations to researchers hoping to bridge the AI chasm for the betterment of patients and entire healthcare systems.

Appendix 2 – References:

1. Kelen GD, Wolfe R, D'Onofrio G, Mills AM, Diercks D, Stern SA, et al. Emergency Department Crowding: The Canary in the Health Care System. Catalyst non-issue content. 2021;2(5).
2. Kareemi H, Vaillancourt C, Rosenberg H, Fournier K, Yadav K. Machine Learning Versus Usual Care for Diagnostic and Prognostic Prediction in the Emergency Department: A Systematic Review. Acad Emerg Med. 2021;28(2):184-96.
3. Keane PA, Topol EJ. With an eye to AI and autonomous diagnosis. npj Digital Medicine. 2018;1(1):40.
4. Yin J, Ngiam KY, Teo HH. Role of Artificial Intelligence Applications in Real-Life Clinical Practice: Systematic Review. J Med Internet Res. 2021;23(4):e25759.
5. Plana D, Shung DL, Grimshaw AA, Saraf A, Sung JJY, Kann BH. Randomized Clinical Trials of Machine Learning Interventions in Health Care: A Systematic Review. JAMA Netw Open. 2022;5(9):e2233946.
6. Lam TYT, Cheung MFK, Munro YL, Lim KM, Shung D, Sung JJY. Randomized Controlled Trials of Artificial Intelligence in Clinical Practice: Systematic Review. J Med Internet Res. 2022;24(8):e37188.
7. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. Lancet Digit Health. 2019;1(6):e271-e97.
8. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over

logistic regression for clinical prediction models. J Clin Epidemiol. 2019;110:12-22.

9. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med. 2022;28(5):924-33.
10. Sounderajah V, Ashrafian H, Golub RM, Shetty S, Fauw JD, Hooft L, et al. Developing a reporting guideline for artificial intelligence-centred diagnostic test accuracy studies: the STARD-AI protocol. BMJ Open. 2021;11(6):e047709.
11. Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ, Spirit AI, et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. Lancet Digit Health. 2020;2(10):e549-e60.
12. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, Spirit AI, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med. 2020;26(9):1364-74.
13. Stewart J, Freeman S, Eroglu E, Dumitrascu N, Lu J, Goudie A, et al. Attitudes towards artificial intelligence in emergency medicine. Emerg Med Australas. 2023.
14. Fujimori R, Liu K, Soeno S, Naraba H, Ogura K, Hara K, et al. Acceptance, Barriers, and Facilitators to Implementing Artificial Intelligence-Based Decision Support Systems in Emergency Departments: Quantitative and Qualitative Evaluation. JMIR Form Res. 2022;6(6):e36501.
15. Sax DR, Sturmer LR, Mark DG, Rana JS, Reed ME. Barriers and Opportunities Regarding Implementation of a Machine Learning-Based Acute Heart Failure

Risk Stratification Tool in the Emergency Department. Diagnostics (Basel). 2022;12(10).

16. Jordan M, Hauser J, Cota S, Li H, Wolf L. The Impact of Cultural Embeddedness on the Implementation of an Artificial Intelligence Program at Triage: A Qualitative Study. J Transcult Nurs. 2023;34(1):32-9.
17. Thornberg R, Charmaz K. Grounded theory and theoretical coding. The SAGE handbook of qualitative data analysis. 2014;5(2014):153-69.
18. Francis JJ, Johnston M, Robertson C, Glidewell L, Entwistle V, Eccles MP, et al. What is an adequate sample size? Operationalising data saturation for theory-based interview studies. Psychol Health. 2010;25(10):1229-45.
19. Sandhu S, Lin AL, Brajer N, Sperling J, Ratliff W, Bedoya AD, et al. Integrating a Machine Learning System Into Clinical Workflows: Qualitative Study. J Med Internet Res. 2020;22(11):e22421.
20. Galarneau LR, Scheuermeyer FX, Hilburt J, O'Neill ZR, Barbic S, Moe J, et al. Qualitative Exploration of Emergency Department Care Experiences Among People With Opioid Use Disorder. Ann Emerg Med. 2023;82(1):1-10.
21. NVivo pro qualitative analysis software. 12.6.0.959 ed: QSR International Pty Ltd; 2018.
22. Charmaz K. Constructing grounded theory: A practical guide through qualitative analysis: sage; 2006.
23. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med. 2014;89(9):1245-51.

24. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care. 2007;19(6):349-57.

Appendix 3. Recruitment Email Sent to Potential Participants of Qualitative Study

Recruitment Email

Subject Line: Invitation to participate in study regarding AI-based clinical decision support in emergency medicine

Hello Dr. [Insert Name],

This email is being sent on behalf of Dr. Hashim Kareemi, Clinical Epidemiology Fellow and Dr. Christian Vaillancourt, Associate Professor at the University of Ottawa, Department of Emergency Medicine at The Ottawa Hospital.

You are receiving this email because you have been identified as a researcher who has previously published studies related to developing or implementing artificial intelligence (AI)-based clinical decision support (CDS) and have e-mailed regarding interest in being a participant in this study.

You are being asked to participate in a research study that we are conducting. Participation is voluntary.

Briefly, the study involves interviewing experts like yourself to gain insight into methodological issues with developing AI-based CDS, how these issues affect successful implementation into clinical workflows in the emergency department, and potential solutions to these issues. The interviews will be conducted virtually using Microsoft Teams and will last approximately 30-45 minutes. The information you share during the interview will be confidential. The Information Sheet is attached for your review.

If we do not hear back from you, we will send two reminder e-mails in two weeks and six weeks from the current date.

Thank you,

Hashim Kareemi
Department of Emergency Medicine, The Ottawa Hospital

Appendix 4. Information Sheet Sent with Recruitment Email to Potential Participants of Qualitative Study

Information Sheet

Study Title: Strategies for Developing Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Qualitative Study

Principal Investigator: Christian Vaillancourt, Emergency Medicine

Co-Investigator: Hashim Kareemi, Emergency Medicine

IS THERE A CONFLICT OF INTEREST?
There are no conflicts of interest to declare related to this study.

WHY IS THIS STUDY BEING DONE?
This is a qualitative study that employs semi-structured interviews with experienced researchers. The goal of this study is to identify methodological issues at the development stage of artificial intelligence (AI)-based clinical decision support (CDS) that may cause challenges with the subsequent implementation into clinical workflows in the emergency department (ED), and to identify potential solutions to these issues.

HOW MANY PEOPLE WILL TAKE PART IN THIS STUDY?
We estimate that 15 participants will be enrolled in the study. This study should take approximately 3 months to complete, and the results should be known in about 6 months to a year.

WHAT WILL HAPPEN DURING THIS STUDY?
Your participation in this study would involve the completion of one interview using MS Teams videoconference, lasting approximately 30-45 minutes. You will be asked to speak about your experience and perspective on developing and implementing AI-based CDS in the ED. The audio from the interview will be recorded and analyzed by the research team to identify common themes amongst the participants.

VOLUNTARY PARTICIPATION AND WITHDRAWAL
Taking part in this study is voluntary. You have the option to not participate at all, or you may choose to leave the study at any time (this is called withdrawal), without having to provide a reason.

If you decide to leave the study, you can ask that the information that was collected about you not be used for the study.

If you withdraw your consent, the study team will no longer collect your personal identifying information and personal views for research purposes, and your responses will be erased and removed from the transcripts, if requested.

WHAT ARE THE RISKS OR HARMS OF PARTICIPATING IN THIS STUDY?
There are no anticipated risks or harms of participating in this study. If for any reason you become uncomfortable while discussing your experiences. You may choose not to answer questions or leave the interview at any time if you experience any discomfort.

WHAT ARE THE BENEFITS OF PARTICIPATING IN THIS STUDY?
You may not receive direct benefit from participating in this study. We hope the information learned from this study will help other researchers in developing their own AI-based CDS in the future.

HOW WILL PARTICIPANT INFORMATION BE KEPT CONFIDENTIAL?
If you decide to participate in this study, we will only collect the information needed for this study.

Records identifying you at this centre will be kept confidential and, to the extent permitted by the applicable laws, will not be disclosed or made publicly available, except as described in this consent document.

Authorized representatives of the following organizations may look at your original research records at the site where these records are held, to check that the information collected for the study is correct and follows proper laws and guidelines.

- Ottawa Hospital Research Institute, the Sponsor of this study, and who oversees the conduct of research at this location.
- The Ottawa Health Science Network Research Ethics Board who oversees the ethical conduct of this study.

Information that is collected about you for the study (called study data) may also be sent to the organizations listed above. Your name, address, email, or other information that may directly identify you will not be used. The records received by these organizations may contain your participant code.

Communication via e-mail is not absolutely secure. We do not recommend that you communicate sensitive personal information via e-mail.

During the discussions, you will be encouraged to refrain from using names. If names or other identifying information is shared during the discussion, it will not be included in the written records.

The audio/video recordings will be stored in a secure location and listened to/viewed only by members of the research team. The recordings will be kept until they have been transcribed (turned into written records), and then they will be destroyed.

If the results of this study are published, your identity will remain confidential. It is expected that the information collected during this study will be used in qualitative analyses, published in a peer-reviewed journal, and presented to the scientific community at conferences, including the Canadian Association of Emergency Physicians annual conference.

Even though the likelihood that someone may identify you from the study data is very small, it can never be completely eliminated.

WHAT IS THE COST TO PARTICIPANTS?
Participation in this study will not involve any additional costs to you.

ARE STUDY PARTICIPANTS PAID TO BE IN THIS STUDY?
You will not be paid for taking part in this study.

WHAT ARE THE RIGHTS OF PARTICIPANTS IN A RESEARCH STUDY?
You will be told, in a timely manner, about new information that may be relevant to your willingness to stay in this study.

You have the right to be informed of the results of this study once the entire study is complete. If you would like to be informed of the results of this study, please contact the research team.

Your rights to privacy are legally protected by federal and provincial laws that require safeguards to ensure that your privacy is respected.

WHOM DO PARTICIPANTS CONTACT FOR QUESTIONS?
In case you have any questions, you can contact the study researcher:

Hashim Kareemi

For questions about your rights as a research participant or about ethical issues related to the study, you can contact The Ottawa Health Science Network Research Ethics Board at 613-798-5555, extension 16719, and speak to someone who isn’t involved in the study at all.

Appendix 5. Verbal Consent Script for Participants of Qualitative Study

Verbal Consent Script

Study Title: Strategies for Developing Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Qualitative Study

Principal Investigator: Christian Vaillancourt, Emergency Medicine

Co-Investigator: Hashim Kareemi, Emergency Medicine

OHSN-REB Number: 20240338-01H

Participant name:	**Person calling:**
Date Called:	**Time Called:**

N.B. Calls will take place with the participant over Microsoft Teams.

Hello, am I speaking with [name of the potential participant]?
***If potential participant is unavailable:**
Is there a better time to call back? Date/time:

***If potential participant indicates they are not interested:**
Thank you for your time. Goodbye.

***If potential participant is respondent:**
Hi, [name of the potential participant] this is Hashim Kareemi. I am calling you because you have agreed to be contacted for involvement in our research project.

Taking part in this study is voluntary. Deciding not to take part or deciding to leave the study later will not result in any penalty.

The study is being conducted by myself under the supervision of Dr. Christian Vaillancourt. You have been invited to participate in this study because you have been identified as a researcher with experience in developing and implementing artificial intelligence-based clinical decision support. Our research team believes your perspective will help us better understand how to create AI tools effectively and safely for use in the emergency department.

Please note that this call is being recorded for the purposes of transcribing and performing thematic analyses.

Is this an okay time to proceed with the interview?

☐ No *If no → Is there a better time? Date/time:
☐ Yes *If yes → *Continue with script below*

Conflicts of Interest:

There are no conflicts of interest to declare related to this study.

Research Activities:

This is a qualitative study that employs semi-structured interviews with experienced researchers. The goal of this study is to identify methodological issues at the development stage of artificial intelligence (AI)-based clinical decision support (CDS) that may cause challenges with the subsequent implementation into clinical workflows in the emergency department (ED), and to identify potential solutions to these issues.

We estimate that 15 participants will be enrolled in the study. This study should take approximately 3 months to complete, and the results should be known in about 6 months to a year.

Your participation in this study would involve the completion of one interview lasting approximately 30-45 minutes. You will be asked to speak about your experience and perspective on developing and implementing AI-based CDS in the ED. The audio from the interview will be recorded and analyzed by the research team to identify common themes amongst the participants.

Do you have questions about the activities this study involves?

☐ No *If no → *Continue with script below*
☐ Yes *If yes → *Answer questions and document all questions and answers before continuing with script below*

Questions: _____

Answers: _____

Other Comments: _____

Voluntary Participation and Withdrawal:

Taking part in this study is voluntary. You have the option to not participate at all, or you may choose to leave the study at any time (this is called withdrawal), without having to provide a reason.

If you decide to leave the study, you can ask that the information that was collected about you not be used for the study.

If you withdraw your consent, the study team will no longer collect your personal identifying information and personal views for research purposes, and your responses will be erased and removed from the transcripts, if requested.

Do you have questions about the voluntary nature and ability to withdraw from of this study?

☐ No ***If no →** *Continue with script below*
☐ Yes ***If yes →** *Answer questions and document all questions and answers before continuing with script below*

Questions: _____

Answers: _____

Other Comments: _____

Potential Risks, Harms, Discomforts:

There are no anticipated risks or harms of participating in this study. If for any reason you become uncomfortable while discussing your experiences. You may choose not to answer questions or leave the interview at any time if you experience any discomfort.

Do you have questions about the potential risks this study involves?

☐ No ***If no →** *Continue with script below*
☐ Yes ***If yes →** *Answer questions and document all questions and answers before continuing with script below*

Questions: _____

Answers: _____

Other Comments: _____

Potential Benefits:

You may not receive direct benefit from participating in this study. We hope the information learned from this study will help other researchers in developing their own AI-based CDS in the future.

Do you have questions about the potential benefits this study involves?

☐ No ***If no →** *Continue with script below*
☐ Yes ***If yes →** *Answer questions and document all questions and answers before continuing with script below*

Questions: _____

Answers: _____

Other Comments: _____

Privacy/Confidentiality:

If you decide to participate in this study, we will only collect the information needed for this study.

Records identifying you at this centre will be kept confidential and, to the extent permitted by the applicable laws, will not be disclosed or made publicly available, except as described in this consent document.

Authorized representatives of the following organizations may look at your original research records at the site where these records are held, to check that the information collected for the study is correct and follows proper laws and guidelines.

- Ottawa Hospital Research Institute, the Sponsor of this study, and who oversees the conduct of research at this location.
- The Ottawa Health Science Network Research Ethics Board who oversees the ethical conduct of this study.

Information that is collected about you for the study (called study data) may also be sent to the organizations listed above. Your name, address, email, or other information that may directly identify you will not be used. The records received by these organizations may contain your participant code.

Communication via e-mail is not absolutely secure. We do not recommend that you communicate sensitive personal information via e-mail.

During the discussions, you will be encouraged to refrain from using names. If names or other identifying information is shared during the discussion, it will not be included in the written records.

The audio/video recordings will be stored in a secure location and listened to/viewed only by members of the research team. The recordings will be kept until they have been transcribed (turned into written records), and then they will be destroyed.

If the results of this study are published, your identity will remain confidential. It is expected that the information collected during this study will be used in qualitative analyses, published in a peer-reviewed journal, and presented to the scientific community at conferences, including the Canadian Association of Emergency Physicians annual conference.

Even though the likelihood that someone may identify you from the study data is very small, it can never be completely eliminated.

Do you have questions about the how your privacy will be protected?

☐ No ***If no →** *Continue with script below*

☐ Yes ***If yes →** *Answer questions and document all questions and answers before continuing with script below*

Questions: _____

Answers: _____

Other Comments: _____

Cost to participation:

Participation in this study will not involve any additional costs to you.

Payment or Reimbursement:

You will not be paid for taking part in this study.

Do you have questions about the costs of participation or payment/reimbursement?

☐ No ***If no →** *Continue with script below*

☐ Yes ***If yes →** *Answer questions and document all questions and answers before continuing with script below*

Questions: _____

Answers: _____

Other Comments: _____

Participant Rights:

You will be told, in a timely manner, about new information that may be relevant to your willingness to stay in this study.

You have the right to be informed of the results of this study once the entire study is complete. If you would like to be informed of the results of this study, please contact the research team.

Your rights to privacy are legally protected by federal and provincial laws that require safeguards to ensure that your privacy is respected.

Questions:

In case you have any questions, you can contact the study researcher:

Hashim Kareemi

For questions about your rights as a research participant or about ethical issues related to the study, you can contact The Ottawa Health Science Network Research Ethics Board at 613-798-5555, extension 16719, and speak to someone who isn't involved in the study at all.

Have all of your questions been answered?

☐ No ***If no →** *Answer questions and document all questions and answers before continuing with script below*

Questions: _____

Answers: _____

Other Comments: _____

☐ Yes ***If yes →** *Continue with script below*

Consent:

Based on the description of the study, would you like to participate? Or would you like some time to think about it?

☐ No ☐ Yes ☐ More time to think about it

***If they do not want to participate:** Thank you for your time. Goodbye.

***If they wanted more time:** When would be a good time for me to call you back to answer any further questions you may have and obtain your decision?
Date/time:

***If they do want to participate:** *Continue with documentation of verbal consent below*

Documentation of Verbal Consent

Study Title: Strategies for Developing Artificial Intelligence-based Clinical Decision Support in the Emergency Department: A Qualitative Study
OHSN-REB Number: 20240338-01H

Name of Participant: _____

Date of Discussion: _____

Duration of Discussion: _____

SIGNATURES

- The participant's questions have been answered,
- The participant understands the information within this Verbal Consent Script,
- Each page of the Verbal Consent Script has been read to the participant.
- The participant agrees to take part in this study.

Investigator or Delegate Statement

I have carefully explained the study to the study participant. To the best of my knowledge, the participant understands the nature, demands, risks and benefits involved in taking part in this study.

_____	_____	_____
Signature of Person Conducting Consent Discussion	Printed Name and Role	Date

Appendix 6. Confirmation of Submission of Qualitative Study Manuscript to *Canadian Journal of Emergency Medicine*.

Date: 05 Aug 2025

From: "Canadian Journal of Emergency Medicine (CJEM)" cjem@caep.ca
Subject: Under Review

Dear Hashim Kareemi,

Your manuscript has passed the initial screen by the editors and has been sent out for peer review. At CJEM we endeavour to make a decision within 60 days of initial submission.

Best regards,

Editorial Office
Canadian Journal of Emergency Medicine

This letter contains confidential information, is for your own use, and should not be forwarded to third parties.

Recipients of this email are registered users within the Editorial Manager database for this journal. We will keep your information on file to use in the process of submitting, evaluating and publishing a manuscript. For more information on how we use your personal details please see our privacy policy at https://www.springernature.com/production-privacy-policy. If you no longer wish to receive messages from this journal or you have questions regarding database management, please contact the Publication Office at the link below.

In compliance with data protection regulations, you may request that we remove your personal registration details at any time. (Use the following URL: https://www.editorialmanager.com/cjem/login.asp?a=r). Please contact the publication office if you have any questions.