A User-Centered Design Approach to an Artificial Intelligence-Enabled Electronic Medical Record Encounter in Canadian Primary Care

by

Krizia Francisco

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2025
© Krizia Francisco 2025

Examining Committee Membership

The following served on the Examining Committee for this thesis.

External Examiner	**Saskia M. Kelders** Associate Professor Psychology, Health and Technology University of Twente
Supervisor	**Catherine M. Burns** Professor Systems Design Engineering University of Waterloo
Internal Members	**John Zelek** Associate Professor Systems Design Engineering University of Waterloo
	Kate Mercer Adjunct Professor Systems Design Engineering University of Waterloo
Internal-External Members	**James Wallace** Associate Professor Public Health Sciences University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

This work was part of a collaboration and research partnership between the University of Waterloo and TELUS Health. For the duration of my research, I was under the supervision of Dr. Catherine Burns. I collaborated closely with Dr. Puneet Seth (Chief Medical Director and primary care physician) and Aaron Brady (Director of Product) at TELUS Health. Puneet reviewed and collaborated on the preliminary designs of the AI-enabled EMR encounter mock-ups and aided with recruitment. Aaron Brady reviewed secondary designs of the AI-enabled EMR encounter mock-ups. A request to present this research was also completed for Dr. Jim Kavanagh, Medical Director at TELUS Health and founder of Practice Solutions Suite EMR.

Sormeh Mehri (master's student) and Sabrina Saiko (master's student) from the Advanced Interface Design Lab (AIDL) supported the coding and thematic analysis component of this research by individually reading through raw interview transcripts, coding independently, and meeting to discuss results and collaborate on preliminary redesign considerations

The following publication has been cited in "Chapter 2: Methods":

- Francisco, K. M., & Burns, C. M. (2024). An approach to potentially increasing adoption of an artificial intelligence–enabled electronic medical record encounter in Canadian primary care: Protocol for a user-centered design. *JMIR Research Protocols*, *13*. https://doi.org/10.2196/54365

The research was also accepted and presented at the following venues:

1. University of Toronto, Temerty Centre for AI Research and Education in Medicine (T-CAIREM). AI in Medicine Conference. October 13th, 2023 Oral Presentation

2. Human Factors and Ergonomics Society (HFES) Healthcare Symposium. March 30th – April 2nd 2025, Poster Presentation.

3. Health by Technology Conference. June 10th-June 11th, Oral Presentation. To be presented by Master's Student Sormeh Mehri; I was a contributing author.

Abstract

Introduction

The Canadian primary care system serves as the first point of contact for patients entering the healthcare system and plays a crucial role in ensuring continuity of care. Primary care clinicians provide a broad range of services, including disease prevention, health promotion, diagnosis, treatment, and coordination of referrals to specialized care.

Despite the widespread adoption of electronic medical records (EMRs) in primary care, their development has historically lagged in addressing the specific needs of primary care providers. For example, chronic disease management (CDM) is a core responsibility in primary care. Yet, EMRs have not evolved to incorporate native, CDM-focused tools that align with the complex, long-term nature of managing chronic conditions. With the emergence of artificial intelligence (AI), there is growing recognition of its potential to enhance clinical decision-making, optimize workflows, and improve patient outcomes. Despite these promising advancements, clinicians remain hesitant to adopt AI due to concerns surrounding trust, usability, and seamless integration into existing workflows.

Primary care is a complex and dynamic environment where clinicians must balance efficiency with patient-centered care, making it critical for AI systems to align with their needs rather than introduce additional cognitive or administrative burdens. Trust in AI is not inherent; rather, it is a critical factor influencing clinicians' willingness to incorporate AI-driven tools into their practice. Without trust, even the most advanced AI systems risk a lack of adoption, as clinicians must have confidence that these technologies will enhance, rather than hinder, patient care. Concerns around reliability, accuracy, and the ability to align with clinical workflows contribute to skepticism, making AI adoption a complex challenge. Clinicians also worry about unintended consequences for both patient outcomes and

professional responsibility, fearing that AI-generated recommendations could lead to issues, including diagnostic errors, inappropriate treatments, or a diminished sense of clinical accountability. Additionally, broader ethical and legal uncertainties further complicate integration, as unresolved questions regarding liability, accountability, and regulatory oversight leave clinicians uncertain about the implications of relying on AI-enabled tools embedded in their EMRs. Without clear governance structures and well-defined safeguards, hesitation around AI use in healthcare will persist, underscoring the need for thoughtful design and policy development. This study investigates how a user-centered design approach can address these challenges, ensuring that AI-enabled EMR encounters enhance, rather than disrupt, the clinician's role in primary care. Through qualitative engagement with practicing primary care clinicians in Ontario, this research explores the design considerations necessary to address the various concerns that may hinder the adoption of AI-enabled tools in primary care clinical interactions.

Methods

This research employed a user-centered design approach, incorporating a two-phase semi-structured qualitative interview process with 14 primary care clinicians practicing in Ontario. In the first phase, scenario-based interviews were conducted in which clinicians interacted with a standard EMR encounter module, allowing for the development of an initial sequence model that mapped out the typical workflow clinicians follow during patient encounters.

In the second phase, clinicians engaged with AI-enabled mock-ups with the same scenarios as the first. This mock-up included an AI-generated summary encompassing a numeric confidence score, the concept of approve/edit/decline buttons, and underlined actionable steps like selecting medications. Clinicians provided feedback on the AI-enabled EMR encounter mock-ups, highlighting aspects they liked, elements they found distracting, areas where they had reservations about AI inclusion, and other general sentiments. A thematic

analysis was conducted from the qualitative interview transcripts, which informed iterative redesigns of the AI-enabled interface.

These themes were further explored with the concepts of ease of use and usefulness through the Technology Acceptance Model (TAM), using these concepts to structure design requirements for revised mock-ups. A second round of semi-structured validation interviews was conducted to assess the effectiveness of the redesigned AI-enabled EMR encounter, with results tabulated to capture clinician feedback. Additionally, a System Usability Scale (SUS) was administered to quantify clinicians' perceptions of the redesigned interface, providing a standardized measure of its usability and potential for adoption in primary care settings.

Results

The initial sequence model was developed to map out a clinician's typical workflow during a clinical encounter, providing a structured understanding of key interactions from patient intake and history-taking to diagnosis, treatment planning, and documentation. This model highlighted key areas in the workflow where AI could support, rather than disrupt, established workflows by aligning with real-world clinician behavior. It also helped identify inefficiencies and moments where AI could provide meaningful assistance without diminishing clinician autonomy.

Through thematic analysis, six key themes were identified, contributing to a greater understanding of how AI-enabled EMR encounters can be designed for adoption, including trust and transparency, clinician authority, workflow efficiency, complexity sensitivity, legal and ethical concerns, and human-centered design.

Clinicians approach AI with skepticism, expecting trust to be earned through high performance and transparency. They prefer AI that reinforces their authority by providing relevant, non-intrusive support rather than acting autonomously. Trust is further influenced by clarity in AI-generated confidence scores and the ability to audit recommendations. Maintaining clinician decision-making authority is essential.

AI is viewed as most beneficial when offering supportive suggestions, such as reminders or preventative care guidelines, while avoiding directive or overly prescriptive actions.

AI adoption also depends on complexity sensitivity, where clinicians favor AI-assisted recommendations in straightforward cases like uncomplicated urinary tract infections but remain cautious in complex, nuanced scenarios such as mental health visits, uncovering a potential pathway for further exploration.

Workflow efficiency and seamless integration into clinical practice are critical, with clinicians expressing their desire for AI-generated summaries that are concise, patient-centered, and structured to minimize cognitive burden. Implementing unnecessary steps or disrupting existing workflows is seen as a barrier to adoption, as clinicians emphasize their time constraints and high-volume workload, requiring minimal disruption and patient care maintaining priority.

Legal and ethical concerns also play a significant role, with clinicians expecting clarity. Clinicians express delineation between AI-generated and clinician-modified content and well-defined policies on liability auditability.

Finally, human-centered design remains paramount, with clinicians emphasizing that AI should enhance rather than replace the clinician-patient relationship. AI should be positioned to operate unobtrusively in the background, ensuring efficiency without diminishing clinical judgment or disrupting the clinician and patient interaction during a visit.

These findings directly informed the redesigned screens, with ease of use and usefulness from the Technology Acceptance Model (TAM) serving as guides for the redesign

requirements. An "AI Assist" button was introduced, allowing clinicians to engage AI support at their discretion rather than imposing AI-driven suggestions on their workflow. This reinforced clinician autonomy and minimized cognitive burden. Additionally, color-coded delineations were implemented to clearly distinguish AI-generated content from clinician-entered data, ensuring transparency while maintaining workflow efficiency. An optional descriptive confidence score feature was incorporated, addressing varying clinician preferences by allowing them to toggle the feature on or off as needed. The language in the AI-generated summary was also adjusted to align with a more supportive narrative in contrast to an overly prescriptive tone.

These insights and the thematic analysis results guided the development of redesigned screens, positioning AI to better align with clinicians' preferences and expectations. The redesigned screens were completed in a way that attempted to reinforce trust, reduce friction between clinicians and AI, maintain flexibility, and reinforce a clinician's role as the decision-maker.

Validation interviews indicated that these design modifications potentially improved usability by positioning AI in the EMR encounter to better align with clinicians' preferences and expectations, as informed by the thematic analysis, ultimately increasing the likelihood of adoption.

Conclusions

AI adoption in primary care depends on thoughtful design prioritizing reinforced clinician autonomy, high performance, and seamless workflow integration. While clinicians recognize AI's potential benefits, they remain cautious about unintended consequences. Designing AI as a supportive, rather than directive, tool is key to fostering trust and improving adoption.

Future research should focus on real-world implementation, longitudinal studies on AI adoption, and regulatory frameworks that address liability and ethical considerations.

Contributions

This study contributes to the academic and practical understanding of integrating AI-enabled clinical tools in Canadian primary care EMRs. From a scientific perspective, it advances knowledge on the barriers and facilitators of AI adoption in EMRs, identifying key design principles that can encourage adoption.

From a practical standpoint, this study provides concrete design recommendations that can inform AI developers, EMR vendors, and healthcare policymakers. Additionally, this study contributes to the legal and ethical discourse surrounding AI in healthcare by highlighting unresolved questions regarding liability, data privacy, and patient consent. The findings call for regulatory frameworks that protect clinicians from undue legal risks while ensuring that AI systems are accountable, interpretable, and aligned with best medical practices.

Finally, this research offers a framework for future AI integration, emphasizing the need for context-sensitive AI models that adapt to case complexity. Through these contributions, this research informs the potential for responsible design and deployment of AI in primary care, ensuring that technology is a sustainable partner in healthcare delivery.

Acknowledgments

I am incredibly fortunate to have taken this journey with some of the best people.

I'd like to start by expressing my deepest gratitude to my research supervisor, Dr. Catherine Burns. You have been an incredible mentor and teacher and have demonstrated unwavering patience, kindness, and encouragement throughout this journey. Your guidance shaped my research and helped me preserve and maintain my passion for learning, even on the most challenging days.

To my collaborators at TELUS Health, including Dr. Puneet Seth, David Taylor, and Aaron Brady. I am incredibly grateful for the freedom and opportunity you provided to explore such an important and impactful topic, one with the potential to drive real-world change. Your support and insights have been invaluable in shaping this work.

To Sormeh Mehri and Sabrina Saiko, and others from the Advanced Interface Design Lab, I sincerely appreciate your hard work and collaboration. I am grateful for our shared knowledge and perspectives.

To my committee members, thank you for challenging me, refining my ideas, and helping me strengthen my work. Your perspectives have been instrumental in making this journey meaningful and worthwhile

Danilo & Imelda Francisco, the best mom and dad out there. Thank you for taking risks and getting us on a plane 20+ years ago to pursue the possibility of a better future in Canada. Your love and sacrifice are everything. To Adrian and Albert, my permanent cheerleaders in the form of brothers, thank you for never doubting my abilities. This climb was made better because of you. To my husband, John, your partnership and love have given me endless comfort and confidence to pursue my passions and dreams. And to my friends whom I consider family, a simple "thank you" could never fully capture my gratitude. Your support and encouragement have filled my cup daily, making this possible.

Dedication

To my family.

Table of Contents

Examining Committee Membership ii

Author's Declaration iii

Statement of Contributions iv

Abstract v

Acknowledgments xi

Dedication xii

List of Figures xvii

List of Tables xviii

List of Abbreviations xix

Chapter 1 1

Primary Care in Canadian Healthcare 1

Research Objectives and Questions 8

Overview of this Dissertation 10

Chapter 2 12

Methods 12

Modifications and Methodological Expansions 12

Modifications 12

Methodological Expansions 13

Background 15

TELUS Health Partnership 17

Phase 1: Initial Interviews 17

Scenarios with the Primary Care Encounter Module 18

Primary Care Physician Initial Interviews (Duration: 60 Minutes Total) 19
Table 1. Key differences between the summary and a traditional EMR encounter. 21
Phase 2: Analysis and Reporting 22
Phase 3: Develop User-Centered Design Requirements 23
Phase 4: Redesign of an AI-Enabled Encounter Module in the CHR 24
Usability Study 25
Phase 5: Validate in a Secondary Interview (Duration: 60 Minutes) 25
Recruitment of Participants 26
Ethical Considerations 26
Conclusion 26
Chapter 3 28
Understanding EMR Use in Primary Care 28
Mental Health Visit Sequence Model 29
Mental Health Visit Key Insights: 31
Uncomplicated UTI Visit Sequence Model 33
Uncomplicated UTI Key Insights: 34
Discussion 36
Conclusion 36
Chapter 4 37
Qualitative Analysis 37
Reflexivity statement 37
Positionality and Potential Bias 38

Conflict of Interest 39
Thematic Analysis 40
Data Saturation Reflection 41
Trust and Transparency 44
Clinician Authority and Decision-Making 50
Complexity Sensitivity 53
Workflow Efficiency and Integration 57
Legal and Ethical Concerns 62
Human-Centered Design 66
Discussion 68
Chapter 5 73
Redesign Process 73
Primary Care Clinician User Persona 86
Mock-Up Redesigns 87
1.0 Redesigned AI-Enabled Encounter Screen for Mental Health Scenario 88
2.0 Redesigned AI-Enabled Encounter Screen for UTI Scenario 93
..... 94
Discussion 95
Conclusion 97
Chapter 6 99
Validation Interviews 99
Discussion 106

System Usability Scale Analysis 109
Conclusion 112
Chapter 7 113
Contributions and Conclusions 113
Research Question Findings 113
Conclusion 121
Contributions 122
Bridging the AI Evaluation Gap 123
Aligning with Primary Care Principles 124
Scientific Contributions 125
Practical Contributions 127
Limitations and Future Research Directions 130
Conclusion 133
References 135
Appendices 145
Appendix A Recruitment Materials 145
Appendix B Pre-Read Material 147
Appendix C Oral Consent Script 150
Appendix D Thematic Analysis 153
Appendix E: System Usability Scale (SUS) 176
Appendix F Part A: Tabulated answers from validation interviews in detail. 179
Appendix F Part B: Tabulated answers from validation interviews in detail. 183

List of Figures

Figure 1 Dissertation outline 11
Figure 2. The encounter module in the TELUS CHR in an empty state 19
Figure 3. The AI-enabled version of the encounter in the CHR shown to clinicians during part 2 of the first interview 20
Figure 4. Visual sequence model of scenario 1 mental health visit. Green bars are activity chunks, blue is the trigger, red intents, and yellow breakdowns. 32
Figure 5. Visual sequence model of scenario 2 UTI visit. Green bars are activity chunks, blue is the trigger, red intents, and yellow breakdowns. 35
Figure 6 Primary Care Physician Persona 86
Figure 7.Redesigned encounter screen resting state mental health. 88
Figure 8. AI populated encounter screen mental health. 89
Figure 9. AI populated encounter screen with confidence score on 91
Figure 10. AI populated encounter screen with confidence score off 92
Figure 11. Redesigned encounter screen resting state – UTI scenario 93
Figure 12. AI populated encounter screen for UTI scenario. 94
Figure 13. “A comparison of mean system usability scale scores by quartile, adjective ratings, and the acceptability of the overall SUS score” by Bangor et al., (2008). 110

List of Tables

Table 1. Key differences between the summary and a traditional EMR encounter. 21
Table 2. Qualitative analysis. 22
Table 3. Sequence model of scenario 1 mental health visit 30
Table 4. Sequence model of scenario 2 urinary tract infection visit. 33
Table 5. Summary of Thematic Analysis 42
Table 6. Qualitative findings to design implications. 76
Table 7 Summary of Tabulated Clinician Responses Organized by Feature 102

List of Abbreviations

AI	Artificial Intelligence
EMR	Electronic Medical Record
TAM	Technology Acceptance Model
UTI	Urinary Tract Infection
LLM	Large Language Models
SOAP	Subjective, Objective, Assessment, Plan
PHQ9	Patient Health Questionnaire - 9
GAD7	Generalized Anxiety Disorder Assessment - 7

Chapter 1
Primary Care in Canadian Healthcare

In the Canadian healthcare system, primary care clinicians have played an integral role in providing continuity of patient care. Primary care clinicians are the first point of contact for patients entering the healthcare system; their scope of work includes illness prevention, referral to specialties, health promotion, diagnosis, treatment, rehabilitation, and counseling (Province of Ontario, 2024). To support their role as clinicians, the adoption of electronic medical records (EMRs) by Canadian primary care clinicians has grown significantly, from 37% in 2009 (Commonwealth Fund Survey, 2020) to 93% in 2023 (Johnson, 2023). An EMR can be defined as "a secure software system and its associated database. It is the record that primary care clinicians use within their practice environment to capture patient information such as a patient's family history, lab requests and results, screening tests, emergency room visits, prescriptions, and more" (OntarioMD, n.d.). Over time, EMRs have gained acceptance and credibility for improving a primary care clinician's ability to provide care (Ayanso et al., 2015; Jones et al., 2017; Raymond et al., 2015). They have become a standard part of care and are further supported through professional colleges and regulatory policies.

As EMRs have become a standard part of care, there is growing interest in further leveraging artificial intelligence (AI) to enhance their functionality. To fully realize AI's potential in this context, it is critical to explore how AI can be thoughtfully designed and integrated into the primary care setting, ensuring it complements clinical practices and ultimately supports the delivery of high-quality primary care (Ayanso et al., 2015).

AI in Primary Care: Challenges and Opportunities

AI offers vast potential for application, making it both an exciting frontier and a complex challenge to implement effectively. In healthcare, integration is further complicated by varying expectations and levels of comprehension among stakeholders. Understanding their perspectives, goals, and concerns is essential to aligning AI innovations with the needs of clinicians, patients, and policymakers.

With its broad scope and generalist approach, primary care presents unique challenges for AI adoption. Unlike specialties such as oncology or diagnostic imaging, where data is highly specific and narrowly scoped, primary care encompasses diverse conditions and patient needs. This requires AI systems to be designed to process highly variable, sensitive, and personal datasets while supporting clinical decision-making without undermining autonomy (Kueper et al., 2020; Baron et al., 2014; Henselmans et al., 2016; Tsai et al., 2022).

How AI interacts with clinicians has significant implications for behavior and decision-making (Gomez-Cabello et al., 2024). Ensuring human acceptance is critical as AI tools influence how clinicians interpret data and make decisions during patient encounters. Understanding these dynamics is key to identifying barriers and enablers of AI adoption in primary care.

These complexities highlight the need to explore technical and human-centered barriers to AI adoption in Canadian primary care, as examined in the following sections.

This study examines a conceptual generative AI application designed to assist primary care clinicians during patient visits by populating the EMR encounter with a summary of the most relevant clinical information. Unlike traditional rule-based or predictive AI tools, this application would leverage a large language model to generate concise, context-sensitive summaries of encounter data in real-time. The primary focus of this tool is to distill information that primary care clinicians would otherwise have to review manually.

Current State of AI Research in Primary Care

A scoping review by Kueper et al. (2020) highlights that AI research in primary care remains in its early stages, with most studies focusing on technical feasibility rather than practical implementation or clinician uptake. This gap underscores the need for further research into how primary care providers envision AI use and strategies to encourage its adoption. The review also emphasizes the absence of standardized evaluation frameworks for assessing

AI's impact on clinical outcomes, highlighting the need for robust methodologies tailored to healthcare to measure clinical effectiveness and workflow integration. Addressing these gaps is essential to advancing AI beyond theoretical applications and fostering tangible improvements in primary care. Additionally, the findings demonstrate that clinician engagement in designing and evaluating AI tools is crucial for ensuring acceptance in real-world settings.

The College of Family Physicians of Canada (CFPC) acknowledges AI's potential to enhance primary care delivery while emphasizing the need for its development to align with high-quality primary care principles (CFPC, 2024). The CFPC advocates for AI initiatives that improve healthcare service delivery, support physician well-being, and promote health equity. They stress that AI tools must be rigorously evaluated, user-centered, and sustainable. This highlights the CFPC's emphasis on ensuring AI enhances clinicians' ability to provide high-quality care (CFPC, 2024).

Despite the rapid advancements in AI for healthcare, its real-world impact remains constrained by sector-specific challenges and concerns about usability (Coulibaly et al., 2025). A recent scoping review of AI developments in Canada found that while AI models are being actively developed for patient triage, diagnosis, and care management, their integration into clinical workflows remains limited due to factors such as data quality issues, lack of clinician trust, and insufficient transparency. These findings align with themes generated in this study, which will be explored further in Chapter 4.

As AI adoption in healthcare accelerates, research has increasingly focused on understanding how these tools interact with primary care workflows. While many studies emphasize technical performance metrics, there is a growing recognition of the need for frameworks that assess AI integration within real-world clinical contexts. A recent conceptual framework (Choudhury, 2022) proposes an ecologically valid approach to analyzing clinician-AI interactions, incorporating human, AI, and contextual factors to better capture the complexity

of AI implementation. This aligns with broader shifts in AI research toward human-centered design and usability considerations, particularly in primary care settings where clinician workflows are highly dynamic. By acknowledging the interplay between user experience, AI functionality, and healthcare system constraints, this framework highlights the necessity of designing AI tools that are accurate but also intuitive and adaptable. Its emphasis on mixed-methods analysis further reinforces the need for qualitative insights into clinician perspectives, a focus that aligns with this study's thematic analysis. As primary care remains a crucial domain for AI integration, this framework provides a structured way to evaluate and optimize AI's role in supporting, rather than complicating, clinical decision-making.

Barriers and Facilitators to Adoption

Adopting AI in clinical decision-making is perceived as increasingly essential for improving healthcare outcomes. Hemmer et al. (2022) conducted a qualitative study identifying six core factors influencing human-AI collaboration: complementary AI capabilities, mutual learning through feedback loops, user adaptability across clinical roles, decision transparency, time efficiency, and maintaining human agency. These findings highlight the multiple considerations from industry and researchers that shape AI's integration into clinical decision-making, particularly as it becomes an increasingly powerful tool in comprehension and reasoning.

One area where these factors have been extensively explored is Clinical Decision Support Systems (CDSS), which have been integrated into primary care EMR systems since their inception. Gomez-Cabella et al. (2024) conducted a systematic scoping review examining six AI-CDSS implementations across diverse settings, including the US, Netherlands, Spain, and China. Their findings demonstrate AI-CDSS's potential to enhance diagnosis, treatment, and adherence to clinical pathways. However, adoption has been inconsistent, primarily due to physician skepticism regarding workflow integration and explainability. The study also highlights the need for greater transparency in algorithm design and stronger collaboration between developers, policymakers, and clinicians to address these barriers.

Darcel et al. (2023) explored barriers to AI adoption in Canadian primary care. Through virtual deliberative dialogues and thematic analysis, they identified key obstacles, including system readiness, algorithmic bias, and inequities in datasets. Trust emerged as a foundational element, with participants advocating for participatory co-design and iterative implementation to build confidence in AI tools (Darcel et al, 2023). The study also highlights the importance of investing in interoperability and equity-focused design to ensure broad and effective adoption. These findings reinforce the value of engaging primary care clinicians in evaluating and refining AI-enabled EMR tools. This process aligns with the participatory co-design approach to developing AI tools that can seamlessly be responsive to primary care clinicians' needs within their existing workflows. Building on this, Hassan et al. (2024) identified key themes critical to successful adoption, including transparency, explainability, governance frameworks, and human-AI teaming. Addressing algorithmic bias, ensuring data quality, and conducting rigorous evaluations are essential to building trust and demonstrating the value of AI tools. These studies advocate for balancing automation with human oversight to foster user acceptance and broader integration. Notably, these studies highlight themes that align with those generated in this study, suggesting a shared emphasis on trust, usability, and human-centered design to encourage AI adoption. Both studies emphasize balancing automation with human oversight to foster user acceptance and broader integration.

Recent literature further reinforces the complexity of AI adoption in clinical settings. A randomized controlled trial found that while large language models (LLMs) outperformed individual physicians in diagnostic reasoning, they did not significantly improve clinician decision-making compared to conventional resources (Goh et al., 2024). This suggests that AI's technical superiority alone is insufficient for adoption, emphasizing the need to align AI with clinical workflows and involve clinicians in shaping its effectiveness in practice. Similarly, another study evaluating GPT-4 assistance in clinical management reasoning found that AI-assisted physicians performed better on expert-developed scoring rubrics, supporting AI's potential to enhance complex decision-making (Goh et al., 2025). However,

AI-assisted clinicians also spent more time per case, raising concerns about workflow inefficiencies and over-reliance on AI-generated recommendations. These findings highlight key tensions between AI's benefits and its practical implementation, reinforcing the need to balance decision-making support with efficiency and cognitive burden.

Building on these challenges, Agarwal, Dugas, and Gao (2023) also identified trust and workflow integration as critical barriers to AI adoption. Physicians remain skeptical about AI's ability and ethical integrity, often fearing that it may compromise clinical judgment rather than support it (Agarwal et al., 2023). These concerns continue to support the notion that the design of AI tools is just as crucial as the technological capabilities they are often marketed for.

Expanding on these concerns, Ullah and Ali (2025) conducted a systematic review analyzing 51 studies on AI's role in healthcare, examining its impact on diagnostics, risk assessment, and clinical workflows. While AI applications demonstrated potential in improving efficiency and decision-making, the review found that adoption challenges, such as skepticism among experienced physicians, workflow disruptions, and lack of transparency, continue to hinder widespread implementation (Ullah & Ali, 2025). The review also highlights the lack of high-quality randomized controlled trials (RCTs) assessing AI's real-world effectiveness, reinforcing the need for more rigorous evaluations to determine the true impact of AI in clinical practice (Ullah & Ali, 2025). These findings support the broader argument that AI's integration into healthcare must go beyond technical advancements and, more specifically, should prioritize clinician trust and evidence-based implementation strategies.

Liu et al. (2024) conducted a longitudinal study evaluating the impact of AI-powered ambient clinical documentation on clinician efficiency. The study assessed the use of Nuance's Dragon Ambient eXperience (DAX) Copilot across 112 primary care clinicians, comparing them with a control group of 103 clinicians not using the tool. While exploratory

analyses suggested small reductions in documentation time for high-usage subgroups, the study found no statistically significant improvements in EMR use or financial outcomes across the broader sample (Liu et al., 2024). The findings highlight the limitations of AI-driven documentation in delivering widespread efficiency gains, reinforcing the need for tailored implementations that align with clinician workflows and support adoption at scale.

These studies collectively highlight the multifaceted challenges of AI adoption in clinical settings. While AI has the potential to enhance efficiency and decision accuracy, its successful implementation hinges on addressing skepticism. Key strategies include ensuring transparency, fostering trust, and prioritizing seamless workflow integration. The evidence suggests that AI must be developed and deployed with a strong focus on clinician trust, usability, and real-world effectiveness rather than relying solely on technical advancements as the primary driver of clinician adoption.

Healthcare Provider Perceptions of AI

Understanding the perspectives of healthcare providers and staff is crucial to successfully integrating AI into primary care. Nash et al. (2023) explored these perceptions through qualitative interviews within community health centers (CHCs). Participants expressed mixed views on AI, with many appreciating its potential to reduce workloads but raising concerns about trust, ethical considerations, and privacy. Poor integration with existing workflows was a common frustration, highlighting the need for clear communication about AI's benefits and limitations to build trust.

Garies et al. (2024) similarly examined provider perceptions through semi-structured virtual interviews. While participants were enthusiastic about AI's collaborative potential, concerns about over-reliance on AI tools and the erosion of clinical judgment persisted (Garies et al., 2024). These studies highlight the importance of iterative implementation and targeted training programs to ensure AI tools complement rather than replace human expertise.

Together, these studies emphasize that AI adoption in primary care hinges on its technical capabilities and provider confidence in its role within clinical practice. While AI offers the potential to ease workloads and enhance collaboration, concerns about workflow disruptions, trust, and the preservation of clinical judgment remain significant barriers. Addressing these challenges requires a thoughtful approach. This research provides insight into how such an approach can be realized by engaging clinicians in evaluating and refining AI tools to align with their needs and workflows.

Research Objectives and Questions

Many studies indicate that while clinicians recognize potential benefits associated with AI, as discussed in the prior sections, clinicians are primarily concerned with trust and workflow integration (Nash et al., 2023; Garies et al., 2024; Gomez-Cabella et al., 2024; Hemmer et al., 2022; Hassan et al., 2024) This emphasizes the need for a user-centered design approach to help drive AI adoption in primary care by embedding AI within the context of clinical practice, iterative prototyping, and user feedback. User-centered design may provide a framework to align AI with clinicians' real-world workflows and expectations, which might help reduce usability barriers and enhance overall acceptance.

This dissertation explores the challenges of redesigning an AI-enabled EMR encounter with a user-centered approach. My approach incorporates clinician feedback and uses thematic analysis as a foundation to develop design ideas. In particular, I have focused on AI-driven care recommendations within the EMR because AI-based recommendations pose challenges for clinician acceptance (Sivaraman et al., 2023). AI-based care recommendations trigger a complex interplay between human judgment and machine input. Additionally, primary care clinicians have expressed concerns about the evidence base behind these recommendations, accountability of AI, and clinical workload, emphasizing the need for collaboration between healthcare professionals and technology organizations to prevent AI from disrupting care delivery (Razai et al., 2024).

Despite advancements in AI research, gaps remain in understanding how AI-enabled tools, such as EMRs, can be designed to align with clinicians' needs. This dissertation addresses key gaps in the literature by focusing on clinician perspectives, usability, and workflow integration, areas where prior studies have been limited. While existing research, such as Sivaraman et al. (2023), highlights clinician skepticism regarding AI's decision-making and transparency, this study engages directly with primary care providers through qualitative analysis to explore their perspectives on how to build trust with AI. Similarly, Razai et al. (2024) emphasize concerns about AI's evidence base, accountability, and workload; this research extends this by incorporating usability testing and validation interviews to systematically assess AI's clinical impact.

A persistent challenge in AI adoption is workflow integration, as noted in previous studies (Hemmer et al., 2022; Gomez-Cabella et al., 2024;), which found that AI systems often disrupt clinician workflows due to poor alignment with existing practices. This study addresses this challenge through a user-centered design approach, introducing AI-enabled concepts to primary care clinicians and using their feedback to directly inform how an EMR tool should be designed to serve their needs in real-world scenarios best.

Another critical gap in the literature is the ethical and legal implications of AI, including liability concerns and the need for human oversight (Naik et al., 2022). This research builds on this by exploring clinician perspectives on AI-generated content and transparency as practical solutions for accountability. Finally, usability issues remain a major barrier, with past studies (Scott et al., 2024; Hassan et al., 2024) noting that many AI systems are designed without sufficient clinician input, leading to poor adoption. This study addresses this by employing contextual design principles and mock-up testing, ensuring that AI features align with clinician expectations.

While previous research has identified concerns such as trust, accountability, and workflow disruption, this dissertation provides practical solutions that enhance AI usability, acceptance, and integration in real-world primary care settings.

This dissertation focuses on the following objectives:

To understand if a user-centered design approach can increase the likelihood of adoption of an AI-enabled encounter module embedded in a primary care EMR. To address this overarching objective, this study investigates the following sub-questions:

1. How would clinicians prefer to interact with an AI-enabled tool embedded in a module they use for daily practice?
2. In what scenarios do clinicians find AI most helpful, and under what circumstances might they perceive it as intrusive?
3. What design considerations are necessary to build trust, enhance usability, and ensure seamless integration of AI systems into primary care workflows?

Exploring these sub-questions will uncover what is important to primary care clinicians when introducing a technology intended to enhance their practice. In the context of the diverse objectives within the Canadian healthcare system, this inquiry can provide valuable insights into a) the expectations primary care clinicians have for AI and b) the specific benefits they prioritize, such as time savings, improved patient outcomes, or enhanced user experience.

Overview of this Dissertation

This dissertation begins by introducing the research context, supporting literature, objectives, and overall structure. Chapter 2 outlines the methodology and approach. Chapter 3 examines the current state of how clinicians approach a clinical encounter in the scenarios described in the previous chapter. Chapter 4 presents a thematic analysis of qualitative interviews with primary care clinicians, highlighting key insights. Chapter 5 explores ease of use and

usefulness through the Technology Acceptance Model (TAM), using these concepts to structure design requirements for revised mock-ups. This chapter also details the step-by-step integration of these requirements. Chapter 6 reports findings from validation interviews with clinicians and presents results from the System Usability Scale (SUS). Finally, Chapter 7 summarizes the dissertation's contributions to the broader body of research.

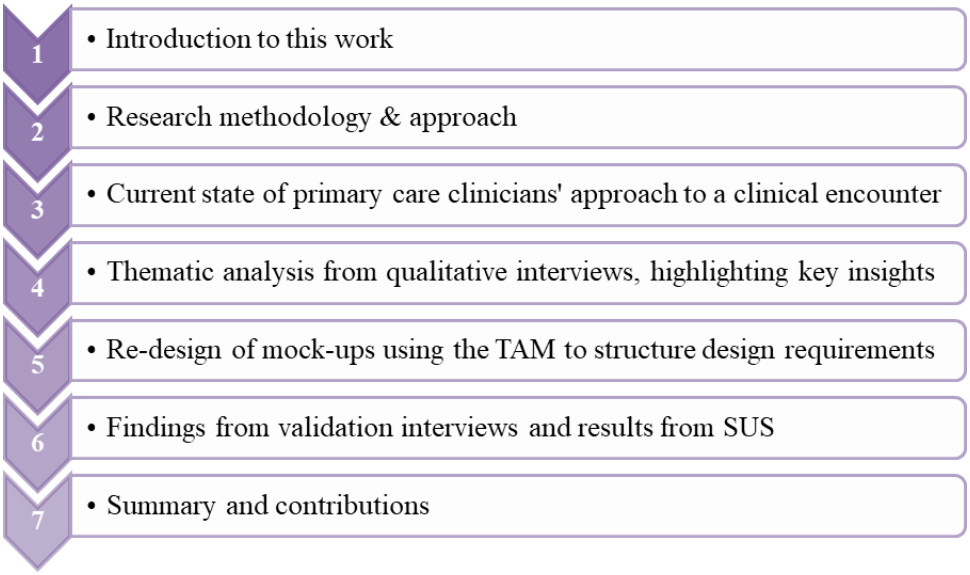
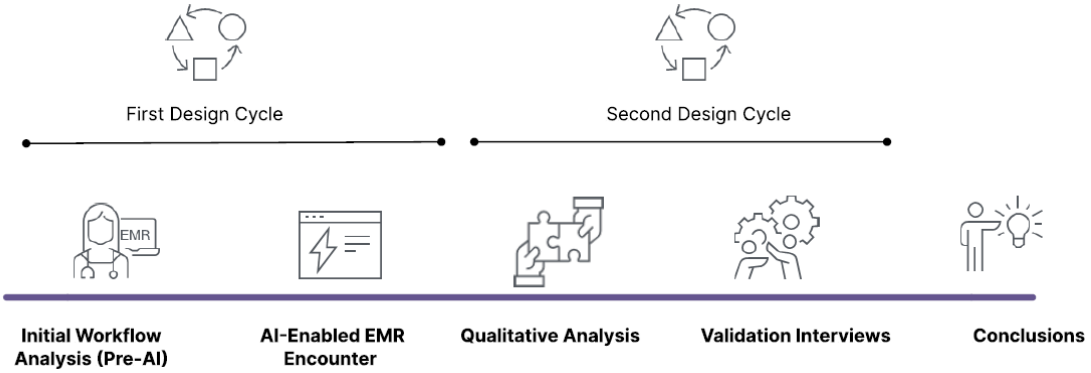
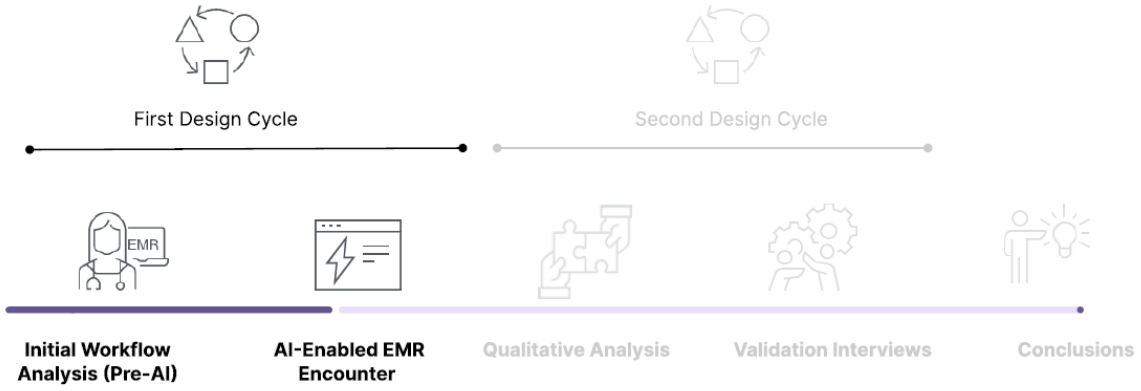


Figure 1 Dissertation outline

To help readers navigate the research and design process, the following diagram will appear at the beginning of each chapter. It will highlight completed sections, the current chapter, and upcoming sections to provide a clear progression of the study.



Chapter 2
Methods



This chapter is based on a previously peer-reviewed, published paper: Francisco, K. M., & Burns, C. M. (2024). *An approach to potentially increasing adoption of an artificial intelligence–enabled electronic medical record encounter in Canadian primary care: Protocol for a user-centered design. JMIR Research Protocols, 13.* https://doi.org/10.2196/54365.

Modifications and Methodological Expansions

This chapter includes additions beyond the scope of the originally published paper. While the core structure and findings remain unchanged, this section outlines post-publication modifications in addition to methodological expansions including an explanation of the triangulation approach used, rationale for semantic coding, and the design basis for the initial AI-enabled EMR encounter mock-up.

Modifications

- The original article stated that a deductive approach was used to sort data into categories established based on the TAM. While this applies to the redesigned mock-ups in Chapter 5, the analysis and reporting in Phase 2 (Chapter 4) followed a different approach. Specifically, the thematic analysis derived from qualitative

interviews was conducted inductively, without pre-established categories. This enabled theme generation as an analytic output by engaging the data and the codes generated during analysis.

- The two introductory paragraphs from the background of the publication were omitted from Chapter 2 as Chapter 1 was intended to be a more thorough introduction to the dissertation.

- While this study employs a user-centered design approach to iteratively refine AI-enabled EMR encounters based on clinician feedback, the broader methodological framework aligns with principles of Contextual Design. Contextual Design emphasizes understanding user needs within real-world workflows, an aspect that informed how design decisions were evaluated and structured in this research. However, direct field observations were not conducted, making this study primarily user-centered rather than fully contextual

Methodological Expansions

Triangulation

This study used triangulation to strengthen the credibility and trustworthiness of the findings, incorporating additional researchers and using complementary strategies like a system usability scale.

First, investigator triangulation was applied through collaborative analysis, where more than one researcher was involved in interpreting the data. A total of three researchers coded the data independently and met to discuss theme development regularly, which promoted reflection throughout the analytic process.

Second, method triangulation was used to cross-validate findings. In addition to the semi-structured interviews, participants were asked to complete a SUS to quantify their perceptions of the redesigned screens. Although there are limitations described on how SUS was used for this study, it still provided a structured, numerical compliment to the qualitative data and helped contextual participant attitudes towards the AI-enabled EMR encounter.

Finally, a set of validation interviews was conducted after the initial round of analysis. The feedback gathered during this phase further supported the coherence and relevance of the identified themes. Together, these strategies contributed to a well-substantiated understanding of clinician perceptions, supporting the credibility of this study's conclusions.

Semantic Coding Rationale

This study employed semantic-level coding as the primary approach to the thematic analysis. Semantic coding focuses on the explicit content of participants' responses (what was said) rather than interpreting latent or underlying assumptions. This approach was well suited to the applied nature of this study, which sought to inform future design decisions for AI-enabled encounter modules by grounding findings in clinicians' own words.

The goal was to understand how primary care clinicians perceive and interact with an AI-enabled EMR encounter, not uncover beliefs or broader social discourse. Semantic coding offered a more direct pathway from data to design implications, aligning with the study's user-centered goals and enabling findings to be translated into actionable insights.

Semantic coding also presents limitations. By focusing on what participants explicitly say, it risks overlooking implicit meaning or emotional nuance that may surface. This limitation made it important to incorporate safeguards against research bias and over interpretation. To address this, the study included multiple layers of reflection: a clear articulation of researcher

positionality, explicit recognition of the potential bias introduced by my professional role, and ongoing reflexivity throughout the design analysis and interpretation stages.

Design Basis for the AI-Enabled EMR Encounter

Details of the AI-enabled EMR encounter are presented in **Table 1** and **Figure 3.** This iteration was co-designed with a practicing primary care clinician who also served as a co-founder of the CHR. The design intent was to introduce a manageable number of AI concepts, enough to elicit meaningful feedback from clinicians without overwhelming them. Rather than presenting a full developed AI interface, the prototype was deliberately scoped to serve as a thinking tool: its primary purpose was to prompt reflection on how generative AI might support, disrupt, or augment clinical workflows.

Specifically, the prototype was designed to simulate how AI could summarize relevant patient information during a clinical encounter. By anchoring the design in existing EMR structures and clinician routines, the aim was to create a realistic, yet forward-looking, model that clinicians could respond to in a grounded way. The module's design functioned as both a testable artifact and a conversation catalyst, revealing the assumptions, expectations, and concerns that clinicians bring to AI integration in primary care.

Background

The contextual design process will be used as the foundation of this research. Contextual design was first invented in 1988 and is continuously used in various industries to drive innovative design (Holtzblatt & Beyer, 2017). Contextual design is a user-centered design process that uses techniques to analyze user data, drive ideation from the data, develop a design based on the data, and iterate these designs with the end users (Holtzblatt & Beyer, 2017).

There are 3 distinct phases to contextual design as follows:

Phase 1: Gathering user data. User-centered design recognizes that innovation starts with an understanding of the user by gaining an in-depth understanding of their tasks, motivations, intents, strategies, and detailed steps (Holtzblatt & Beyer, 2017).

Phase 2: Deriving insight from the data. Qualitative methods of understanding the data collected in Phase 1 are used to drive inductive insight into the design. This phase is further supported by other contextual design tools, like affinity diagrams or experience models, to better articulate user design requirements (Holtzblatt & Beyer, 2017).

Phase 3: Taking data to design. This key phase involves translating the data to drive design thinking and visioning. Prototypes and mock-ups are redesigned with users to better fit their activities (Holtzblatt & Beyer, 2017).

This process will be supported by the technology acceptance model (TAM) as a frame of reference. Although the TAM model has evolved in the last 30 years, the core of TAM remains focused on perceived ease of use and perceived usefulness (Davis & Granić, 2024). Perceived ease of use and perceived usefulness influence attitudes and intentions, which can influence and predict user acceptance and adoption (Davis & Granić, 2024).

There is great potential and significant risk to advancing AI technologies in health care. The risk of embedding AI in an EMR encounter is disrupting an actively accepted and used tool that they perceive as a benefit to their performance (Jones et al., 2017; (Raymond et al., 2015). Inappropriately introducing AI in the clinical encounter could impact its perceived usefulness, which could further deter the willingness and uptake of AI-enabled tools by primary care clinicians (Holden & Karsh, 2010; Davis, 1993).
For the context of this research, the users are primary care health care providers (eg, primary care physicians, and nurse practitioners) in Ontario, Canada.

TELUS Health Partnership

To accomplish this, a partnership has been established with an industry partner, TELUS Health, to use their EMR, the collaborative health record (CHR). The CHR is TELUS Health's cloud-based EMR solution. The CHR has standard features like the cumulative patient profile, free-text charting, custom forms, and electronic booking. However, it does have advanced functions, including streamlined clinical encounters, embedded virtual care, and real-time clinical and business intelligence.

This proposed work can demonstrate how to better optimize the use of AI-enabled products and services in the primary care setting, which is important given how the health care environment continues to evolve today. This can act as a key study that can inform some fundamental approaches to incorporating AI-enabled tools or services.

The approach of user-centered design requires qualitative interviewing to gain a clear understanding of users' approaches, intentions, and other key insights to inform the design process. This approach will be supported by qualitative methods to achieve the research objective. Several similar studies use a qualitative approach to user-centered design to inform technology acceptance in health care (Holden & Karsh, 2010; Zhu & Zhang, 2021; Vereenooghe et al., 2020; Abdullah et al., 2016; Yarbrough & Smith, 2007).

Preread material on AI and its capabilities within Canadian primary care has been developed for primary care clinicians to review in advance of the interviews.

Phase 1: Initial Interviews

All interviews will take place individually using a semi structured format comprising a standard set of scenarios with the ability to ask follow-up questions based on the participant's answer. This approach to interviewing enables a balance to provide purpose to the interview through the standard question set while also focusing the conversation on ideas and responses that can significantly inform the research and design considerations.

Several qualitative studies are conducted to inform contextual design in a clinical interaction. A number of these studies have demonstrated that anything over a total of 20 participants did not produce any significantly unique answers, which will inform our ability to reach data saturation with similar methods (Zhu & Zhang, 2021; Vereenooghe et al., 2020; Abdullah et al., 2016; McAlearney et al., 2004) For this reason, we will recruit up to 20 participants in total. The participants will be comprised of primary care clinicians currently practicing in Ontario, Canada. Each participant will be asked to participate in the interview structure outlined below, with the potential for subsequent interviews to be requested based on input gathered at the secondary interview, primarily focused on validating findings from the first interview. All interviews will occur virtually using Microsoft teams, where audio will be recorded and automatically transcribed through Microsoft teams. Immediately after the interview, the interviewer will correct minor errors in the transcription.

Scenarios with the Primary Care Encounter Module

The goal is to understand a clinician's decision-making process through a clinical interaction with a patient and what key elements of workflow exist (eg, when do they document the clinical encounter, what do they consider distracting on their EMR screen, etc). Given the breadth and ambiguity of presentation in primary care, specific scenarios have been chosen focused on general approaches to primary care physician workflows. The primary module within the CHR that will be used for these interviews is "Encounter." The Encounter module is where primary care clinicians spend the majority of their time documenting, even during a clinical visit. It captures information on patient responses to questionnaires, health history, examinations, assessment and planning, prescriptions, attachments, referrals, injections, follow-up items, and billing. Encounters require comprehensive information capture, making it an ideal module to focus on for added AI and its potential to impact how a physician captures a clinical encounter. Figure 1 shows the current structure of the Encounter module within the CHR.

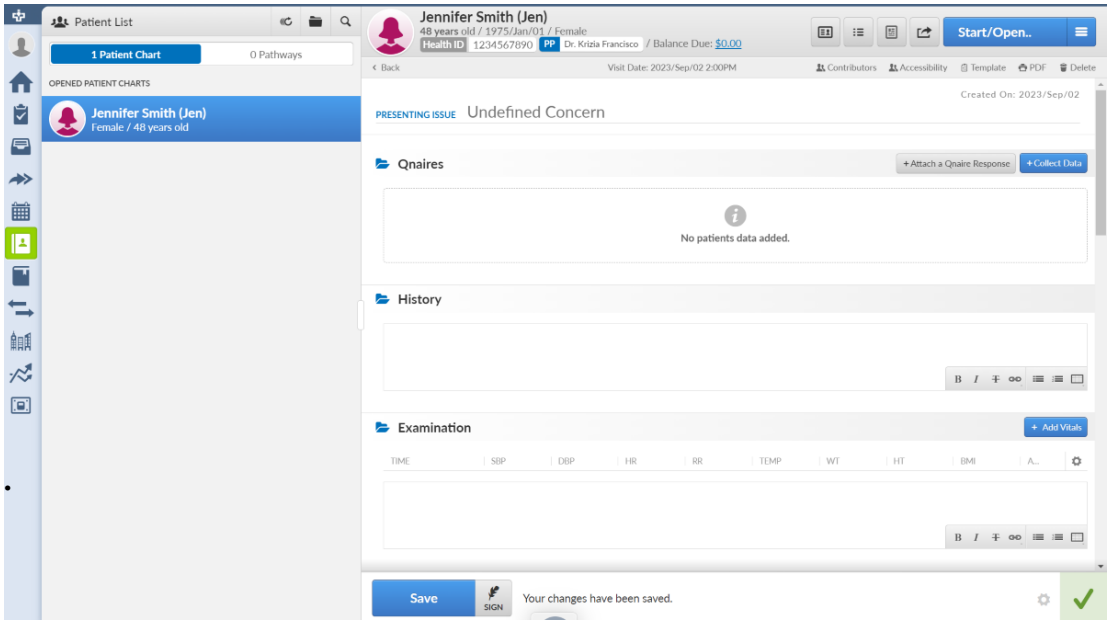


Figure 2. The encounter module in the TELUS CHR in an empty state

Primary Care Physician Initial Interviews (Duration: 60 Minutes Total)

Part 1 (30 Minutes) Current State: Clinician. There will be an initial interview with the primary care physicians to understand their current approach to documenting clinical encounters during a visit. The goal is to establish a workflow of what clinicians use in their existing EMR for that particular scenario.

Scenario 1: Taylor Jones, a female patient aged 29, called your office 2 weeks ago requesting a visit related to her mental health. In preparation for her appointment, she was given a GAD7 to fill out at home; the results were then sent to your office, with a score of >15, indicating severe anxiety. This patient is now sitting in your office, and the encounter screen is open on your screen. What flow would you engage in? (eg, which sections would you expand first, what would you chart, and what are the questions you ask next).

Scenario 2: Alice Smith, a female patient aged 48, called your office with a suspected UTI. The last time she was prescribed an antibiotic for a UTI was over 2 years ago. The patient is now sitting in your office with this encounter screen open on your screen. What flow would

you engage in? (eg, which sections would you expand first, what would you chart, and what are the questions you ask next).

Part 2 (30 Minutes; Scenario Exposure CHR Enabled With AI). Physicians will be presented with the same scenarios as in Part 1 of the interview, but now with an AI-enabled version of the CHR encounter. The objective is to understand how physicians will interact with the AI and whether it prompts differences in their workflow. Figure 2 shows example mock-ups that will be shown to primary care clinicians, demonstrating an AI-enabled version of the encounter module in the CHR. Scenarios 3 and 4 are the same as scenarios 1 and 2, except for adding an AI-enabled EMR encounter.

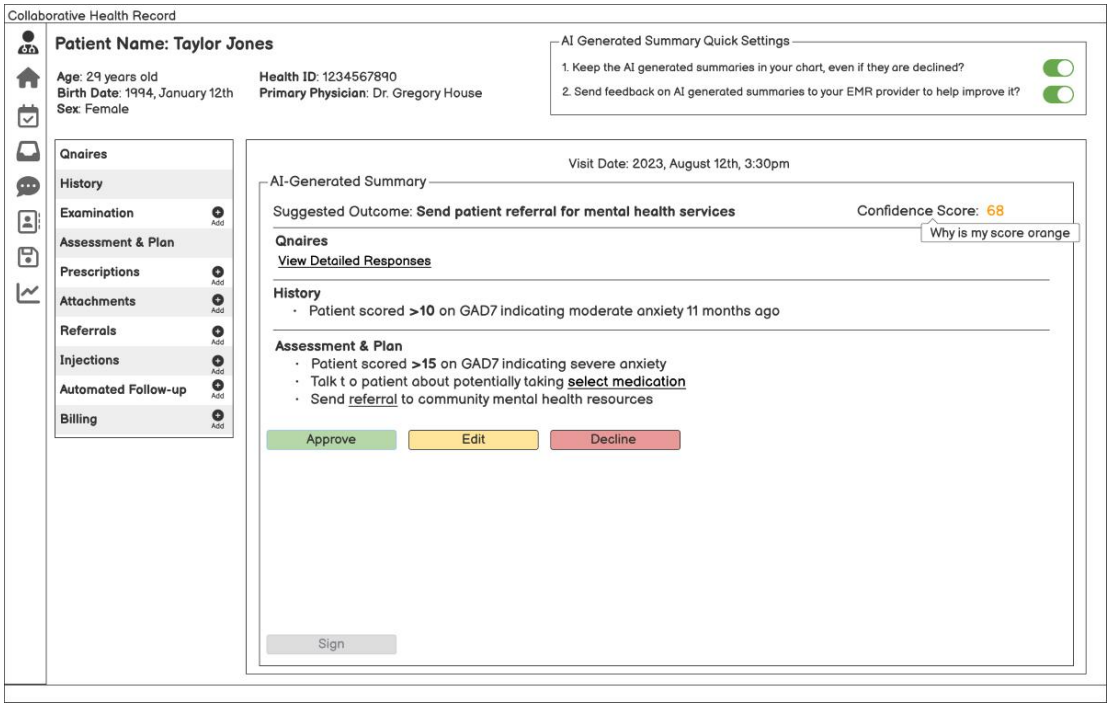


Figure 3. The AI-enabled version of the encounter in the CHR shown to clinicians during part 2 of the first interview.

Scenario 3: Taylor Jones, a female patient aged 29, called your office 2 weeks ago requesting a visit related to her mental health. In preparation for her appointment, she was given a GAD7 to fill out at home; the results were then sent to your office, with a score of >15, indicating

severe anxiety. Your EMR is now supported by AI, and before the patient's appointment, it processes the results and prepopulates your encounter. This patient is now sitting in your office, and the encounter screen is open on your screen. What flow would you engage in with the following screen? (eg, which sections would you expand first, what would you chart, and what are the questions you ask next).

Scenario 4: Alice Smith, a female patient aged 48, called your office with a suspected UTI. The last time she was prescribed an antibiotic for a UTI was over 2 years ago. Your EMR is now supported by AI, and before the patient's appointment, it processes the results and prepopulates your encounter. The patient is now sitting in your office with this encounter screen open on your screen. What flow would you engage in with the following screen? (eg, which sections would you expand first, what would you chart, and what are the questions you ask next).

This mock-up introduces the concept of an AI-generated summary for primary care clinicians. Key differences between this summary and a traditional EMR encounter are provided in **Table 1.**

Table 1. Key differences between the summary and a traditional EMR encounter.

- A numeric confidence score. It is intentionally unspecified whether or not the suggested outcome confidence score is based on a large language model data set informed by best practice, the clinician's practice, or the organization's data set, as this is also a key design consideration that can be further explored in the interview.
- This also introduces the concept of "approve, edit, or decline" buttons in the EMR encounter to further understand how primary care clinicians want to interact with AI from a liability lens (eg, if a clinician declines what the AI is suggesting, do they want it to remain a part of the record? Or, if a clinician decides to edit the summary, do they want the chart to delineate what was AI-generated versus what

was manually edited? Or do they view the note after it's been approved as their record is not distinguishable from what the AI suggested?).

- There are underlined actionable steps in this encounter (eg, "select medication"), demonstrating that the AI has not taken the further step of assigning a more specific medication to prescribe. Further probing clinicians on whether or not they think the AI has taken its suggestion too far, or not far enough.

Phase 2: Analysis and Reporting

The data collected throughout this research will be qualitatively analyzed concurrently with the interviews. This process aims to be deductive with the approach, supported by TAM, codes being sorted into one of 3 categories: perceived usefulness, perceived ease of use, and general attitudes towards using AI in health care practice. The coding techniques being applied are a result of the goals of the research being conducted. The codes discovered throughout the research will inform the various redesign approaches to an AI-enabled version of the CHR encounter. To provide further guidance related to AI-enabled primary care EMRs, key principles from design can be informed by a thematic analysis that results from the categories determined. A reporting guideline will be completed, the consolidated criteria for reporting qualitative research (COREQ), to further support the qualitative analysis (Tong et al., 2007).

The analysis will take place in the following way (Table 2):

Table 2. Qualitative analysis.

- **Preliminary coding:** At this stage, the coding does not have to be accurate or final. The goal is to familiarize myself with the data that have been collected. Immediately after an interview, I will conduct preliminary coding, highlighting any significant quotes. At this stage, capturing ideas for analytic consideration while the study progresses is essential. These are vital first impressions that can provide a transitional link between raw data and codes (Saldaña, 2021).

- **First cycle coding:** The goal of the first cycle, once all interviews have been conducted, is to provide an initial analysis. Semantic coding will be used to do this. Specific segments of the interview data will be sectioned and labeled using either a word or a short phrase from that particular section of the data. Semantic codes do not require any interpretation beyond what the participants have said. These codes can be viewed as a description that only represents the content of the data as it has been communicated by the participant (Saldaña, 2021). It is important at this stage that these initial codes provide enough context.
- **Sorting into categories:** Categories were established based on the TAM and are indicative of the consolidated meaning that is being analyzed in the data. I am hopeful that the categories developed will be informative about the various design changes that can occur with an AI-enabled version of the CHR encounter.
- **Thematic analysis:** The outcome of the second and third steps is theme development. Coding and theming are not mutually exclusive procedures (Saldaña, 2021). Themes are more descriptive versions of the categories discovered in the data. Theming data categorically provides descriptive detail about the patterns observed and constructed (Saldaña, 2021). With the thematic analysis becoming a macro-representation of the patterns, these can be translated into fundamental principles that inform the future design of other AI-enabled digital tools in the primary care setting.

The second and third steps in Textbox 2 are iterative and will be visited and then revisited numerous times until they become more refined and, as a result, more conceptual.

Phase 3: Develop User-Centered Design Requirements

The contextual inquiry methodology will be used to develop the user-centered design requirements. The principles of contextual inquiry enable the right type of interviewing to elicit design requirements, which include understanding the context, developing a partnership with

the users, and engaging in the interpretation process. This method requires the researcher to assume a role similar to an "apprentice," where those being interviewed are experts in their own experiences. The introduction of AI technology within the CHR encounter for these common primary care visits is just 1 key component of a much broader process and experience that needs to be understood and considered. Contextual design is driven by the fact that products are always part of a larger picture, making this a suitable method to complement the proposed research (Holtzblatt & Beyer, 2017). The researcher minimizes the introduction of bias by moving through a continuous interpretation process; by engaging users in the interview process in sharing interpretation, they are typically quick to correct any misguided or misinterpretations that are shared.

The interview will provide initial exposure to an AI-enabled version of the CHR. Participants will be prompted to provide feedback on what they do and do not like about the interface or what they do and do not find helpful as a part of the interface. The interviews will also highlight critical experiences from participants and, through the contextual design approach, reveal intentions of what is not explicitly being said. This information will be interpreted and translated into design requirements which the redesigned mock-up interfaces will be based on. This information will be used to build personas to inform how these designs will be mocked up.

Phase 4: Redesign of an AI-Enabled Encounter Module in the CHR

Both categories and themes identified in the thematic analysis will be used to inform the redesign of an AI-enabled encounter module within the CHR. The goal is for the data to inform user-centered design requirements that accurately reflect the motivations, thoughts, and attitudes of the particular audience or participant. For example, what features on the output may be seen as adding friction and not providing value or purpose that can be presented differently?

Redesigns of the AI-enabled CHR will be completed in a low-fidelity wireframe using Balsamiq, focusing on functional design rather than aesthetic design. After phase 4 is completed, high-resolution wireframes that are consistent with TELUS Health's branding will be completed. TELUS Health will be involved in the iterative design process and will review the mock-up iterations before engaging in a usability study to refine designs further.

Usability Study

A usability study is required at this point in the process to select which of the redesigned AI-enabled mock-up interfaces are most usable and accepted. Based on other usability studies conducted with similar methods, 8-12 participants are considered reliable (Alroobaea & Mayhew, 2014; Macfield, 2021).

Participants will be given access to interactive wireframes using Figma, accompanied by a system usability scale. As participants navigate the wireframes, open-ended questions will be asked to encourage further feedback on the redesigned interfaces. This system usability scale will be used to determine which of the redesigned outputs are best. Once this has been completed, the best designs will be mocked up using Figma, a higher fidelity software that can more consistently reflect TELUS Health's CHR initial user-interface branding. These will then be presented in the secondary interview process. This method is well supported and efficient and has a considerable amount of existing research to support its reliability.

Phase 5: Validate in a Secondary Interview (Duration: 60 Minutes)

Data gathered from the initial interview will be used to inform potential improvements that can be made. Alternative mock-ups will be redesigned, and presented to clinicians along with the scenarios they were presented during the initial interview. The goal is to understand whether or not the redesigned output was seen as an improvement from the initial designs in the first interview.

Recruitment of Participants

Targeted communications with well-known primary care networks will be used for recruitment (eg, primary care digital health networks, Association for Family Health Teams), in addition to direct emails to physicians or direct contact through social media platforms like LinkedIn. The inclusion criteria for this study related to clinicians is that they are primary care clinicians actively practicing in Ontario. In addition, any other requirements outlined by the University of Waterloo's Office of Research Ethics will be implemented upon review of this research protocol application.

All participating primary care clinicians will be provided a one-time $75 honorarium in the form of a VISA gift card in exchange for their valuable time and input.

Ethical Considerations

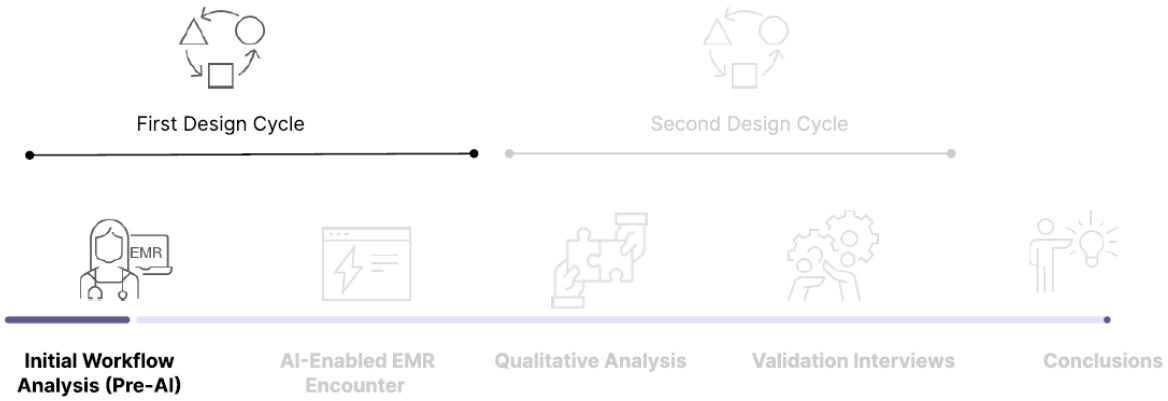
This study has been reviewed and received ethics clearance through the University of Waterloo Research Ethics Board (REB [43922]). This application includes the formal consent letter that will be used to support the study and reiterate the confidentiality of participation, in addition to an oral consent script to be used with participants. Participants will be informed that transcripts of their interviews will be recorded and stored on an encrypted hard drive for a 7-year retention period in alignment with the University of Waterloo's best practice guidelines. Based on the feedback provided by this office, there may be minor changes made to the approach of this research study.

Conclusion

This chapter outlined the methodology used in this dissertation and any modifications made after it was published. Chapter 3 applies the contextual design method of sequence modeling to analyze clinician workflows. This chapter presents findings from the initial phase of clinician interviews, where participants describe how they typically approach the two clinical scenarios discussed earlier. By mapping these interactions, the sequence models provide a structured representation of the processes, information flow, and key considerations in primary care encounters from a primary care clinician’s perspective. These insights serve as a

foundation for understanding existing workflow patterns and identifying opportunities for AI integration in clinical practice.

Chapter 3
Understanding EMR Use in Primary Care



A key component of this research was developing and analyzing a sequence model to examine how primary care clinicians conduct patient visits. As described in the methods chapter, Part 1 of the interviews involved presenting clinicians with a mental health scenario and asking them to walk through how they would conduct the visit. They were then given an uncomplicated UTI scenario and asked to do the same. To facilitate discussion, **Figure 2** was provided as a reference, allowing clinicians to articulate the actions they would take using a familiar EMR system. As a core element of contextual design, the sequence model captures the triggers, intentions, and step-by-step actions clinicians take, both with and without technology. In this study, it was used to define the as-is scenario based on data collected during Part 1 of the scenario-based interviews. To construct the sequence diagrams, all Part 1 interview transcripts were analyzed, with the data segmented into key processes and further broken down into discrete steps. Once developed, the diagrams were validated with two clinician participants to ensure accuracy; their feedback was minimal, with no refinements required upon review.

This phase of the research aimed to map out how primary care clinicians navigate a clinical encounter, making the sequence model the most suitable approach for capturing these structured workflows. The model provided valuable insights into clinicians' engagement with

EMRs across two distinct scenarios: a mental health visit, which introduces additional complexity, and a routine visit for an uncomplicated UTI. Despite differences in visit type, clinicians demonstrated consistent workflows and priorities, with minimal variability observed in their approach.

Beyond the scope of this study, other contextual design models could offer complementary perspectives on EMR use during clinical encounters. A flow model could have mapped the exchange of information between clinicians and staff, shedding light on how EMRs facilitate or obstruct communication (Holtzblatt & Beyer, 2017). A cultural model might have revealed institutional factors, such as resistance to EMR adoption or compliance-driven documentation habits (Holtzblatt & Beyer, 2017). An artifact model could have explored the impact of different interfaces, desktops, tablets, or paper-based backups, on usability and clinician efficiency. A physical model would have provided a spatial perspective, analyzing how technology placement in exam rooms affects clinician-patient interactions. Lastly, a consolidated work model could have synthesized multiple viewpoints to better understand decision-making, workflow patterns, and collaboration (Holtzblatt & Beyer, 2017). While this study centered on the sequence model to define the as-is scenario, these alternative models highlight the depth and versatility of contextual design in examining EMR use in primary care.

Mental Health Visit Sequence Model

The following section focuses on Scenario 1: The Mental Health Visit. To provide context, a detailed description of this scenario is included below as a refresher, with Table 3 presenting the corresponding sequence diagram.

Scenario 1: Taylor Jones, a female patient aged 29, called your office 2 weeks ago requesting a visit related to her mental health. In preparation for her appointment, she was given a GAD7 to fill out at home; the results were then sent to your office, with a score of >15, indicating

severe anxiety. This patient is now sitting in your office, and the encounter screen is open on your screen. What flow would you engage in? (eg, which sections would you expand first, what would you chart, and what are the questions you ask next).

Table 3. Sequence model of scenario 1 mental health visit

Trigger of this sequence
Complete a mental health visit
Activity: Pre-Appointment
Intent: Gather all necessary information for an effective and efficient visit
Review patient history
Check for prior mental health concerns or documented diagnoses
Understand if this issue has been previously addressed or ongoing
Review substance use history
Review any history relevant to MH including previous medications or therapy
Send GAD7 PHQ9 or some variation of it beforehand
Nurse takes vitals like blood pressure
Optional: Briefly review GAD7 and or PHQ9 scores and identify key concerns to guide questioning
Activity: During Visit
Greet patient
Intent: Build a comprehensive understanding of the patient's mental health needs
Open a SOAP note
Intent: Maintain a personable and empathetic conversation to encourage openness
BD: If patient expresses suicidal ideation, initiate crisis services
Engage the patient, use GAD7 and PHQ9 to guide the conversation
Intent: Prioritize active listening and patient engagement before documentation
Don't chart while patient is there
If lots of information flowing and need to type, try your best to maintain eye contact
Go through line of questioning understand holistic view of patient
Try and rule out other sources of anxiety and/or depression (e.g. anemia, thyroid etc)
Optional: If possible use real-time scribe to reduce documentation burden
Intent: Decrease documentation burden
Intent: Determine next steps based on comprehensive understanding the patient
Communicate findings, next steps, and follow-up actions
Provide materials or resources if needed
Optional: Begin referral process to other care if necessary
Book follow-up appointment with patient
Conclude visit
Activity: Post Visit
Intent: Ensure accurate and complete documentation to support continuity of care
Aim to write and complete notes before next patient
BD: Too busy to complete notes
Write down critical points immediately and fill in details with bulk charting at end of day

Mental Health Visit Key Insights:

Clinicians emphasized the importance of preparing for mental health appointments in advance to ensure that visits were both efficient and patient-centered. This preparation often involved reviewing patient history, prior diagnoses, and relevant mental health data, such as GAD-7 and PHQ-9 scores, to provide a well-informed foundation for the discussion.

During the visit itself, clinicians prioritized empathetic engagement, recognizing that building trust and openness was essential to effective mental health care. They were particularly mindful of maintaining eye contact and minimizing documentation during the conversation to foster a more supportive and non-disruptive environment. However, while real-time documentation was often reduced in favor of patient engagement, clinicians acknowledged that documentation remains essential for ensuring continuity of care and adherence to professional standards.

A recurring challenge across all clinicians was time constraints and competing demands, which often led to bulk charting at the end of the day. This practice, while sometimes necessary, carried the risk of compromising the detail and accuracy of notes, particularly in complex mental health cases. Despite these challenges, clinicians relied on structured tools like the GAD-7 and PHQ-9 to help guide discussions and support clinical decision-making during visits. Additionally, they placed strong emphasis on follow-up actions, ensuring that next steps were clearly communicated to the patient and properly documented, either during or after the encounter, to maintain high-quality patient care.

Figure 4 provides a visual sequence model of Scenario 1: The Mental Health Visit, illustrating the key steps and interactions within the clinical encounter. This figure corresponds to **Table 3,** which outlines the detailed sequence diagram, providing further insight into the structured flow of the visit.

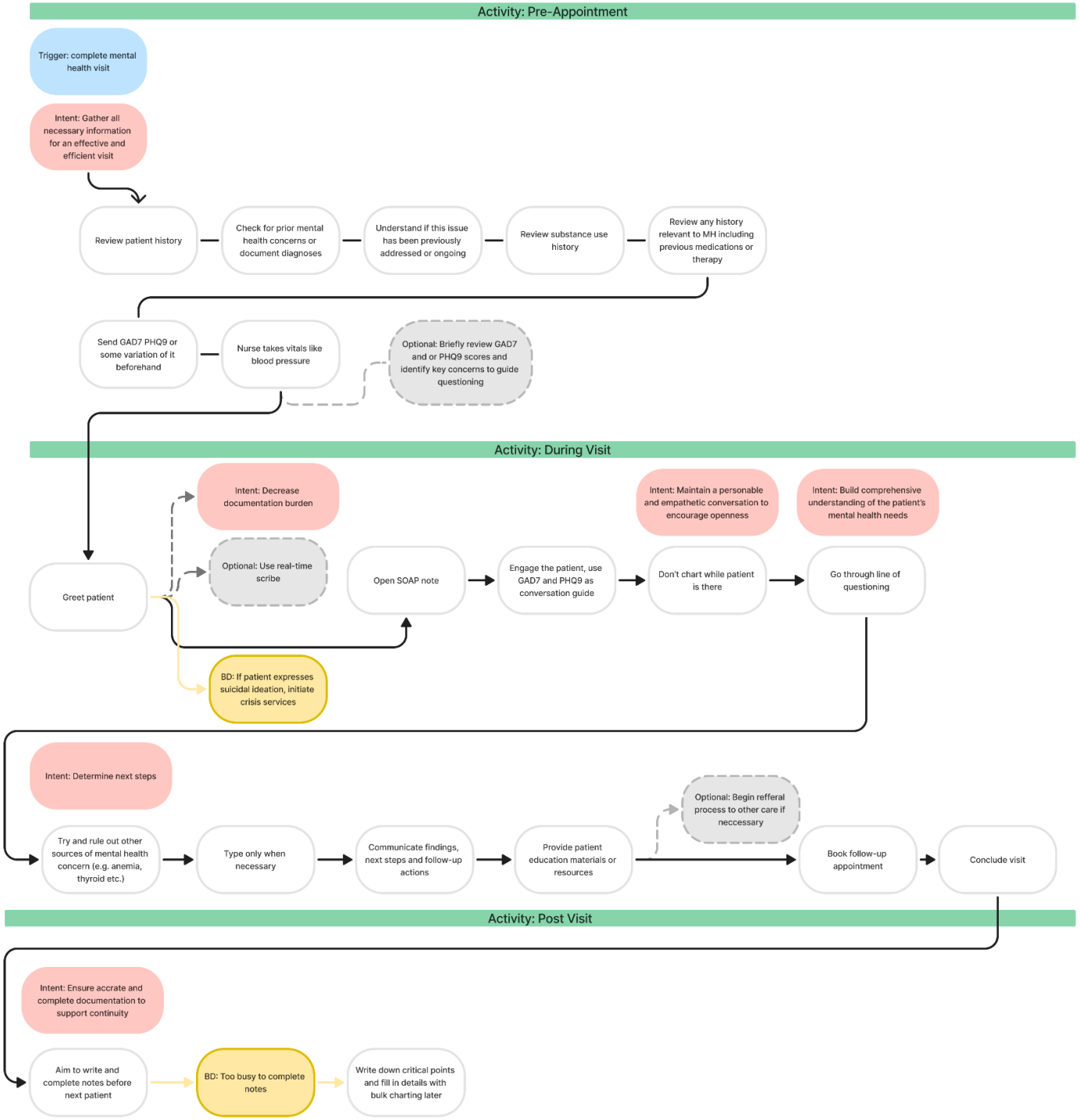


Figure 4. Visual sequence model of scenario 1 mental health visit. Green bars are activity chunks, blue is the trigger, red intents, and yellow breakdowns.

Uncomplicated UTI Visit Sequence Model

The following section focuses on Scenario 2: The UTI Visit. To provide context, a detailed description of this scenario is included below as a refresher, with Table 4 presenting the corresponding sequence diagram.

Scenario 2: Alice Smith, a female patient aged 48, called your office with a suspected UTI. The last time she was prescribed an antibiotic for a UTI was over 2 years ago. The patient is now sitting in your office with this encounter screen open on your screen. What flow would you engage in? (eg, which sections would you expand first, what would you chart, and what are the questions you ask next).

Table 4. Sequence model of scenario 2 urinary tract infection visit.

Trigger of this sequence
Complete an uncomplicated UTI visit
Activity: Pre-Appointment
Intent: efficient patient care for straightforward cases like this
Nurse does urine dipstick
BD: results aren't available pre-appointment
Nurse populates relevant history
BD: Upon review, case is straightforward
Skip the in-person visit and send prescription directly to pharmacy
Activity: During Visit
Greet patient
Intent: Confirm diagnosis and ensure efficient comprehensive care while avoiding unnecessary antibiotic use
Engage with patient to understand symptoms
Review history of prior UTIs, sexual history, or other relevant factors
Rule out other systemic symptoms - fever, chills
Asses symptom duration
Optional: conduct relevant physical exam
Optional: open SOAP note to capture findings or recommendations
Discuss hydration and antibiotic stewardship
Intent: Do not overuse antibiotics
Check past antibiotics prescribed and allergies
Send prescription directly to pharmacy for patient to pick-up
Conclude visit
Activity: Post Visit
Intent: Complete documentation
Complete billing of this visit
Charting UTI is a minute or two at most so do it quickly before next patient
BD: too busy to chart
Write down critical points immediately and fill in details with bulk charting at end of day

Uncomplicated UTI Key Insights:

In the case of an uncomplicated UTI, efficiency was the primary focus, with clinicians aiming to deliver care as quickly and seamlessly as possible. Given the straightforward nature of these cases, many clinicians expressed a preference for skipping in-person visits altogether when appropriate. Instead, if a patient's history and test results aligned with an uncomplicated UTI diagnosis, clinicians were inclined to send prescriptions directly to the pharmacy without requiring an additional consultation.

Despite the streamlined nature of charting for uncomplicated UTIs, which typically involves minimal documentation, busy schedules and competing demands often led clinicians to delay charting until the end of the day. While this did not significantly impact the efficiency of care, it underscored the broader challenges of balancing administrative tasks with clinical responsibilities.

Figure 5 provides a visual sequence model of Scenario 2: The UTI Visit, illustrating the key steps and interactions within the clinical encounter. This figure corresponds to **Table 4,** which outlines the detailed sequence diagram, providing further insight into the structured flow of the visit.

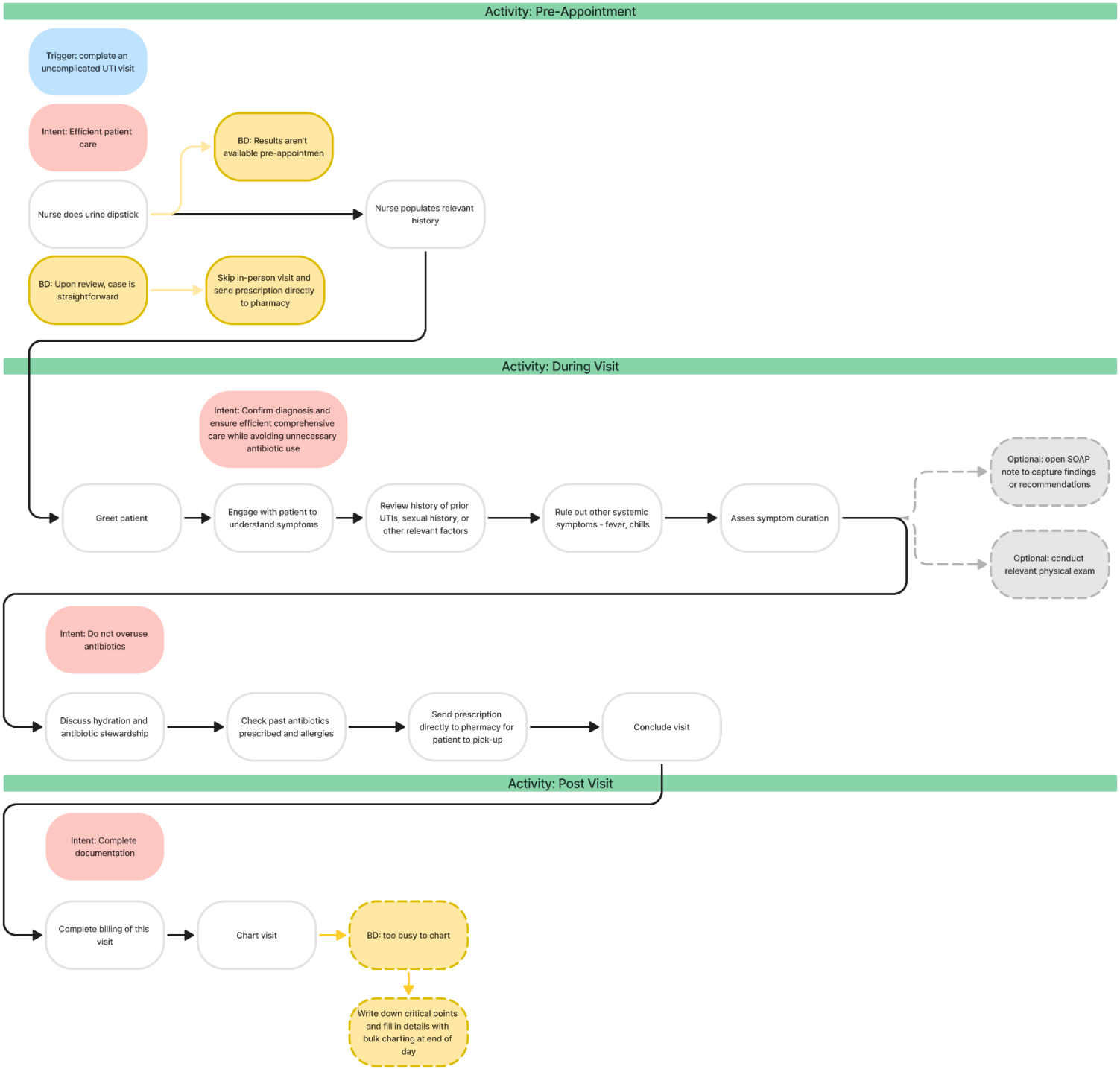


Figure 5. Visual sequence model of scenario 2 UTI visit. Green bars are activity chunks, blue is the trigger, red intents, and yellow breakdowns.

Discussion

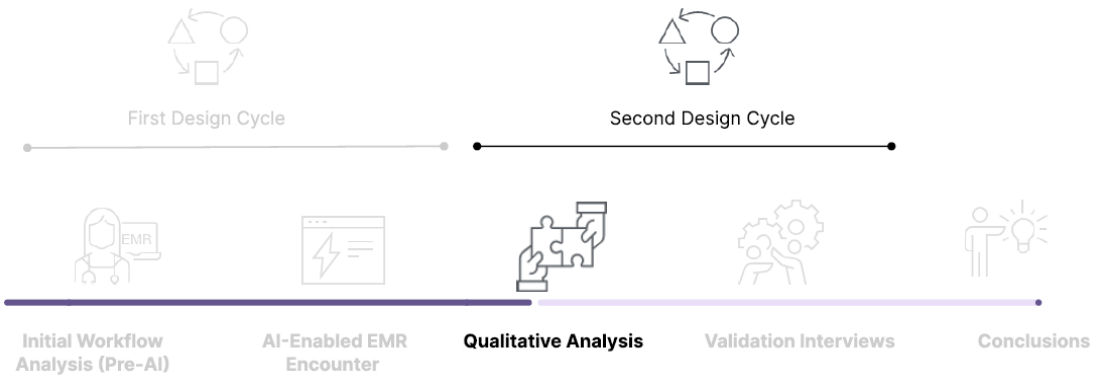
Both mental health and uncomplicated UTI visits required clinicians to balance efficiency and documentation, though the nature of each visit led to distinct approaches. In both scenarios, clinicians faced time constraints and competing demands, often leading to delayed charting until the end of the day. Additionally, ensuring continuity of care through proper documentation and follow-up actions was a shared priority across both visit types.

However, the two visit types differed significantly in workflow, patient interaction, and documentation complexity. Mental health visits required more preparation, with clinicians reviewing patient history and mental health screening scores (GAD-7, PHQ-9) in advance to guide discussions. These visits also emphasized empathetic engagement, where clinicians prioritized maintaining eye contact and minimizing real-time documentation to build trust. In contrast, efficiency was the central focus of UTI visits, with clinicians aiming to deliver care as quickly as possible. Many preferred skipping in-person visits in straightforward cases, opting to send prescriptions directly to the pharmacy when the diagnosis was clear. Additionally, while mental health visits involved structured tools to support some decision-making, UTI cases followed a more standardized and streamlined protocol, making charting quick and routine. These differences highlight how clinicians adjust their workflow, interaction style, and reliance on technology based on the complexity and nature of the patient encounter.

Conclusion

Understanding how clinicians navigate these two distinct clinical scenarios provided a critical foundation for assessing the potential impact of AI integration on their workflow. By analyzing their processes, documentation practices, and interaction styles, this section of this study provides valuable insights and considerations into how AI could be designed to support existing clinical processes.

Chapter 4
Qualitative Analysis



Reflexivity statement

As the principal researcher in this study, I acknowledge my dual role as both a senior product manager at TELUS Health and an independent researcher conducting this study. Given that TELUS Health served as the research partner, this position presented potential challenges related to positionality, bias, and influence over the research process.

To ensure research integrity and independence, I implemented several safeguards. First, a formal research agreement was established with TELUS Health, explicitly allowing me to operate independently for the purposes of this study and owning the results of this research. This agreement ensured that my role as a researcher remained distinct from my professional responsibilities. Additionally, to mitigate bias in data analysis, I incorporated two independent master's students as coders. They are members of the Advanced Interface Design Lab and the Department of Systems Design Engineering; this relationship was established for the purposes of this study. They conducted preliminary coding without my direct influence, contributing to triangulation and helping to ensure that the thematic analysis was driven by the data rather than my prior knowledge or expectations. Their input helped refine the

interpretation and organization of the data, enhancing the validity of the thematic framework.

By maintaining a transparent and structured approach, I aimed to minimize the impact of my professional affiliation on the study's findings while upholding academic rigor and objectivity. These reflexivity measures strengthen the credibility of the research and assure that the results reflect participant perspectives rather than preconceived notions or organizational priorities.

Positionality and Potential Bias

As part of maintaining transparency and research integrity, this section acknowledges my professional background and potential biases that may have influenced this study.

As a researcher I acknowledge my positionality and potential bias stemming from my professional role at a large EMR vendor that is actively exploring the use of AI in healthcare. My current position within TELUS Health involves strategic work related to the future of clinical data use and infrastructure. In this role, I am often exposed to forward-thinking discussions on how generative AI could be leveraged to enhance health system efficiency, improve the clinician experience or improve patient outcomes. This proximity to innovation naturally predisposes me to see AI as a powerful opportunity.

At the same time, I recognize the importance of academic research as a form of thoughtful pause and an opportunity to critically examine what can be built, in addition to what should be built. Primary care clinicians in Canada, especially those operating under increasing complexity, administrative burden, and time constraints,

require solutions that are not only technologically feasible but contextually appropriate and ethically sound.

I also acknowledge that working within a vendor organization could introduce a bias toward seeing AI-enabled tools as inevitable or intrinsically valuable. However, my experience has also made me deeply aware of the responsibilities of designing and deploying such technologies. At TELUS Health, we operate under rigorous frameworks for privacy, security, and consent, guided by the principle that data belongs to patients and is stewarded, not owned by vendors. This understanding informs my research lens: AI is not a shortcut to bypass clinical judgment, but a potential facilitator of thoughtful, consent-based, and secure data exchange in future care models.

Through this study I sought to remain grounded in the voices and lived experiences of clinicians, many of whom face substantial challenges in their day-to-day practice. By maintaining a user-centered approach, and by reflecting critically on my dual identity both as a health system employee and a researcher, I aimed to balance the promise of AI with a grounded understanding of its real-world implications in primary care settings.

Conflict of Interest

Although this dissertation was undertaken as a part of independent academic research, it is important to formally acknowledge the potential for a perceived conflict of interest stemming from my professional affiliation with TELUS Health. As an employe of an EMR vendor engaged in the development of AI-enabled health technologies, my role intersects with domains that are directly relevant to the subject

matter of this study. I held this role at the outset and throughout the duration of the research. While I remained committed to academic independence, I acknowledge that my professional experiences naturally informed how I engage with the topic, and it would be unrealistic to suggest complete detachment, free from bias, from the perspectives shaped by that context.

To mitigate this conflict, all aspects of this research including the design, recruitment, data collection, analysis, and interpretation were conducted independently and without the use of proprietary information or internal resources from TELUS Health.

This disclosure is intended to provide readers and future researchers with appropriate context for interpreting the study's findings and to uphold the ethical standards of academic research by acknowledging the dual roles I occupy as both a scholar and professional in the digital health sector.

Thematic Analysis

Fourteen actively practicing primary care clinicians from Ontario were recruited for interviews. The group included 10 male-identifying and 4 female-identifying physicians, with an average of 17 years of clinical experience. Participants' time practicing in primary care practice ranged from 5 to 37 years. Of the clinicians, 8 were affiliated with a Family Health Team (FHT), 2 with a Family Health Organization (FHO), 1 with a Family Health Network (FHN), 2 practiced independently, and 1 was associated with a Community Health Centre.

The research team conducted coding using QSR NVIVO 12 and Microsoft Excel, identifying 121 unique codes (Appendix D), which were then organized into 6 overarching themes and 22 sub-themes. An inductive thematic analysis was employed, enabling theme generation from the data and codes collected. In **Table 5,** I

provide a summary of these themes and sub-themes. Each main theme is discussed in greater detail following the table. Data saturation was reached after 14 interviews, indicating that no new patterns emerged with additional participants.

Data Saturation Reflection

As described in the background section, a target of 20 clinician participants was initially set to align with the expectations for reaching data saturation, based on sample sizes used in similar qualitative studies.

Despite extensive recruitment efforts, a total of 14 clinicians were interviewed. While this may appear to be a relatively small sample size, data saturation was reached. One possible explanation is that the AI-enabled encounter presented to participants was deliberately constrained in scope. By focusing on a limited number of features, the design narrowed the range of concepts clinicians needed to process and respond to, which in turn concentrated the analytic focus of the interviews.

Rather than seeking breadth across many AI functionalities, the study prioritized depth of reflection on a core interaction: generative AI to summarize clinically relevant information during an encounter. This narrow but targeted design likely contributed to early saturation as participants consistently returned to similar themes. Given the quality and depth of conversation were emphasized over quantity. Participants were given space to engage meaningfully with the concept, allowing the study to capture rich, experience-based insights within a manageable and focused sample.

Table 5. Summary of Thematic Analysis

Main Theme	Sub-Themes
Trust and Transparency	Clinicians' trust is linked to high-performance expectations of AI. AI can build trust with clinicians by reinforcing their authority and providing relevant, nonintrusive support. Transparency in how confidence scores were determined and training data could improve trust. Clinicians felt uncertain about unclear button functions associated with AI, which led to hesitation.
Clinician Authority and Decision-Making	AI can be helpful for supportive suggestions like reminders, preventative care guidelines, or flagging allergies. Suggestions should align with realistic, helpful actions that are evidence-based and practical. AI should offer useful suggestions without appearing to make autonomous decisions. Clinicians should retain final decision-making authority, and AI's role should be to support, not dictate or override clinical decisions.
Complexity Sensitivity	Context-sensitive suggestions: AI-suggested medications are acceptable for simple cases, while complex cases require heavier clinician discretion.

	Contextual information provided by clinicians in prior interactions is important. Clinicians want AI that understands past interactions and can capture nuanced contexts. Summaries need more comprehensive patient histories, with previous scores and relevant screening tools displayed.
Workflow Efficiency and Integration	Clinicians desire actionable, patient-centered summaries with appropriate depth and focus. Clinicians expect prompts for preventative measures (e.g. mental health risk factors). Clinicians are cautious of features that add workload or disrupt their workflow. AI needs to fit smoothly into the existing workflow without adding steps. Clinicians primarily saw the confidence score as a potential distraction but were curious about its utility, which may serve well in selective use.
Legal and Ethical Considerations	Clinicians require a visual representation that distinguishes what they say vs. what the AI generates before a note is finalized. Concerns on how legal frameworks are going to support the introduction of AI in clinical practice.

	Ultimately viewed AI as a tool that does not require direct disclosure to patients of its use.
Human-Centered Design:	AI must maintain a human element in patient interactions by not distracting clinicians and performing reliably. AI should support human-centered care without creating dependence on AI.

Trust and Transparency

Trust is widely recognized as a foundational construct in human-AI interaction literature. It has been conceptualized across disciplines as a dynamic, relational, and context-dependent phenomenon. In the context of AI in healthcare, trust is especially critical given the high stakes, professional liability, and deeply embedded norms around clinical autonomy and accountability. Trust can be defined as "the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control the other party."(Ueno et al., 2022).

This evolving view of trust is supported by both organizational and automation-focused literature. Schoorman et al (2007) describe trust as a willingness to accept vulnerability grounded in assessments of ability, benevolence, and integrity; three factors that influence whether a party or system is deemed trustworthy. These dimensions map closely to clinicians' evaluations of AI systems; technical competence, alignment with patient care goals, and conformance to professional norms. From a human-AI interaction standpoint

Kohn et al. (2021) emphasized that trust is shaped by a feedback loop involving observed system behavior, r8isk environment, and ongoing user experience. Trust builds or erodes over time as users refine their mental models of how the system behaves under varying conditions. Both frameworks highlight hat trust in AI is not binary or immediate; it is learned through reliable performance, predictable outcomes and systems that can communicate their internal logic clearly enough to support human oversight.

Clinicians in this study largely echoed the need for trust to be earned rather than assumed. Consistent with existing research, they expressed initial skepticism toward AI, highlighting that trust must be developed through consistently accurate performance, transparency in how decisions or summaries are generated, and auditability. Trust in this setting was closely tied to clarity of the interface and interpretability of outputs; clinicians preferred tools that made reasoning visible and allowed them to stay in control.

The following sections explore how these dimensions of trust and transparency were experienced and articulated by participants in relation to the AI-enabled encounter module.

Clinicians' trust is linked to high-performance expectations of AI.

A key factor in trust is AI's ability to perform at a high level. Clinicians define high performance as AI providing immediate, reliable support without adding cognitive or operational burdens. If AI consistently delivers accurate and helpful insights, clinicians are more likely to trust it, reducing the need for constant oversight. Given

the fast-paced and high-pressure nature of clinical environments, tools must integrate seamlessly without slowing down workflows.

However, high error rates can erode trust and slow adoption. Clinicians prioritize reliability over innovation, preferring proven and dependable AI systems rather than experimental features.

> *"It can't be performing worse than my own medical students." (P4)*

Another dimension of high-performance that was consistently mentioned was "AI accuracy" and how it is key to building trust. Clinicians have a low tolerance for AI errors, viewing accuracy as a baseline requirement for adoption. Errors raise concerns about the reliability and maturity of AI in healthcare.

> *"If I start declining what the AI suggests, I'd probably just turn the summaries off". (P4)*

> *"If it suggested something to me with low confidence, I would probably just turn the AI suggestions off, altogether" (P5)*

If AI provides incorrect or unreliable recommendations, clinicians will likely disengage from the system entirely, highlighting the need for consistently precise and well-calibrated suggestions.

AI can build trust with clinicians by reinforcing their authority by providing relevant, nonintrusive support.

AI can earn clinicians' trust by reinforcing their authority rather than undermining it. Trust grows when AI provides relevant, non-intrusive support, ensuring clinicians remain in control. Non-intrusive AI means that AI does not act autonomously without

explicit clinician input. For instance, automatically sending a referral without notifying the clinician would be considered intrusive. Trust is strengthened when AI operates as a collaborative assistant rather than an independent decision-maker, aligning with clinicians' expertise instead of competing with their judgment.

> *"It would be bad if this led to overprescribing or inappropriate referrals." (P2)*

Transparency in how confidence scores were determined and training data could improve trust.

Clinicians consistently emphasize the need for transparency in AI's training data, decision-making processes, and confidence scores. They expect AI to be accountable for its outputs, and clear explanations of how AI reaches its conclusions can help provide clinicians with peace of mind. Clinicians expressed that they did not need to see how AI arrives at a conclusion at all times but would want to have the option to view and audit this information as needed.

Transparency about training data was seen as a way to link AI recommendations to evidence-based practices or real-world clinical experience, making it easier for clinicians to trust its suggestions. More than half the participants mentioned that they would expect training data to be based on best practice guidelines.
Confidence scores must be introduced carefully, providing enough context to be useful without disrupting workflow efficiency or undermining clinician autonomy.

Clinicians felt uncertain about generic button functions, which led to hesitation.
Clinicians expect clarity and predictability in their tools. Unclear AI functionality, such as vague button labels, can create confusion and slow adoption.

For example, generic button labels such as “Accept,” “Delete,” and “Edit” caused hesitation among all participants:

> *“What happens if I press accept? Is it going to sign off for me?”(P1)*

> *“What am I accepting? The outcome? The diagnosis? It isn’t clear.”(P1)*

> *“The other two buttons are useless; it should always be ‘edit.’ What does ‘accept’ even mean?”(P4)*

> *“If I accept what the AI says, is it going to trigger another workflow I don’t know about, like sending a referral?”(P9)*

These concerns stem from uncertainty about AI’s influence on clinical decisions and patient safety. Clinicians fear unintended consequences, both positive and negative, when integrating AI into their EMRs. Their reactions range from skepticism to cautious optimism, reflecting the importance of clear policies, well-defined workflows, and intuitive AI interfaces to ensure trust and usability.

Key Takeaway

Trust in AI is not inherently granted, it is earned through consistent high performance, transparency, and alignment with clinicians' needs. Clinicians approach AI with skepticism, expecting it to enhance rather than disrupt their workflow. High accuracy, clear decision-making processes, and non-intrusive support are critical factors in fostering trust. When AI operates reliably, provides explainable and auditable recommendations, and respects clinician authority, it can become a valuable partner in healthcare rather than a source of uncertainty. However, poor design choices, high error rates, and opaque decision-making can erode trust and lead to

disengagement. As AI continues to evolve in healthcare, ensuring clarity, predictability, and user control will be essential in driving adoption and making AI an indispensable tool for clinicians.

Clinician Authority and Decision-Making

Clinicians are clear that they must remain the final decision-makers, with AI serving only as a supportive tool. Features like confidence scores and AI-generated suggestions should be presented in ways that reinforce, not override, clinician autonomy.

AI can be helpful for supportive suggestions like reminders, preventative care guidelines, or flagging allergies.

Clinicians could see value in AI augmenting their memory or catching overlooked details, viewing this function type of support as additive rather than directive. There was an acknowledgment that the breadth of issues that they are presented with on any given day is vast, making it extremely difficult to keep up with best practice guidelines that are constantly being updated. In addition, primary care clinicians appreciate if AI can act as a safeguard for guidelines they sometimes forget (e.g., checking if a female patient is pregnant before prescribing a specific antibiotic for a UTI).

> *"I think it would be helpful to have that little light devil's advocate, where it says some symptoms could also be caused by something I have not thought of" (P7)*

> *"One time, I had a patient come in, and she was over 80 years old; I had sent her for a referral, thinking she potentially had a less acute issue at the time. When I went home and got to bed, that's when I realized that it actually could have been something worse, and I needed to send her for an immediate referral elsewhere. It would be helpful if AI at the time could have made that suggestion to me instead of me trying to remember all these things." (P13)*

> ***"****You know, patient's here for a UTI, but then they also bring up their heart failure and their high blood pressure and their birth control or whatever" (P10)*

Suggestions should align with realistic, helpful actions that are evidence-based and practical.

Actionable insights reinforce clinicians' preference for AI to address real-world challenges directly and effectively, emphasizing practicality over theoretical or overly technical features. Clinicians supported the idea that the information they see on the screen needs to be succinct and practical. Reading large bodies of text or suggestions for services that aren't even available in a patient's immediate vicinity were seen as potential distractions that can compromise their effectiveness in a clinical visit.

AI should offer useful suggestions without appearing to make autonomous decisions.

A tension point for clinicians was when AI was perceived to act autonomously rather than foster collaboration. This further highlights the expectation that AI should support, not supplant, their expertise and decision-making.

> *"It would be good to look over a note and make sure you are agreeing with the AI before you approve." (P3)*

All clinicians clearly emphasized that they should retain final decision-making authority. AI was seen as a tool that should not be relied upon too heavily, as its quantitative ability should not overshadow clinical judgment. Two participants noted that AI-enabled tools like this in the context of training future primary care clinicians should be done with caution, with the fear of overdependence being an unintended consequence.

The expectation for clinician final authority underscores a theme of control, with AI's role firmly positioned as a supporting tool rather than a decision-maker.

Clinicians should retain final decision-making authority, and AI's role should be to support, not dictate or override clinical decisions.

Many clinicians wanted the ability to dismiss AI they disagreed with, reflecting their control in decision-making and autonomy. This further supports ensuring AI feels like a collaborative assistant rather than an authoritative entity.

The overarching desire for autonomy emphasized clinicians' need to feel empowered by AI, not constrained by it.

"Ultimately, the decision is for me and I'm the one saying yes or no." (P5)

Key Takeaway

Clinicians strongly affirm that AI should function as a supportive tool, not a decision-maker, reinforcing rather than undermining their authority. While AI has the potential to enhance clinical workflows by offering reminders, flagging overlooked details, and providing evidence-based suggestions, it must do so in a way that respects clinician autonomy. The value of AI lies in its ability to augment human expertise without imposing rigid directives or autonomous actions.

To be truly effective, AI must prioritize actionable, practical insights that align with real-world clinical needs. Overly complex or irrelevant suggestions can distract from patient care rather than enhance it.

AI should be a reliable safeguard, a helpful second opinion, and an efficient tool, but it should never substitute for human judgment. As AI continues to integrate into healthcare, maintaining clinician authority must remain a guiding principle for its development and implementation.

Complexity Sensitivity

Clinicians recognize the potential for AI to assist in medical decision-making but emphasize the importance of context-sensitive suggestions that adapt to the complexity of each case. While simple cases, such as uncomplicated UTIs, were considered well-suited for direct AI assistance, more complex cases, like mental health conditions, require clinician discretion and heightened sensitivity. Clinicians expressed a desire for AI systems to have the ability to account for past interactions and patient histories, ensuring continuity of care and providing more relevant recommendations. Ultimately, AI was seen as effective in a clinical encounter by tailoring its involvement based on case complexity and patient context.

Context-sensitive suggestions: AI-suggested medications are acceptable for simple cases, while complex cases require heavier clinician discretion.

Clinicians want AI involvement tailored to the complexity of cases. Simple cases like the UTI scenario are suitable for more direct AI assistance (e.g., suggesting straightforward treatment pathways). At the same time, complex cases like mental health are too complex to rely on AI-generated summaries. As seen in the existing sequence models, mental health requires more clinician input to reach the next steps that may not even resolve the issue within that visit. The distinction between simple and complex cases underscores the clinician's role as the ultimate decision maker, requiring the use of AI to be flexible when they consider it appropriate.

> *"If AI can help me with the hard decisions that aren't intuitive, that would be fantastic" (P1)*

> *"Mental health is a whole different story" (P14)*

"I *have this patient; he's been with me his whole life. Every few years, he complains about his mental health, but I know he has very turbulent relationships. To AI, it may seem like he has depression, not realizing that this person typically has low-mood episodes because of life circumstances. I'm not saying he shouldn't be here for his mental health; I'm just saying as his primary care physicians, I can figure out whether he needs someone to listen to him or an actual prescription, and I don't know if AI can do that". (P12)*

Recognizing variations in complexity across conditions, clinicians have different expectations for AI's involvement. The interviews highlighted the clear difference in approaches for complex vs. simple cases for clinicians. Mental health cases are seen as too nuanced for AI, whereas more straightforward cases like UTIs are more acceptable.

Contextual information provided by clinicians in prior interactions is important. Clinicians want AI that understands past interactions and can capture nuanced contexts.

Clinicians value AI systems that enable the continuity of care they are expected to provide, showing an understanding of individual patient histories and nuanced relationships within clinical contexts.

> *"Nuance always needs to be accounted for; it is actually very important when going through a clinical encounter" (P2)*

> *"If it is mental health, then I'd need to know the details of if this has happened before, if they're declining, or if this is a brand new complaint they're seeing me for. The nuance and history of their mental health matters." (P13)*

The ability to "remember" contextual details from past interactions

demonstrates an AI's adaptability and relevance, increasing its perceived usefulness. Clinicians expect the AI to align its actions with their prior decisions, making it feel more intuitive.

> *"A tool can indicate severe anxiety, but I will want to know what the contextualization is with their past mental health history. Has it happened before? Is this a new thing?" (P5)*

Summaries need more comprehensive patient histories, with previous scores and relevant screening tools displayed.

Primary care clinicians place a high value on historical continuity. Their core role within the healthcare system is to provide longitudinal continuity of care, ensuring that patient history, evolving health needs, and past interventions inform ongoing decision-making and long-term management. Comprehensive patient records are essential to contextualizing AI suggestions. The inclusion of historical data supports clinicians' analytical processes and decision-making. Summaries must complement how clinicians engage in their natural thought processes to further foster adoption and usability.

Mixed preferences on medication suggestions, depending on the complexity of the case.

Preferences for medication suggestions reflect clinicians' nuanced approach to balancing efficiency with accuracy. Clinicians view medication suggestions as a form of support to enhance their decision-making, not as directives. Clear, evidence-based recommendations are particularly valuable in straightforward cases like treating an uncomplicated UTI. However, greater caution is required when suggesting medications for more complex and nuanced conditions, such as mental health, where clinician discretion and judgment are paramount. One participant did note their skepticism about AI-generating medication suggestions and their concerns about

pharmaceutical bias and whether or not the vendor had pre-existing motives or relationships that may influence the medications being prescribed.

> *"AI suggesting medications would be helpful" (P6)*

> *"Learn what I prescribe. Especially for something straightforward like a UTI." (P10)*

> *"Is there bias in what medications are being suggested? I would worry about EMR vendors potentially pushing a specific pharmaceutical" (P8)*

> *"If we're talking mental health, I'm not visiting medications in the first visit". (P4)*

Key Takeaway

AI's role in clinical decision-making should be context-aware and adaptive while offering evidence-based recommendations without overriding clinician judgment. While AI can enhance efficiency and provide valuable insights, its effectiveness depends on ensuring transparency, maintaining clinician authority, and respecting the nuances of patient care. Thoughtfully integrating AI into clinical workflows while recognizing its limitations will be key to fostering adoption and gaining trust with primary care clinicians.

Workflow Efficiency and Integration

Clinicians seek actionable, patient-centered AI summaries that provide the right depth of information without overwhelming them. AI should present usable insights that align with their natural reading flow and charting standards, ensuring efficiency in real-time decision-making. Clinicians value concise, visually organized summaries with features like color coding to highlight key information. However, they are cautious about anything that adds workload or disrupts workflow, emphasizing the importance of seamless AI integration that minimizes unnecessary steps. Additionally, the ability to personalize AI tools enhances usability, ensuring AI functions as a collaborative assistant rather than a rigid system.

Clinicians desire actionable, patient-centered summaries with appropriate depth and focus.

Actionability was a critical factor for the designs presented. Clinicians prefer AI to present usable insights directly tied to what can be done for a patient in an interaction (e.g., if an AI-generated recommendation was to send a patient for a referral, have the referral pre-populated contextually within the EMR). Overloading information risks diminishing AI's perceived value in real-time decision-making, further reinforcing that the only information that should be presented to clinicians is what is immediately relevant to their context at that time.

Some design considerations consistently mentioned include the desire for color-coded summaries that align with clinicians' reading flow and charting standards. Clinicians want concise, visually organized information that complements their natural reading and decision-making flow (e.g., SOAP format (Subjective, Objective, Assessment, Plan). Color coding can help them quickly prioritize key information, emphasizing clarity and efficiency to make actionable decisions.

"Color coding - Probably if there was something like if it was a different color font while you're still editing it, and then once it's done and you approve it at all, looks the same that would, that would probably be helpful." (P10)

"I think in SOAP format, so the tool would need to align with that. That's why I use stamps now" (P3)

Clinicians expect prompts for preventative measures (e.g., mental health risk factors).

Primary clinicians highlighted the importance of offering preventative care and not being limited to treating chronic diseases within their practices. Preventative care suggestions align with primary care's goals of delivering forward-looking patient care. If AI can offer proactive suggestions, it could enhance long-term patient health.

Half of the participants expressed that their role is primarily responsive, focusing on addressing immediate patient concerns during visits. However, they also conveyed a strong desire to shift toward proactive care, emphasizing wellness conversations and preventive strategies to reduce the onset of disease, rather than primarily managing chronic conditions once they develop.

"If you diagnose someone with anxiety, lifestyle measures are always a discussion, counseling is always a discussion." (P10)

"There needs to be the flavor of AI that will mention wellness and well-being." (P6)

"If someone is here for something like a straightforward UTI and I know I'm going to have some extra time in the visit, I will definitely use that as an opportunity to talk about other preventative stuff I don't usually get to". (P12)

Clinicians are cautious of features that add workload or disrupt their workflow. Features that disrupt workflows are viewed as barriers to efficiency, reinforcing clinicians' need for AI tools to integrate seamlessly into their routines without adding cognitive or procedural burdens.

> "*Don't give the doc more things to do, please." (P1)*

Even the addition of new buttons on their existing encounter screen was seen as having no added value or usefulness. Conversely, the association of the buttons with AI triggered hesitation and concern, potentially adding unnecessary cognitive load where it would not have otherwise been present. Clinicians perceive unnecessary steps as counterproductive, reflecting streamlined processes prioritizing efficiency over redundant interactions. Resistance to unnecessary clicks underscores clinicians' broader need for intuitive systems that respect time constraints and minimize workflow friction.

> *"In the interest of time, unless this adds value, I don't need extra clicking". (P9)*

> *"No one's going to use it if there's too much stuff". (P6)*

AI needs to fit smoothly into the existing workflow without adding steps. Clinicians expect AI to enhance efficiency without adding workload. Streamlined, intuitive interfaces are crucial to prevent distractions during patient interactions. Features should minimize clicks and mental load, allowing clinicians to maintain focus on patients rather than navigating the interface. Seamless integration reflects clinicians' preference for systems that enhance their current practices without requiring significant adjustments, reinforcing themes of efficiency and usability.

The desire for personalization was consistently highlighted, reflecting the clinicians' need for AI to adapt to their unique workflows and preferences. Customizability strengthens the perception that AI can be a flexible tool.

The focus on efficiency reflects clinicians' broader expectation that technology should simplify their work rather than complicate it, aligning with their professional priorities of time management and effectiveness.

> *"It's important for us to get the information we need while we're trying to move quickly." (P10)*

> *"If I introduce an AI scribe into practice, I'd expect something automated like this to work with it and not compete with it." (P11)*

Initial learning may be intensive but will smooth out with time

Clinicians recognize that introducing new technologies and adopting new systems requires an initial time investment. There is a willingness to engage with tools that promise long-term efficiency gains, provided the learning curve is manageable and the benefits are clear.

> *"Introducing something like this is probably going to be more work at the beginning, but I'd expect everything to become faster over time." (P5)*

Clinicians primarily saw the confidence score as a potential distraction but were curious about its utility, which may serve well in selective use.

The ambivalence toward confidence scores reflects broader concerns about balancing transparency with practicality. Clinicians prefer systems that provide trust-enhancing features without overloading or distracting them from clinical tasks. They prioritize

simplicity in straight-forward cases (e.g., uncomplicated UTIs) and would value a system that can adapt to varying levels of decision complexity.

> *"For easy diseases, confidence scores are useable." (P1)*

> *"Confidence score needs to be interpretable. This would be a net new concept to introduce to clinicians, we would need training to understand it. " (P13)*

> *"This might be useful for newer grads just starting a practice" (P2)*

> *"Maybe the confidence score would encourage me to probe further if it wasn't a high confidence score? I don't know. Maybe that's useful." (P5)*

The confidence score introduced clinicians to unfamiliar concepts and terminology, signaling a broader discomfort with opaque or overly technical elements of AI systems. With clinicians valuing simplicity and relevance and AI that can enhance their decision-making, the confidence score was seen as a distraction from core tasks.

Key Takeaway

For AI to be widely adopted, it must integrate seamlessly into clinical workflows, enhancing efficiency without adding cognitive or procedural burdens. Clinicians prioritize streamlined, intuitive interfaces that minimize clicks and distractions, allowing them to stay focused on patient care. While confidence scores were met with mixed reactions, clinicians expressed curiosity about their potential utility, particularly if presented in a simplified, interpretable format.

The expectation is that AI should be a flexible and adaptable tool, supporting decision-making without dictating it. By prioritizing clarity, usability, and personalization, AI can become a valuable asset in clinical settings.

Legal and Ethical Concerns

When discussing the integration of AI in EMR encounters, legal safeguards, patient consent, and clear role delineations are recurring concerns for clinicians. While they recognize AI as a helpful tool, they emphasize the need for clear policies to protect themselves and their patients. Clinicians expect AI-generated suggestions to be transparent, auditable, and easily distinguishable from their own input, reinforcing their role as the final decision-makers. Additionally, privacy, security, and consent concerns highlight the importance of minimizing administrative burdens while maintaining transparency. Establishing clear legal and policy frameworks ensures AI adoption enhances clinical practice without introducing new risks.

Clinicians require a visual representation that distinguishes what they say vs. what the AI generates before a note is finalized.

As clinicians engage with an AI-enabled EMR encounter, there was an expressed desire to actively distinguish between clinician and AI while charting.
The desire for clear distinctions between clinicians' input and AI suggestions reflects clinicians' concerns over accountability and transparency. It further reinforces the expectation that AI should augment, not overshadow, their professional judgment.

> *"I would want there to be a way to establish a distinction between the physician and the AI role in the EMR. Even if it's something as simple as giving the roles different font colors" (P10)*

Clinicians want the ability to review or audit AI suggestions. The theme of audibility emphasizes the importance of accountability and transparency in AI systems, especially ones that are integrated into the clinical context. Clinicians want to be able to verify and trust AI outputs without additional burden.

"I don't necessarily need to audit every patient I would use AI with. I would want the option there if I needed it, though. It's just good to have because you never know if one of your patients ends up with a legal issue." (P4)

Concerns on how legal frameworks are going to support the introduction of AI in clinical practice.

Clinicians know AI's legal implications and desire clear guidelines to mitigate risks. There is also a call for policies supporting AI use in healthcare while protecting clinicians and patients. With the number of governing bodies (e.g., Ontario Medical Association, College of Physicians & Surgeons Ontario, Canadian Medica Association, etc.), there is an expectation that these governance structures employ policies that align with AI use in primary care and how clinicians are ultimately seen as the overall decision-maker in a clinical encounter.

Clinicians are concerned about the legal implications and the adequacy of existing safeguards for AI use. No clear policies exist to protect clinicians in cases of adverse outcomes as a result of AI. There is a belief that policy will eventually support the use of AI in this way, but it currently lags.

Ultimately, they viewed AI as a tool that does not require direct disclosure to patients of its use.

All clinicians generally view AI as a clinical support tool rather than an independent decision-maker, meaning its use does not require direct disclosure to patients. Similar to decision-support algorithms, spell-checkers, or workflow automation tools already integrated into healthcare systems, AI is seen as an enhancement to existing processes rather than a separate entity that demands transparency to patients.

"Explaining AI to a patient would be time consuming" (P3)

" I wouldn't communicate it to the patient, it's no different than if I'm using the internet to look something up. I don't tell them that." (P1)

"I wouldn't really go out of my way. It doesn't really matter, I think people just want the service when they're here." (P7)

Privacy, Security, and Consent

Clinicians emphasize that patient consent for AI use should be clear, transparent, and minimally disruptive to their workflow. They support a one-time, easily understood consent process that does not require frequent reauthorization or interfere with daily clinical tasks. Ongoing pop-ups or repetitive consent prompts are seen as unnecessary administrative burdens that could detract from patient care rather than enhance it.

"Consent needs to be clear, but I want to agree to it once then never see it pop up in EMR again." (P4)

While clinicians recognize the importance of privacy and security in AI-driven healthcare, they advocate for a streamlined approach to consent that ensures compliance without increasing workload. They believe that AI should align with existing data privacy regulations and best practices without requiring additional oversight or manual intervention from clinicians. Ultimately, the expectation is that AI integrates seamlessly into clinical workflows, respecting both patient rights and provider efficiency while maintaining the highest standards of security.

Key Takeaway

Clinicians seek policies that define accountability, ensuring they are not held liable for AI-generated suggestions without proper safeguards. The ability to review and audit AI recommendations is a fundamental requirement, reinforcing transparency

and clinician oversight. While patient consent should be clear and straightforward, it should not add unnecessary administrative burdens. As AI adoption progresses, regulatory frameworks must evolve to provide clinicians with confidence in its legal and ethical use, ultimately ensuring AI serves as a collaborative tool that enhances, rather than complicates, patient care.

Human-Centered Design

Clinicians emphasize that AI should enhance, not replace, the human connection in patient care. While AI can streamline workflows and assist with decision-making, empathy and communication remain central to clinical interactions. Clinicians want AI to function as a support tool that improves efficiency without compromising the patient experience or reducing their cognitive engagement. The challenge lies in balancing AI's analytical strengths with human-centered care, ensuring that technology supports rather than interferes with the clinician-patient relationship.

AI must maintain a human element in patient interactions by not distracting clinicians and performing reliably.

Clinicians view AI as a tool to enhance, not replace, the human connection in patient care. Empathy and communication are paramount in ensuring AI's alignment with clinical values. Maintaining a human element in care means AI must function seamlessly in the background, providing reliable assistance without creating distractions or interfering with natural patient-clinician interactions. If AI is too intrusive or mechanized, it risks eroding the personal connection and emotional intelligence that are essential for patient confidence and satisfaction.

> *"The human element matters. AI could deal with a straightforward UTI, write a script, and say see you later 5x a day. But would people feel cared for? Probably not". (P5)*

AI should support human-centered care without creating dependence on AI.

AI should enhance the clinician-patient relationship rather than replace or overshadow it. Maintaining human-centric care ensures AI remains a tool for empowerment, not a crutch that diminishes clinical judgment.

Some participants expressed concern that AI could reduce clinicians' cognitive engagement over time if they become overly dependent on it, which could interrupt their critical thinking.

> *" I'm afraid of AI leading to less thinking for clinicians" (P2)*

Key Takeaway

While AI can assist with routine tasks and provide valuable insights, over-reliance on AI is seen as a risk for diminishing clinical judgment and critical thinking over time. Maintaining a balance between automation and clinician engagement ensures AI remains a tool for empowerment rather than dependency. Ultimately, AI should be designed to strengthen, not weaken, the patient-clinician connection, reinforcing trust, empathy, and high-quality care.

Discussion

While this study provides valuable insights into clinicians' perspectives on AI-enabled EMRs, several limitations must be acknowledged. One key constraint is the immaturity of the AI-enabled EMR market in primary care, requiring research to be conducted through scenario-based interviews rather than real-world implementation.

The lack of established AI integration in clinical workflows meant that mock-up interfaces and guided discussions were necessary to simulate use cases, potentially limiting the depth of participant interaction and feedback.

Another limitation stems from the study's scenario-based nature, which restricts the ability to observe clinicians' natural behaviors when interacting with AI-enabled tools. In a real-world setting, contextual factors, such as time constraints, patient load, and cognitive workload, may shape how clinicians engage with AI. While the structured interview format allowed for targeted exploration of key themes, further research with longitudinal field studies in live EMR environments would provide a more comprehensive understanding of AI's impact on the factors explored in this study.

Additionally, the participant pool may not fully represent the broader clinician population. Those who volunteered for the study may have a higher level of technical literacy or openness to AI, potentially skewing findings toward greater acceptance or confidence in AI integration. A more diverse sample, including clinicians with varying levels of experience, digital comfort, and specialty focus, would help ensure a more representative analysis of how AI is perceived and utilized across different practice settings.

Despite these limitations, the study contributes valuable insights into the conditions necessary for AI adoption in primary care, particularly regarding trust, clinician autonomy, workflow integration, and legal/ethical considerations. As AI-enabled tools continue to evolve, ongoing usability testing, real-world pilot implementations, and clinician feedback loops will be essential in refining AI's role within EMR systems to align with real clinical needs and constraints.

The findings of this study align with recent research on AI integration in clinical settings, reinforcing the broader challenges and considerations identified in the field. One of the key themes that emerged from this study was clinician trust and autonomy, with participants emphasizing that AI should act as a supportive tool rather than replace clinical decision-making. This concern is echoed in Coulibaly et al. (2025), which identified clinician skepticism, usability challenges, and a lack of transparency as major barriers to AI adoption in Canadian healthcare. Similarly, Al-Ansari et al. (2024) found that AI adoption is often hindered by poor integration into existing workflows, resulting in a lack of user confidence and reluctance to engage with AI-assisted decision-making.

Another major theme from this study was the importance of seamless workflow integration, as participants stressed that AI tools must complement, rather than disrupt, existing clinical processes. This is consistent with the AI for IMPACTS framework (2025), which highlights workflow integration, interoperability, and clinician burden as critical factors in AI evaluation. Similarly, Al-Ansari et al. (2024) noted that many existing AI evaluation methodologies focus heavily on technical performance without adequately considering how AI fits into real-world clinical settings, further supporting the need for usability-driven AI design.

In addition to usability concerns, this study underscored the need for AI evaluation approaches that extend beyond accuracy metrics to assess practical usability and clinician experience. While many AI studies focus primarily on model performance, Al-Ansari et al. (2024) emphasized the necessity of user-centered evaluation methodologies to ensure AI tools are genuinely beneficial in clinical practice. This aligns with the AI for IMPACTS framework, which advocates for evaluating AI's effect on clinical decision-making and patient outcomes rather than viewing AI assessment purely as a technical exercise.

A final key finding in this study was the importance of context-aware AI design, ensuring that AI-generated recommendations are specific to the clinical scenario rather than providing generic, one-size-fits-all outputs. Participants expressed frustration with AI models that fail to consider the nuances of individual patient cases, reinforcing concerns raised by Coulibaly et al. (2025) regarding AI's difficulty in adapting to real-world clinical variability. The AI for IMPACTS framework also emphasizes the need for context-aware AI, ensuring that AI-generated insights are interpretable, relevant, and actionable for clinicians.

While these studies reinforce the themes identified in this research, this study adds a unique contribution by providing granular, clinician-driven insights into how these challenges manifest in practice. Unlike broad AI evaluation frameworks that focus on abstract principles, this research offers specific design implications that directly address clinician concerns, such as ensuring AI-generated recommendations maintain clinical autonomy, embedding AI seamlessly into existing workflows, and enhancing AI usability through context-aware design. By bridging qualitative findings with practical AI design considerations, this study contributes to the growing body of research aimed at making AI tools more effective, clinician-friendly, and aligned with real-world primary care needs.

Conclusion

Integrating AI into primary care EMR encounters presents opportunities and challenges, requiring careful consideration of trust, transparency, workflow efficiency, and clinician autonomy. Clinicians generally recognize AI's potential to enhance decision-making, improve efficiency, and support patient-centered care, but they remain cautious about overreliance, legal implications, and workflow disruptions. Trust in AI is not automatic; it must be earned through high accuracy and clear transparency.

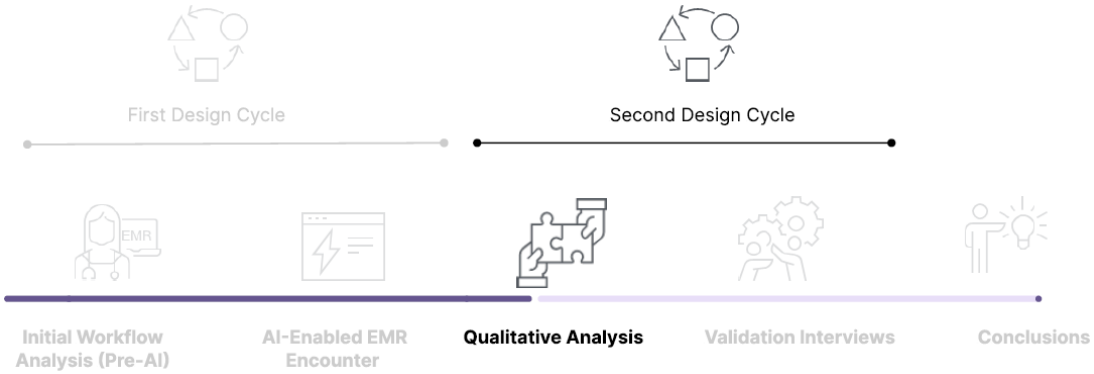
The findings underscore the importance of context-sensitive AI implementation that adapts to case complexity, aligns with clinical decision-making, and preserves human-centered care. While AI can be particularly effective in automating straightforward cases, more nuanced or high-stakes scenarios require greater clinician discretion and control. Additionally, seamless workflow integration is crucial, AI tools should reduce cognitive burden rather than add friction to clinical encounters.

Looking ahead, policy development and regulatory frameworks must keep pace with AI adoption to safeguard both clinicians and patients. Addressing concerns around legal accountability, audibility, and patient consent will be key to fostering confidence and widespread adoption. Future research should focus on real-world testing in live clinical settings.

Ultimately, AI in primary care must be designed to empower rather than replace clinicians, enhance rather than disrupt care delivery, and complement rather than compete with human expertise. By prioritizing trust, usability, and ethical safeguards, AI can become a valuable and sustainable tool in healthcare, supporting clinicians without compromising the essential human element of patient care.

These findings will directly inform the next phase of this research, which involves redesigning the AI-enabled interface with these principles in mind. The next chapter will outline the revised interface, detailing how these insights were translated into concrete design changes.

Chapter 5
Redesign Process



Given the feedback from primary care clinicians, their perspectives were carefully integrated into the redesign process to ensure that the updated interfaces aligned with their needs and expectations. Insights from the thematic analysis conducted in this study directly informed the mock-up redesign. By systematically analyzing user feedback, the study identified key factors influencing perceived ease of use and usefulness, both of which play a critical role in shaping clinician adoption of AI-enabled tools. This chapter outlines the iterative design process.

Perceived ease of use is a significant determinant of user acceptance, not only for its direct impact but also through its influence on perceived usefulness (Davis & Granić, 2024). A system that is difficult to navigate can create friction, leading to frustration and reduced engagement, even if it offers meaningful benefits. However, perceived usefulness remains the primary driver of adoption, users may tolerate some level of difficulty if a system provides essential functionality that substantially enhances their work (Davis & Granić, 2024). In the context of this study, scenario-based interviews incorporating mock-ups provided a valuable mechanism for evaluating perceived usefulness, even in the absence of hands-on interaction. Research has shown that TAM-based assessments of usefulness remain stable and predictive of user

acceptance even when based on non-interactive system demonstrations (Davis & Granić, 2024).

Table 6 presents a high-level framework for translating qualitative findings into design implications by categorizing clinician perspectives and assessing their impact on ease of use and usefulness. Appendix E expands on these insights by detailing specific redesign requirements derived from the thematic analysis, ensuring alignment between qualitative findings and practical system modifications. While user-centered design methodologies typically emphasize usability heuristics and iterative testing, fewer artifacts explicitly integrate qualitative thematic analysis with ease of use and usefulness considerations. This table addresses that gap by providing a structured yet flexible approach to aligning clinician perspectives with system redesign requirements.

Affinity diagrams are valuable for clustering qualitative data and identifying patterns, but they primarily serve as sense-making tools rather than directly informing design decisions (Abascal et al., 2015). While they surface broad themes, they do not inherently guide how those themes could be operationalized into system features or interface elements, making them insufficient for the purposes of this task within this research. In contrast, the qualitative findings-to-design implications table ensures that research insights lead to concrete design interventions, making the AI-enabled EMR encounter development process directly responsive to clinician needs. By incorporating additional criteria, such as ease of use and usefulness, the table refines design considerations in a structured manner. Rather than stopping at thematic organization, it prescribes actionable recommendations, such as integrating tooltips for confidence scores, maintaining clinician autonomy in AI-assisted decision-making, and ensuring AI suggestions are context-aware. This structured research-to-

design mapping helps bridge the gap between qualitative findings and practical AI system improvements.

While alternative approaches, such as a work consolidation model, could have been considered, they were not ideal for the specific problem this study sought to address. A work consolidation model synthesizes multiple contextual design frameworks, such as Flow Models, Sequence Models, and Cultural Models, to create a holistic view of work practices across an organization's design (Holtzblatt & Beyer, 2017). However, this study focused more narrowly on clinician-AI interactions within an EMR-enabled encounter in a clinical visit rather than attempting to consolidate all aspects of clinical work. As a result, a work consolidation model would have been too broad and less actionable for guiding AI design in this specific context.

Similarly, while highly effective for refining interface usability, a strictly UI-focused approach, linking individual user observations to specific interface components such as button placement or workflow responsiveness, would not align well with this study’s broader goal of understanding how AI can integrate into clinical workflows without disrupting existing practices. Instead, the thematic mapping table provides a structured, high-level research-to-design translation, capturing key clinician concerns like trust, decision-making authority, and workflow efficiency while ensuring that the AI system aligns with real-world clinical needs.

Using this approach, the study ensures that clinician insights are categorized and analyzed and transformed into meaningful design implications, bridging the gap between qualitative research and AI system development in primary care settings.

Table 6. Qualitative findings to design implications.

Main Theme	Sub Theme	Clinician Perspective	Functional Requirement	Impact on Ease of Use	Impact on Usefulness
Trust & Transparency	Clinicians' trust is linked to high-performance expectations of AI.	Clinicians expect AI to consistently deliver accurate, reliable insights to earn trust.	Ensure AI delivers precise, high-performance suggestions to maintain clinician trust.	This is a back-end AI specific consideration. (Establish consistent, high-trust results of AI-driven summaries before phasing the approach for rolling it out. Deliver small pieces of high value at a time).	
	AI can build trust with clinicians by reinforcing their authority by providing relevant, nonintrusive support.	Clinicians have low tolerance for AI errors and demand explainability.	Provide clear, auditable decision-making processes with optional visibility.	Make audit trails available and easy to generate.	Include robust audit trails that distinguish clinician and AI actions, with clear roles and permissions.

	Transparency in how confidence scores are determined and training data could improve trust.	Clinicians want the optionality to understand confidence scores better	Provide tooltips, education, or modular support for confidence scores	Do not use large amounts of real estate on the EMR screen to explain confidence scores at the time of charting.	Include onboarding materials or guidance for this learning phase, with reassurance of efficiency benefits over time.
	Clinicians felt uncertain about generic button functions, which led to hesitation.	Unclear buttons create hesitation and usability concerns.	Use explicit button labels and tooltips to improve clarity.	Simplify and clearly label buttons with descriptions of subsequent actions.	Consider removing buttons to reduce the required clicks and avoid added mental load.
Clinician Authority & Decision Making	AI can be helpful for supportive suggestions like reminders, preventative care guidelines,	Clinicians see value in support around a patient (e.g. they are offering core service with AI reminding them other valuable services to include)	Use reminders in a suggestive way, make sure language does not sound explicit or punitive.	Make reminders easy to dismiss, isolate them as an unobstructive toast notification of the screen and not in the middle of the interface.	Only flag relevant reminders necessary for that visit. (e.g. do not prompt growth charts for patients past the age of 2)

	or flagging allergies.				
	Suggestions should align with realistic, helpful actions that are evidence-based and practical.	Practical , action-oriented items are prioritized	Ensure AI offers recommendations but never enforces actions.	Make recommendations easy to dismiss, isolate them as an unobstructive toast notification of the screen and not in the middle of the interface.	Align with best practice guidelines that are constantly changing to help clinicians stay up-to-date.
	AI should offer useful suggestions without appearing to make autonomous decisions.	Clinicians may require a visual separation to prompt and use AI on an encounter by choice	Provide a button to prompt the use or trigger AI, do not by default have it look like it's running without choice	Introduce a user-friendly button that prompts the initiation of using AI (e.g. “AI assist”). To give clinicians control over AI-generated recommendations.	Ensure suggested outcomes are framed as optional, contextual guidance rather than definitive actions.

	Clinicians should retain final decision-making authority, AI's role should be to support, not dictate, or override clinical decisions.	AI must reinforce, not undermine, clinician autonomy.	Integrate AI-driven alerts without enforcing decisions.	Implement a visual or structural delineation within the interface to make it easy to distinguish between clinician and AI inputs.	Emphasize AI as a support tool, not a decision-maker. This could involve passive reminders rather than prescriptive prompts
Complexity Sensitivity	Context-sensitive suggestions: AI-suggested medications are acceptable for simple cases, while complex cases require	AI should adjust involvement based on case complexity.	Implement thresholds where AI suggests vs. provides context-aware recommendations.	Provide clear warning labels if a more nuanced situation requires further clinician probing (e.g. it can also be indicated using a low interpreted confidence score)	Tailor AI involvement based on case complexity, potentially adjusting the extent of AI suggestions based on the condition type

	heavier clinician discretion.				
	Contextual information provided by clinicians in prior interactions is important. Clinicians want AI that understands past interactions and can capture nuanced contexts.	Clinicians want AI to remember past patient interactions for contextualized support.	Enable AI to retain patient history and clinician preferences. "In-chart context" helps.	Keep this within the encounter screen without having to navigate other areas of the chart.	Incorporate context memory and ensure transparency about AI's training, potentially via tooltips or "learn more" links.
	Summaries need more comprehensive	Historical information is critical to	Explore how clinicians can prompt	Keep this within the encounter screen without	Surface context-relevant scoring (e.g. If patient is here

	patient histories, with previous scores and relevant screening tools displayed.	understanding a patient	extensive patient histories	having to navigate other areas of the chart.	for MH then show trend of previous scores)
	Mixed preferences on medication suggestions, depending on the complexity of the case.	Clinicians expect AI to refine recommendations based on prescribing habits.	Develop AI that learns from clinician choices to refine medication suggestions.	Explicitly clarify AI medication suggestions, avoiding brand-specific influences unless justified to maintain clinical trust.	Emphasize medication support features that align with clinicians' existing workflows and safety standards.
Workflow Efficiency and Integration	Clinicians desire actionable, patient-centered summaries with appropriate depth and focus.	Clinicians value have patients at the center of the encounter	Figure out what the right amount of information is	Consider collapsing key points for a streamlined overview, use bullets, be concise.	Provide options for clinicians to customize the summary's content level.

	Clinicians expect prompts for preventative measures (e.g. mental health risk factors).	Close alignment with existing clinical pathways is necessary	Ensure AI suggestions are aligned to current best practice	Provide these suggestions in unobtrusive way, label as suggestions, potentially a checkbox.	Explore the feasibility of a floating, interactive AI panel for easy access.
	Clinicians are cautious of features that add workload or disrupt their workflow.	Clinicians reject AI features that add workload or disrupt workflow.	Ensure AI integrates smoothly with minimal extra steps.	AI interactions need to feel seamless and not add extra steps. Minimising any additional tasks tied to AI can improve acceptance.	Design the workflow to offer optionality, ensuring clinicians can engage with AI only as needed, reinforcing a seamless experience.
	AI needs to fit smoothly into the existing workflow without adding steps.	In-context of the encounter does not require clinicians to open new windows or tabs	Ensure AI integrates smoothly with minimal extra steps.	Keep AI related content within the encounter.	

	Clinicians primarily saw the confidence score as a potential distraction but were curious about its utility, which may serve well in selective use.	Clinicians find confidence scores distracting and unclear.	Allow users to toggle confidence scores and provide clear explanations.	Offer optional visibility for confidence scores, potentially with simplified terminology. Consider providing the ability to show how scores were generated.	Introduce selective visibility for confidence scores, allowing clinicians to customize when and where they see them.
Legal and Ethical Considerations	Clinicians require a visual representation that distinguishes what they say vs. what the AI generates before	Clinicians need clear visual separation of AI vs. clinician-generated content.	Use different colors or labels to differentiate AI input.	Consider tools like colour coding and flexible fields for customization. Ensure summaries meet professional standards (e.g., OMA). Place outcomes at the bottom to support natural workflow.	Colours should be distinct at the time of charting, and for the purposes of audit. But when clinicians "sign off" on an encounter it can all be seen as one finalized note in one colour.

	a note is finalized.				
	Concerns on how legal frameworks are going to support the introduction of AI in clinical practice.	Clinicians want policies that support AI use without adding administrative burden.	Develop governance frameworks ensuring AI transparency without over-regulation.	Consider incorporating legal disclaimers and protections that clarify clinician roles and safeguard patient outcomes.	
	Ultimately viewed AI as a tool that does not require direct disclosure to patients of its use.	Clinicians might be concerned about causing unnecessary concern or involvement from patients on their digital health choices	Consent management frameworks need to enable the use of AI outside of requiring direct patient consent for its use	Create low-touch, user-friendly consent processes, ensuring patients are aware without unnecessary complexity.	If needed, provide an optional disclosure message. Clinicians generally don't see it as a high priority.

Human-Centered Design	AI must maintain a human element in patient interactions by not distracting clinicians and performing reliably.	AI should enhance care without depersonalizing patient interactions.	Ensure AI operates in the background, complementing human interactions.	Contain how AI works within the chart, do not enable large pop-ups, warning, or indications of it acting autonomously without the clinician's involvement.	
	AI should support human-centered care without creating dependence on AI.	Clinicians fear over-reliance on AI could erode clinical judgment.	Position the AI enabled encounter as an assistive tool rather than an authoritative system.	Balance automation with interactive elements that encourage clinicians to remain actively engaged in decision-making.	This is a training consideration that can be impacted depending on the clinician's practice experience (e.g. more seasoned clinicians might not have the risk of over reliance as much as new clinician graduates might).

Primary Care Clinician User Persona

The thematic analysis and functional requirements gathering findings were synthesized to develop a user persona, a character-driven representation of the target user. Personas help maintain alignment by ensuring the redesign focuses on real user needs and expectations (Lowdermilk, 2013). Given the limited sample size, only one persona was created; however, the needs and challenges identified through thematic analysis were consistent enough across participants to justify a generalized representation for the purpose of redesign. This fictional persona embodies the characteristics and challenges of the clinicians engaged in this study, serving as a reference point to guide redesign decisions.

Figure 6 Primary Care Physician Persona

Primary Care Physician: Dr. William | Experience: 12 years in primary care **|** Ontario, Canada
Work Setting: Works in a busy urban family practice clinic, managing 30-40 patient encounters per day with a roster size of approximately 1800 patients.
Tech Proficiency: Moderate comfort with EMRs but skepticism about AI-driven tools embedded in their encounter.
Key Responsibilities:

- Diagnose and manage diverse patient conditions
- Review and document patient history efficiently
- Balancing administrative tasks with clinical decision-making
- Provide evidence-based, patient-centered care

Pain Points & Challenges

❌ **High cognitive load:** Balancing patient care, documentation, and decision-making is overwhelming
❌ **Time pressure:** Limited time per patient (average 10-15 minutes per consult)
❌ **EMR Frustrations:** Prefers simple, efficient interfaces but finds existing EMRs cumbersome
❌ **AI Skepticism:** Concern about over-reliance on AI, transparency in recommendations, and liability issues

Design Needs & Preferences

✓ **Trust & Transparency:** AI recommendations must be explainable, auditable, and not prescriptive
✓ **Clinician Authority:** Prefers AI as a supportive tool, not a decision-maker
✓ **Minimal Disruption:** AI should integrate seamlessly into workflows without adding extra clicks
✓ **Complexity Sensitivity:** Open to AI-assisted suggestions for routine cases (e.g., UTIs) but prefers autonomy in complex cases (e.g., mental health)
✓ **Customization & Personalization:** Wants AI tools to adapt to individual practice styles rather than enforcing rigid workflows

Mock-Up Redesigns

The insights and design implications outlined above provide a strong foundation for refining the system's interface and functionality. By addressing key concerns such as trust, usability, workflow integration, and clinician autonomy, the redesign aims to create a more intuitive and effective user experience.

The following section translates these considerations into concrete design decisions that prioritize clinician control, transparency, and efficiency. The updated mock-ups incorporate solutions that enhance ease of use, improve interpretability, and seamlessly integrate AI support without disrupting established clinical workflows. Specific adjustments, such as explicit AI activation triggers, improved labeling, and context-aware recommendations, are designed to reduce cognitive load while maintaining clinician oversight.

Each redesign choice is guided by the core themes identified earlier, ensuring alignment with real-world needs and expectations. The goal is to create a system that meets usability and ethical standards and actively supports clinicians in providing high-quality patient care. Through this approach, the interface becomes more than just a technological enhancement it evolves into a tool that fosters clinician confidence, impacting the likelihood of adoption.

The following mock-up redesigns represent the second iteration of the design process. The first iteration involved developing the AI-enabled EMR encounter based on the original CHR encounter screen, as shown in **Figure 3**, with key differences outlined in **Table 1**. This redesign incorporates insights from the thematic analysis, further refining and reimagining the AI-enabled EMR encounter based on clinician feedback and user-centered considerations.

1.0 Redesigned AI-Enabled Encounter Screen for Mental Health Scenario

1.1 Encounter Screen Resting State

The thematic analysis and redesign requirements related to clinician authority and decision-making emphasized the need for AI to function as a supportive, suggestive tool rather than an automatic feature within clinical encounters. Similarly, insights from complexity sensitivity reinforced the importance of giving clinicians control over when and how AI is used.

To address this, **Figure 6** shows a friendly and approachable "AI Assist" button was introduced as a clear trigger, ensuring AI activation remains clinician-driven. The deliberate use of the term "Assist" was chosen to reinforce AI's role as a supportive aid rather than a directive force, ensuring clinicians remain at the center of decision-making. The PHQ-9 and GAD-7 questionnaires are still displayed at the top, indicating completion. Unlike the original mock-ups, this redesigned screen provides an explicit, actionable trigger, reinforcing that AI operates under clinician control rather than functioning autonomously in the background.

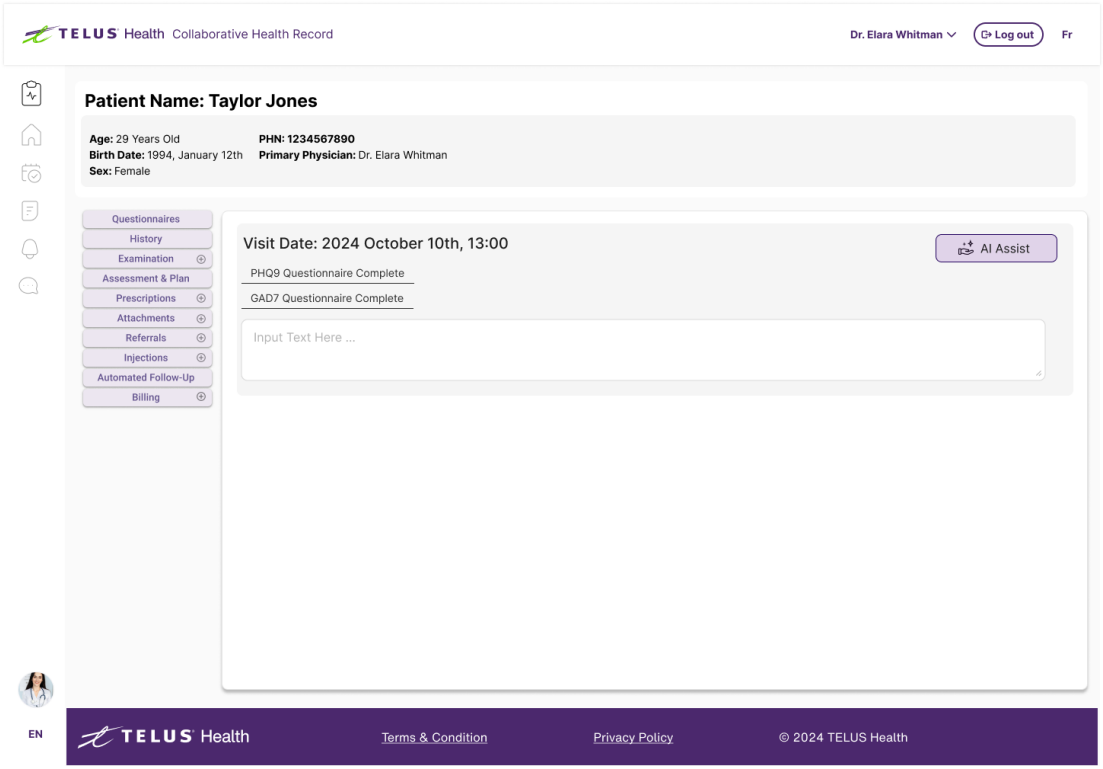


Figure 7.Redesigned encounter screen resting state mental health.

1.2 AI Populated Encounter Screen

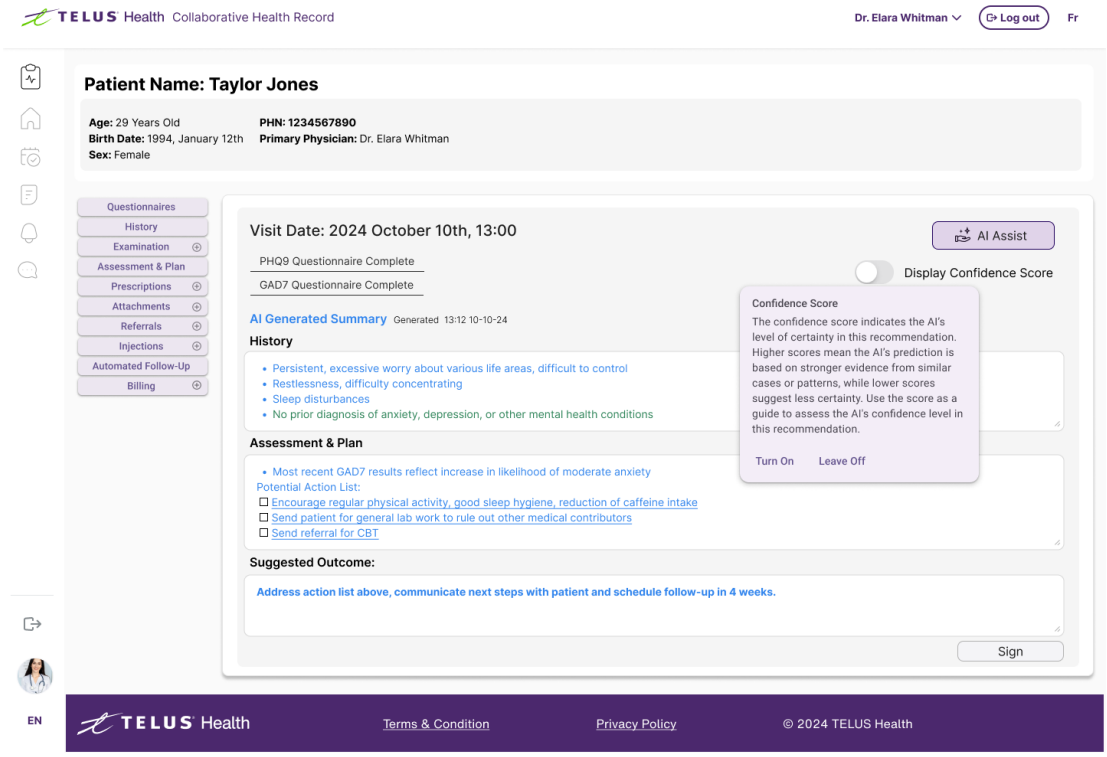


Figure 8. AI populated encounter screen mental health.

Figure 7 demonstrates what happens after clicking the "AI Assist" button. The button populates the encounter for the clinician with the title "AI-Generated Summary" and a date/time stamp indicating when it was generated. While subtle, these design choices play a crucial role in reinforcing auditability, ensuring clinicians can track when AI-generated content was introduced. This directly supports design requirements for trust and transparency by providing a clear, traceable record of AI involvement.

To address legal and ethical considerations, a color-coding system was implemented to visually distinguish content sources:

- AI-generated content appears in blue, ensuring it is easily identifiable.

- Clinician-entered text appears in green, reinforcing human oversight and control.

A confidence score toggle was introduced to accommodate clinicians' varying preferences, aligning with workflow efficiency and design integration. Some clinicians expressed curiosity and willingness to engage with confidence scores, while others found them distracting. The toggle allows users to enable or disable this feature as needed. Additionally, this design decision aligns with complexity sensitivity for nuanced cases, clinicians may find the confidence score valuable as an additional reference point, prompting further investigation. Conversely, the score may provide little added value in straightforward cases (e.g., an uncomplicated UTI). A tooltip explaining the confidence score's meaning and interpretation was included to enhance usability, ensuring clinicians have the necessary context.

Further refinements strengthened trust and transparency while maintaining clinician authority and decision-making. The language used throughout the interface was intentionally adjusted to be suggestive rather than prescriptive, reinforcing AI's role as a supportive tool rather than a directive force.

To emphasize that AI augments rather than dictates clinician decision-making, the assessment and plan section now includes a "Potential Action List" instead of rigid recommendations. Additionally, suggested outcomes were deliberately phrased more ambiguously and flexibly, further reinforcing that clinicians retain full control over the clinical encounter. The positioning of AI-generated suggestions was also adjusted, shifting them to the bottom of the encounter to align with typical clinical workflows. This mirrors the SOAP format, where clinicians review patient data first and make treatment decisions toward the end of the encounter.

1.3 AI populated encounter screen with confidence score on

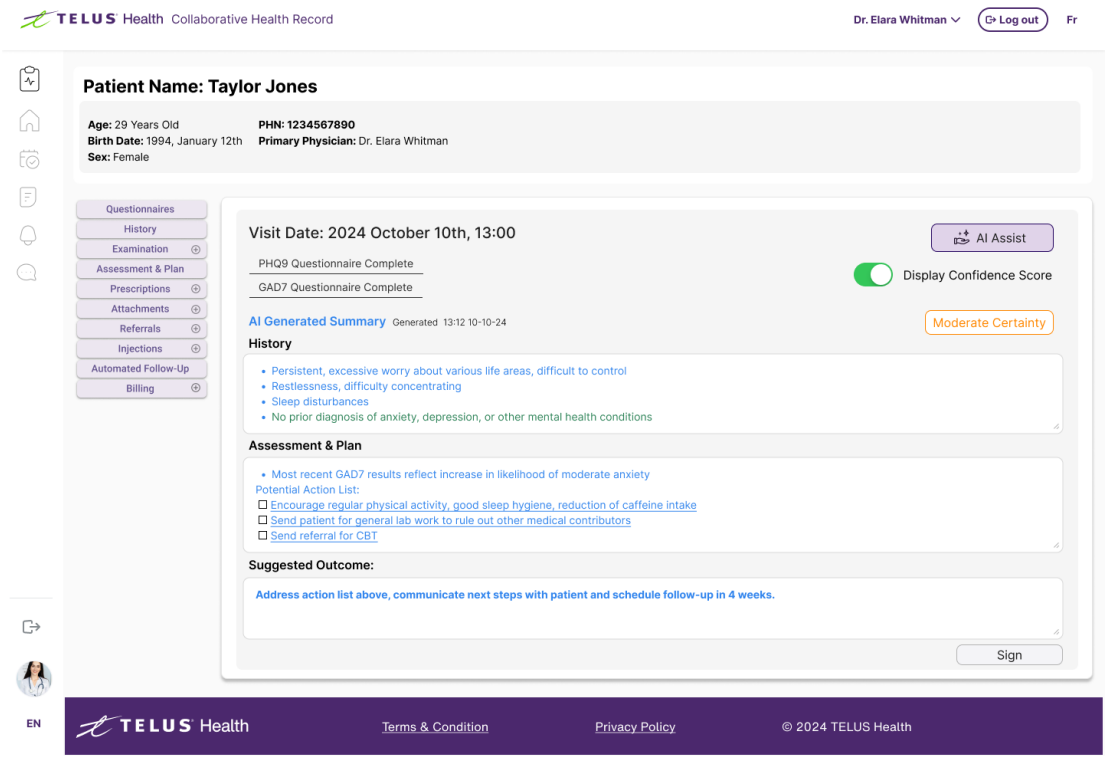


Figure 9. AI populated encounter screen with confidence score on

When the “Display Confidence Score” toggle is selected and in the “on” state, **Figure 8** demonstrates that the confidence score is now displayed using qualitative labels “Low Certainty,” “Moderate Certainty,” or “High Certainty” with corresponding color coding. This approach was designed to minimize the cognitive load by avoiding numerical representations, which clinicians found distracting and effortful to interpret. Instead, the simplified labels provide a quick, intuitive assessment of certainty without requiring additional mental processing.

This design decision directly supports workflow efficiency and integration by ensuring that clinicians can engage with confidence scores in a way that complements their decision-making process rather than disrupting it. By making the confidence

score an optional, easily digestible visual cue, it respects clinicians' varied preferences. Those interested can enable it for additional context, while those who find it unnecessary can turn it off.

Additionally, while there was general curiosity about confidence scores and their potential impact on clinical decision-making, there was no immediate resistance or demand for their complete removal. Instead, the ability to toggle the feature on or off provides flexibility, ensuring that AI support remains a seamless, non-intrusive addition to the clinical workflow rather than an imposed distraction.

1.4 AI populated encounter screen with confidence score off

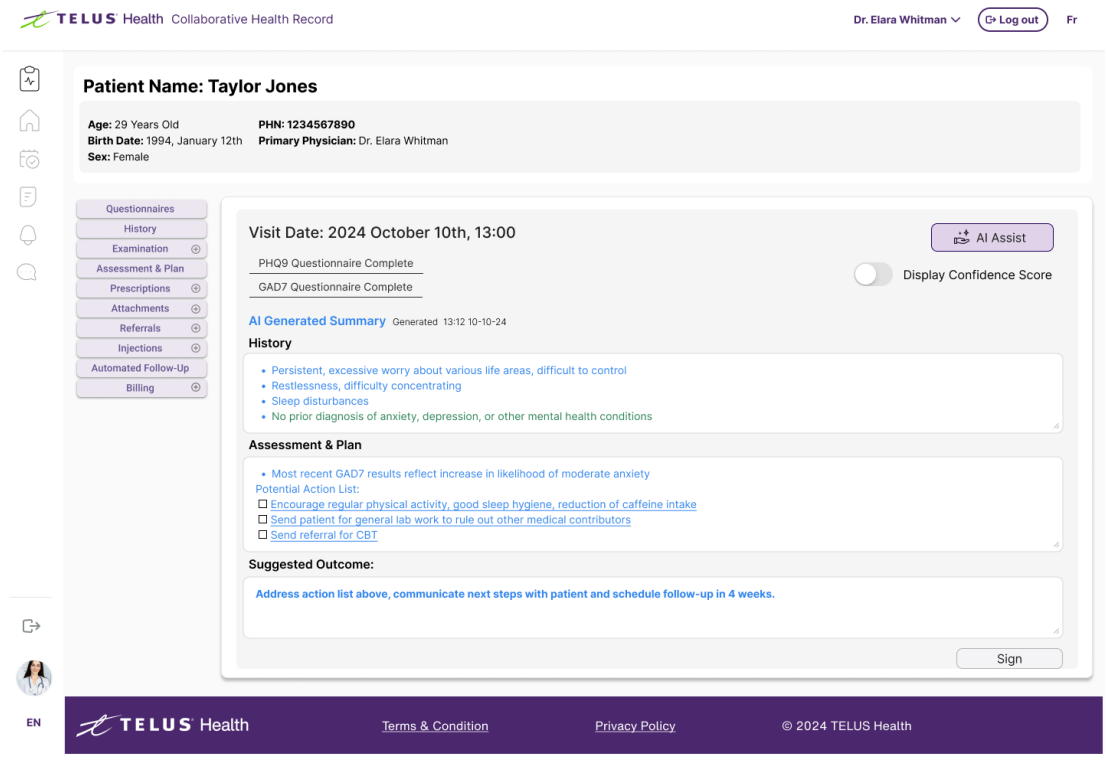


Figure 10. AI populated encounter screen with confidence score off

Clinicians were given the option to enable or disable the confidence score feature. This approach addressed the feedback that, while clinicians were curious about the

potential long-term utility of the confidence score, it was not considered a primary factor in adopting AI tools.

2.0 Redesigned AI-Enabled Encounter Screen for UTI Scenario

3.1 Encounter Screen Resting State

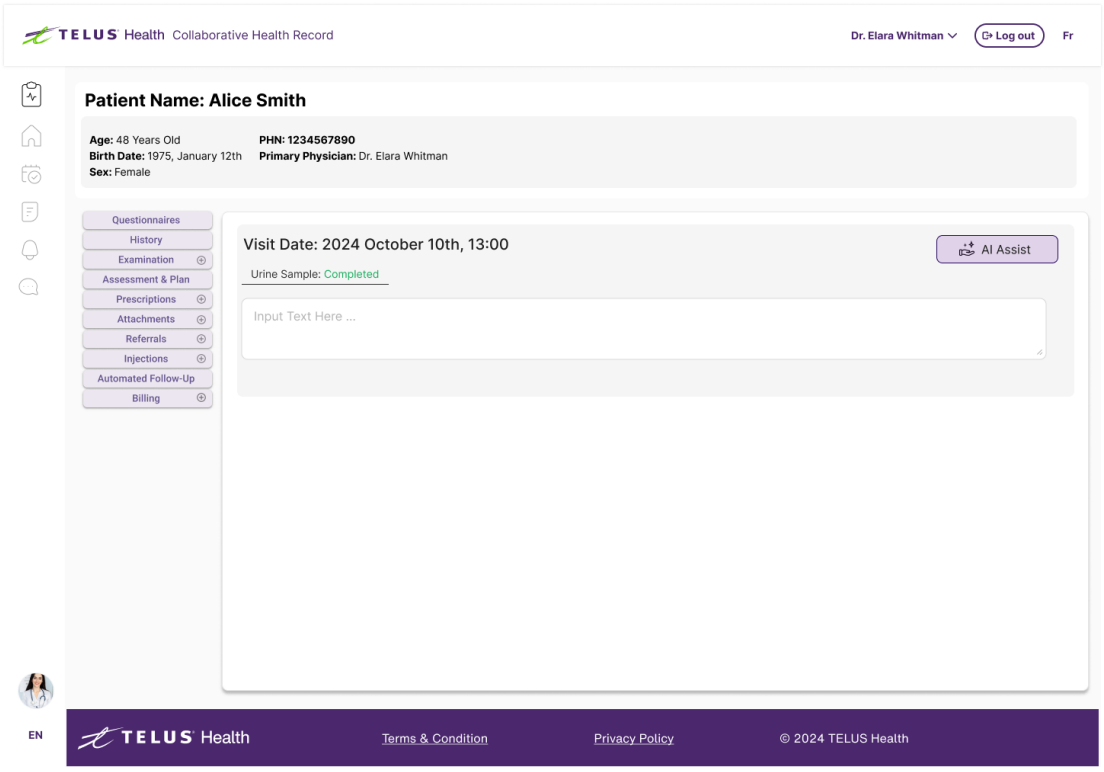


Figure 11. Redesigned encounter screen resting state – UTI scenario

Similar to the mental health scenario, **Figure 10** shows the redesign encounter screen for the UTI scenario with the AI Assist button. This button was implemented to ensure clinicians have the option to engage AI support as needed rather than having it automatically applied. This aligns with workflow efficiency and integration by allowing AI to function as a flexible tool that adapts to clinician preferences rather than an imposed system feature.

Additionally, interview findings highlighted the importance of clearly indicating key steps that had already been completed by support staff before the clinician saw the patient. To address this, the screen was designed to display explicitly the completion of pre-visit tasks, such as collecting a urine sample. This ensures clinicians do not waste time duplicating efforts and can efficiently focus on decision-making, reinforcing AI's role in streamlining workflows rather than introducing additional complexity.

3.2 AI Populated Encounter Screen – UTI Scenario

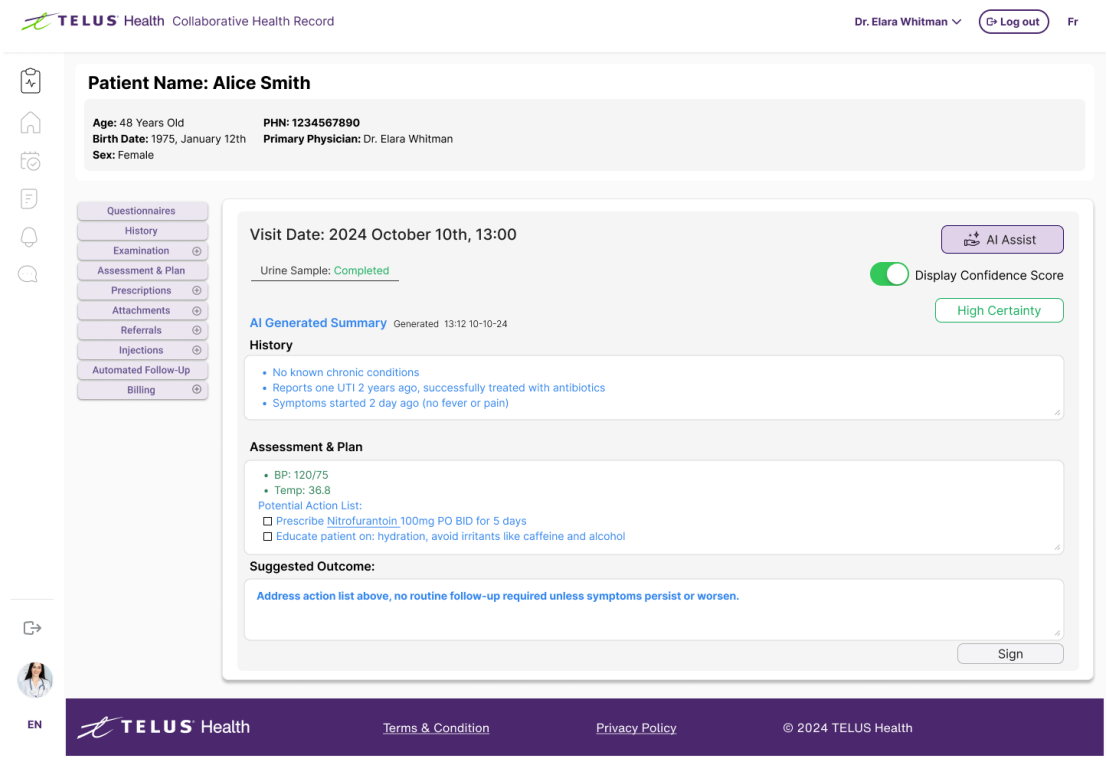


Figure 12. AI populated encounter screen for UTI scenario.

The same design considerations from the mental health scenario were applied to the UTI scenario, with one key distinction. Thematic analysis revealed clinicians viewed the UTI scenario as more suitable for AI-driven prescription suggestions, as antibiotic choices are typically standardized. In contrast, mental health treatments were seen as

too nuanced for AI to reliably recommend a medication regimen, making prescription suggestions less appropriate in that context.

To reflect this, the mock-up presents an AI-generated suggestion for prescribing nitrofurantoin, a commonly used first-line antibiotic for UTIs, including dosage and duration. This ensures that AI supports clinical efficiency by streamlining routine decision-making, particularly in cases where treatment pathways are well-established. The selection of medications in this scenario directly addresses design requirements within the complexity sensitivity theme, as clinicians expressed a need for AI recommendations to be context-aware, providing structured guidance for straightforward cases while avoiding overreach in nuanced scenarios.

Additionally, this screen maintains clinician control and flexibility by allowing them to toggle the confidence score on or off, with this scenario displaying a “high confidence” rating. Including clear, algorithmically approved treatment suggestions ensures that AI recommendations align with best practices and established clinical workflows, addressing key workflow efficiency and integration themes. Rather than introducing unnecessary complexity, the AI-generated prescription suggestions reinforce decisions clinicians would have made independently, acting as a supportive tool.

Discussion

The various design decisions made throughout this study collectively address the human-centered design theme. A key consideration in AI-enabled EMR encounters is ensuring that technology enhances, rather than disrupts, the patient-clinician interaction. Digital health tools must be designed to support clinicians without

introducing unnecessary distractions or cognitive burdens. By prioritizing clinician control, transparency, and seamless integration, the redesign maintains a balance between leveraging AI for efficiency and preserving the clinician's ability to engage meaningfully with patients.

The thematic analysis conducted in this study revealed key redesign considerations that directly informed the development of an AI-assisted clinical interface. By addressing clinician authority, workflow efficiency, trust, transparency, and complexity sensitivity, the iterative design process aimed to create an interface that is functional and aligned with real-world clinician needs and expectations.

Although these refinements significantly improve the interface's usability and integration into clinical workflows, it is important to acknowledge that no interface can be considered "perfect" upon initial implementation. Instead, the goal was to design an AI-enabled EMR encounter that aligns with clinicians' mental models, decision-making processes, and workflow preferences in a way that fosters gradual adoption over time. AI in clinical settings often faces resistance due to concerns about automation bias, lack of transparency, and disruptions to existing routines. By embedding suggestive rather than prescriptive AI guidance, maintaining clinician autonomy, and ensuring seamless workflow integration, the redesign seeks to minimize friction and increase trust.

A key takeaway from this study is that clinician adoption of AI tools depends not solely on technical accuracy but also on how well the system integrates into their existing workflows. The intentional design decisions, such as allowing clinicians to toggle AI-generated confidence scores, providing clear distinctions between AI and human input, and ensuring that medication suggestions align with established clinical

decision pathways, reflect a broader human-centered approach to AI adoption. The interface encourages progressive engagement rather than immediate reliance by offering control, transparency, and optionality.

Future iterations of this design will require continued user feedback and usability testing to refine the interface further and identify any unintended challenges or resistance points that emerge in real-world clinical settings. As AI continues to evolve in healthcare, system design must remain adaptive, responsive to clinician concerns, and grounded in user-centered design principles and ethical AI integration.

Conclusion

This study explored key design considerations for embedding an AI-enabled tool into clinicians' daily workflows, guided by three central research questions. Through thematic analysis and iterative redesign, several insights were generated that directly address these questions and inform the development of a clinician-centered AI system.

First, regarding how clinicians prefer to interact with AI, findings emphasized the importance of optional engagement rather than default automation. Introducing an “AI Assist” button ensures that clinicians can choose when and how to interact with AI rather than having AI-generated suggestions imposed upon them. This aligns with the need for clinician autonomy, reinforcing that AI should serve as a supportive rather than a directive tool in daily practice.

Second, in examining when AI is most helpful versus when it may feel intrusive, the analysis revealed that clinicians are more receptive to AI in structured, well-defined scenarios, such as the UTI case, where treatment pathways are standardized.

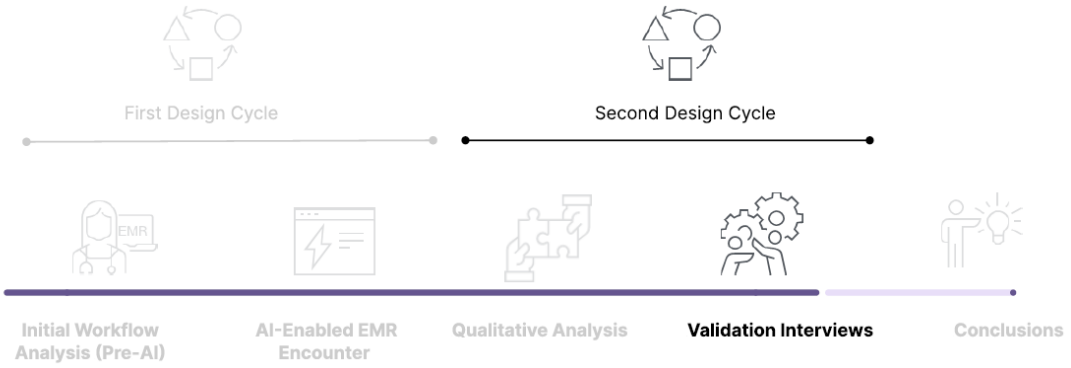
Conversely, clinicians expressed concerns about AI-generated treatment suggestions being overly simplistic or prescriptive in complex cases, such as mental health scenarios. To address this, the redesign ensures that AI recommendations are context-aware, reinforcing clinician expertise rather than replacing it.

Finally, this study highlights several critical elements in response to the question of what design considerations are necessary to build trust, enhance usability, and ensure seamless integration. Clear labeling, audit trails, and color-coded distinctions between AI-generated and clinician-entered content reinforced trust and transparency. Usability enhancements included a confidence score toggle to accommodate varying preferences. At the same time, workflow integration was supported through non-disruptive AI engagement that aligns with existing clinical practices, such as the SOAP format.

While this redesign does not represent a perfect or final solution, it reflects a human-centered approach that prioritizes gradual adoption, clinician autonomy, and seamless AI integration into primary care. Future research and usability testing will be essential to refine these design elements further and evaluate their impact on real-world clinician adoption and patient outcomes.

Chapter 6

Validation Interviews



To evaluate the effectiveness of the redesigned AI-enabled EMR interface, a series of validation interviews were conducted with primary care clinicians who participated in the initial round of interviews. These sessions aimed to determine whether the updated mock-ups had improved from the initial AI-enabled mock-ups initially shown in the first interview.

The validation interviews were conducted to assess clinician perceptions of the redesigned AI-enabled EMR encounter, with 10 out of the 14 original participants agreeing to participate. Each interview followed a semi-structured format and was conducted virtually via Microsoft Teams for approximately 30 minutes. Participants were first reintroduced to the original scenarios (Mental Health and UTI), shown with the AI-enabled EMR encounter from **Figure 3**, to refresh their memory. After, they were guided through a walkthrough of the redesigned scenarios as seen in **Figure 6** to **Figure 11**, where key design changes and their underlying rationale were explained. Clinicians were encouraged to provide candid feedback on usability, workflow integration, and the overall effectiveness of the redesign.

All interviews were automatically transcribed, and no formal coding process was applied. Instead, key thoughts and takeaways were documented immediately after each session to capture clinicians' perspectives. Responses were not systematically categorized into themes but reviewed holistically to identify common patterns, concerns, and approval areas. Since the interviews primarily consisted of answers to direct redesign areas, feedback was summarized in a tabulated format, capturing clinician insights on specific aspects of the redesign. The findings are presented as observations and direct quotes, highlighting recurring sentiments and individual perspectives. Additionally, clinicians completed a System Usability Scale (SUS) assessment to quantify their perceptions of the redesigned experience.

While this process was not intended to serve as a definitive validation of the redesigns, it provided valuable insight into how clinicians perceived the updated interface. By revisiting the original mock-ups before reviewing the redesigns, participants could reflect on specific improvements and articulate whether the changes aligned with their expectations and workflow needs. These discussions helped identify the remaining areas of friction and potential refinements. This iterative approach ensured that the redesign process remained user-driven and responsive to real-world clinical considerations rather than purely theoretical design choices.

The findings from these validation interviews reaffirmed many of the themes identified in the initial interviews. Clinicians continued to gravitate toward features perceived as most useful, emphasizing AI's potential to enhance efficiency, support how they practice, and streamline workflows. Overall, the redesigned mock-ups were well-received, and clinicians expressed optimism about AI's ability to save time, reduce cognitive load, and improve patient care delivery.

"You need to make it easier for people to find family practice interesting and not make it feel overwhelming. This could really be one way to do it" (P13)

"If AI can save me time on charting and documentation, it's a win. I like how the summary looks here" (P10)

To effectively capture clinician feedback from the validation interviews, responses were organized by feature, as seen in **Table 7**. This structure aligns with the interview approach, where each session focuses on specific elements of the AI-enabled EMR redesign. Responses were grouped based on individual features like the AI Assist button, confidence scores, color-coded AI suggestions, and workflow customization. The tabulated summary highlights clinician preferences, usability concerns, and areas for improvement in a structured and actionable manner. A more detailed version of the tabulated results of this interview are available in **Appendix F part A and Appendix F part B.**

This organization allows for a direct comparison of feedback across different features, illustrating which aspects were well-received and which required further refinement. Additionally, this format ensures that each feature is evaluated independently, providing clearer insights into specific aspects of the redesign. Direct quotes were included to preserve clinician perspectives.

Table 7 Summary of Tabulated Clinician Responses Organized by Feature

Feature	Summary of Clinician Feedback	Quotes
AI Assist Button (Opt-in AI Assistance)	Clinicians strongly preferred controlling when AI is activated rather than AI suggestions appearing automatically. They valued the ability to engage AI only when needed. This button was well received by nine of the participants. One participant specified that he did not see the button as necessary to the workflow.	*"I think this looks a little bit cleaner, having the AI assist button where you can enable that but if you don't want it, you don't have to."* (P3) *"Technology should assist, not dictate, having control over AI activation makes a big difference"* (P6)
Color-Coded AI vs. Clinician-Entered Text	All participants found it helpful to distinguish between AI-generated content and their own entries visually. Four clinicians specified that once they sign off on an AI-enabled encounter, they want all of the text to be one color since they made the final decision.	*"I love the fact that you're distinguishing what's mine versus what was AI generated."* (P10) *"It makes reviewing notes easier—I don't have to second-guess what I wrote versus*

		what AI suggested." (P4)
Confidence Scores	Similar to the initial findings in the thematic analysis, the confidence scores were still met with uncertainty for their utility. 6 of the participants stated that they liked the way it was displayed in the redesign in comparison to the numeric value of the first mock-up.	*"I don't really know what the confidence score means, so I'm not sure if I'd trust it."* (P3) *"It might be useful, but only if it's explained clearly and optional."* (P10) "*I appreciate how much clearer reading this confidence score actually is"* (P1)
AI-Generated Summary	Similar to the findings in the thematic analysis, context sensitivity was still a factor when clinicians reviewed the AI-generated summary. All clinicians stated were open to AI suggestions for structured, routine cases (e.g., UTIs) but felt that AI should not be as	*"For UTI prescriptions, this is useful. For mental health, I'd still rather make my own assessment."* (P5) *"Language guides thinking. If AI suggests something,*

	prescriptive in complex cases (e.g., mental health).	*people might follow it without questioning."* (P6)
Customization & Personalization	Clinicians emphasized the need to tailor AI recommendations to their workflow, noting that every doctor has a different practice style. An API for additional customization was seen as a valuable future feature.	*"Doctors like to customize things their own way. You need flexibility in the format of documentation and decision support."* (P1) *"Having an API that people could use to build their own things, I think, would be very helpful. "*(P3)
Potential Volume of Clicks with Redesigned Workflow	No negative comments were made about the proposed click-through workflow with the mock-ups.	*"The fewer fields you have, the easier it is to navigate, fewer clicks, less work."* (P1) *"If AI can save me even an hour a day, that's time I can*

		spend with my family." (P13)
Volume of Information Presented in AI-Generated Summary	All clinicians stated their comfort with the volume of information in the redesigned AI-enabled EMR encounter.	*"The more information that's there upfront, the more the clinician feels responsible for it."* (P6) *"Show me something actionable—share the handout or follow up—but don't overcomplicate it."* (P7)
AI-Generated Summary – Decision Support	The information provided in the generated summary from a decision-support perspective was well-received by 9 of the 10 participants. 1 of the participants did indicate that they did not see the summary as useful anymore, given how quickly they evolved using an AI scribe within their EMR in the last few months. This highlights how quickly the	*"Avoid being too prescriptive. It should be a suggestion, not an order."* (P6) *"I think it largely does need to sit as an assistive augmented tool with a driver."* (P4)

	technological landscape can shift with AI.	
Medication Suggestion for UTI Scenario (Listing a Generic Drug Recommendation)	Five of the clinicians explicitly said that they were good with the AI making generic suggestions of nitrofurantoin. The other five clinicians cautioned against specifically naming medications, concerned that the AI was taking a step too far for their comfort.	*"There's a potential here to improve quality. For instance, nitrofurantoin is a better first-line choice than ciprofloxacin, which is overused."* (P6)

Discussion

While the validation interviews provided valuable insights into clinician perceptions of the redesigned AI-enabled EMR interface, several limitations must be acknowledged. These constraints may impact the generalizability of the findings and highlight areas for future research and refinement.

Although 10 of the 14 original participants agreed to participate in the validation interviews, this still represents a relatively small sample size. The perspectives captured, while insightful, may not fully represent the diversity of primary care clinicians. Additionally, the four participants who did not participate may have had different experiences or perspectives that were not accounted for.

All interviews were conducted via Microsoft Teams, meaning clinicians interacted with static mock-ups rather than a fully functional prototype. While the walkthroughs helped illustrate the intended AI workflow, participants could not engage with the system in a real-world clinical environment. This may have limited their ability to

fully assess usability, particularly regarding efficiency, cognitive load, and real-time decision-making. Future studies should incorporate live prototype testing within clinical settings to validate usability in practice.

Unlike the initial round of interviews, the validation interviews were not formally coded or analyzed using a thematic framework. Instead, feedback was summarized holistically and tabulated by feature. While this approach effectively captured feature-specific feedback, it may have missed underlying themes that could have emerged through a more systematic qualitative analysis. A more structured thematic approach could provide deeper insights into recurring concerns, behavioral patterns, and clinician expectations beyond individual feature evaluations.

One notable insight from the interviews was that clinician expectations and experiences with AI tools are evolving rapidly. For example, one participant indicated that their perception of AI-generated summaries had already shifted within a few months due to exposure to AI scribes within their EMR. This highlights the challenge of assessing AI adoption in a dynamic technological landscape where user familiarity and expectations may change over time. Future research should consider longitudinal studies to track how clinician attitudes toward AI evolve with continued exposure.

While the iterative design process ensured clinician feedback was incorporated into the redesigns, these interviews remain an early-stage evaluation. The real-world implementation of AI-enhanced EMRs could uncover new usability challenges, unintended workflow disruptions, or adoption barriers that were not apparent in static evaluations. A next step would be conducting pilot deployments of AI-assisted EMRs

in clinical environments, allowing for real-time feedback loops and iterative refinements based on usage patterns.

Despite these limitations, the validation interviews provided valuable clinician-driven insights into AI adoption in primary care EMRs. The findings reaffirm the importance of customization, clinician control, workflow alignment, and transparency in AI-assisted decision support tools. However, future studies should incorporate larger, more diverse samples, live usability testing, and longitudinal assessments to further refine AI's role in enhancing clinical workflows. Addressing these limitations will be essential for ensuring that AI-driven tools are not only technically effective but also clinically meaningful and widely adopted in real-world practice.

System Usability Scale Analysis

The System Usability Scale (SUS) is a widely validated 5-point Likert scale that provides an efficient and reliable method for measuring users' perceptions of a system's usability. Its simplicity and adaptability suit various technologies and contexts, including novel concepts such as AI-enabled electronic medical records (EMR). For this study, SUS was chosen to evaluate the redesigned AI-enabled EMR encounter screens due to its interface-agnostic nature and ability to offer meaningful insights even in the absence of AI-specific usability scales.

All participants in the validation interviews completed the SUS via a Google Form immediately after their interviews, typically within three minutes of concluding. The participant pool consisted of 10 of the 14 clinicians participating in the first interview round. Based on prior research, sample sizes of 8–12 participants are generally considered reliable for usability evaluations (Alroobaea & Mayhew, 2014; Macfield, 2021).

Final SUS scores range from 0 to 100, with higher scores indicating better usability. The scoring process involves summing adjusted responses for odd-numbered questions (1, 3, 5, 7, and 9, subtracting 1 from each score) and even-numbered questions (2, 4, 6, 8, and 10, adding 1 to each score). The total is then multiplied by 2.5 to produce the final score. While SUS provides a holistic measure of usability, individual analysis of specific questions is less meaningful, as prior studies indicate minimal differentiation among responses to individual statements (Brooke, 2013)

This study employed an empirically derived framework (Bangor et al., 2008) to interpret scores. Scores below 50 signify serious usability challenges and potential rejection of the system. Figure 12 provides a visual representation of this interpretive framework**.**

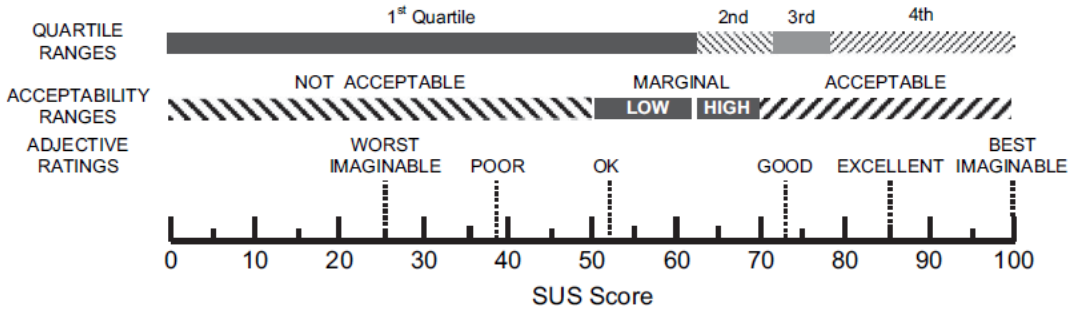


Figure 13. "A comparison of mean system usability scale scores by quartile, adjective ratings, and the acceptability of the overall SUS score" by Bangor et al., (2008).

The System Usability Scale (SUS) was chosen to evaluate the AI-enabled EMR encounter screens due to its interface-agnostic nature, making it adaptable across various technologies and contexts (Brooke, 1996). The screens presented to clinicians introduced an AI-enabled EMR encounter, which currently lacks standardized, AI-specific usability evaluation tools. In this context, SUS provided a reliable and well-established framework to assess usability without bias toward a particular interface type, allowing for an initial, objective evaluation of clinicians' interactions with AI in the EMR environment. Its flexibility ensured that insights could still be derived about usability in this setting, even in the absence of specialized AI usability scales.

However, while SUS is effective as a general usability metric, it has inherent limitations when applied to AI-driven healthcare interfaces. The binary nature of its response scale and lack of granularity can overlook domain-specific concerns such as interpretability, trust, and cognitive burden—all critical factors in clinician adoption of AI tools (Kortum & Bangor, 2013). Additionally, SUS does not explicitly account for transparency, explainability, or decision support, which are crucial for AI usability in clinical decision-making (Lindgaard & Dudek, 2003). This limitation suggests that while SUS provides a broad usability benchmark, it may not fully capture the nuanced interactions between clinicians and AI systems, especially

regarding trust in AI-generated recommendations and workflow integration challenges.

The SUS evaluation of the AI-enabled EMR encounter screens resulted in an overall score of 68.33, placing it within the "good" usability range, leaning toward acceptability (Sauro, 2011). While this score does not indicate a “best imaginable” system, it provides a solid foundation for usability in an emerging AI-integrated EMR context, offering valuable insights for further refinement. Given that AI in EMRs is still an evolving area, this usability assessment serves as a baseline measure for future iterations rather than an endpoint.

The validation interviews focused on clinician perceptions, preferences, and expectations rather than quantifiable performance metrics. While the System Usability Scale (SUS) assessment provided some degree of structured evaluation, usability and efficiency improvements were not empirically measured through task-based testing. Future research should incorporate comparative studies where clinicians complete identical tasks with and without AI assistance to assess time savings, error reduction, and cognitive load.

Despite SUS’s applicability as a foundational usability measure, its limitations highlight the need for additional AI-specific evaluation tools to capture critical elements like interpretability, transparency, and clinician trust. Future research should explore complementary usability frameworks—such as the AI Trustworthiness Scale (Zhang et al., 2020) or the Explainability Satisfaction Score (Eiband et al., 2018)—to provide a more holistic understanding of how AI is perceived and adopted in clinical settings. By integrating usability benchmarks with AI-specific dimensions, future iterations of AI-driven EMRs can be more precisely tailored to clinician needs, ultimately enhancing trust, efficiency, and adoption in healthcare environments.

Conclusion

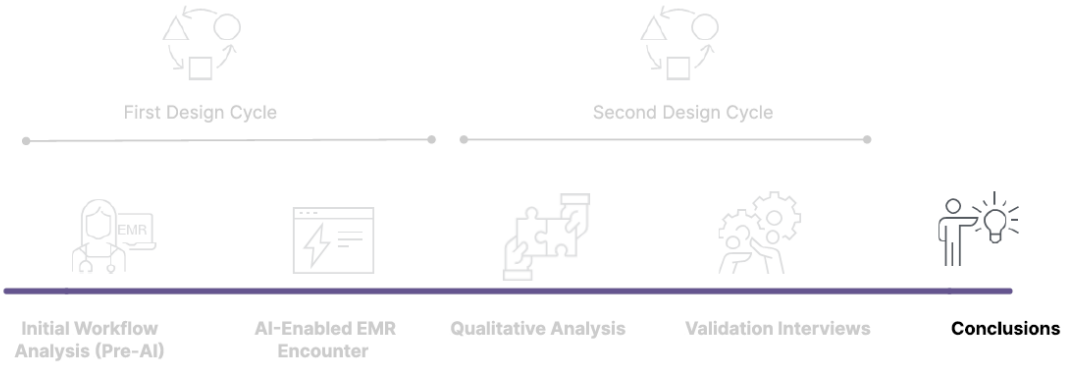
The validation interviews provided key insights into clinician perceptions of the redesigned AI-enabled EMR interface, emphasizing usability, workflow integration, and clinician autonomy as critical factors in AI adoption. Clinicians were generally optimistic about AI's potential to improve efficiency, reduce cognitive load, and enhance decision-making but highlighted the need for clear differentiation between AI-generated and clinician-entered content, transparency in recommendations, and customization options to align with individual practice styles.

The structured feedback approach identified well-received features, such as opt-in AI assistance and color-coded AI-generated text, while revealing areas for improvement, including the interpretation of confidence scores and balancing automation with clinician control. The System Usability Scale (SUS) assessment, which produced a score of 68.33, indicated good usability but room for refinement, particularly in addressing AI-specific concerns like transparency and interpretability.

Several limitations must be acknowledged, including the reliance on static mock-ups instead of live system testing and the absence of structured thematic coding in response analysis. Future research should explore larger clinician samples, real-world usability testing, and longitudinal studies to track AI adoption over time.

These findings reinforce that AI must be transparent, customizable, and seamlessly integrated into clinical workflows to earn clinician trust. Moving forward, pilot implementations in real clinical settings, integration with evolving AI evaluation frameworks, and ongoing refinement will be essential for ensuring that AI-driven EMRs effectively support primary care delivery.

Chapter 7
Contributions and Conclusions



This dissertation explores whether a user-centered design approach can enhance the adoption of an AI-enabled encounter module within a primary care EMR. It provides valuable insights into primary care clinicians' expectations of AI and its potential impact on their ability to integrate it into clinical encounters. Qualitative feedback from interviews generated key themes that informed the redesign.

Research Question Findings

To support the overarching goal of this research, this study explored three key sub-questions, which are examined in the following section.

How would clinicians prefer to interact with an AI-enabled tool embedded in a module they use for daily practice?

Clinicians expressed a strong preference for AI tools that are optional, non-intrusive, and seamlessly integrated into existing workflows rather than automated or disruptive. Depending on the clinical context, they valued the ability to engage with AI at their discretion. For example, their approach to using AI in a mental health visit, which often requires deeper, nuanced patient interactions, differed significantly from how they would use it in a straightforward UTI visit, where efficiency and structured decision-making take priority. This need for context-aware flexibility supports the

importance of AI tools that clinicians can activate when needed rather than being imposed upon their workflow.

Beyond optionality, clinicians also emphasized the importance of suggestive rather than prescriptive AI functions. They preferred AI-generated language that provides recommendations rather than dictates decisions, reinforcing their clinical autonomy. There was particular concern that overly directive AI outputs could create medical legal risks, especially given the lack of formal guidance, regulation, or support from governing bodies regarding AI use in clinical practice. The fear of liability and accountability issues made it clear that AI must act as an assistive tool rather than an authoritative entity.

To align with these preferences, the redesigned encounter module introduced an "AI Assist" button, ensuring that AI activation remains entirely clinician-driven. This feature reinforced their sense of control and minimized unnecessary interruptions in patient encounters. Additionally, clinicians favored color-coded elements and structured summaries, particularly those aligned with their existing documentation habits, such as the SOAP format (Subjective, Objective, Assessment, Plan). These design choices helped ensure that AI outputs were intuitive and easy to interpret without requiring a steep learning curve. Context-sensitivity was also considered in how AI-generated outputs made recommendations. For example, in the redesigned interface, the UTI scenario provided a generic drug recommendation to prioritize efficiency, whereas the mental health scenario avoided explicit recommendations, instead encouraging discussion.

Transparency was another critical consideration. Clinicians wanted clear visibility into how AI-generated suggestions were formed, particularly regarding confidence scores and data sources. However, they also recognized the need to strike a balance; while some clinicians desired detailed insights into AI-generated recommendations,

others preferred a more streamlined interface that did not overwhelm their workflow. The introduction of AI confidence scores was met with curiosity, but its usefulness depended on how easily interpretable it was at the moment. To address this, the ability to toggle confidence scores on or off was introduced, allowing each clinician to customize AI visibility based on their workflow preferences and decision-making style.

By integrating these clinician-driven design elements, the study highlights the importance of flexibility, transparency, and autonomy in AI-enabled clinical tools, ensuring they function as supportive aids rather than intrusive systems.

One limitation of this study is the potential lack of diversity in the clinician sample, as participants may have been drawn from specific practice settings or geographic regions, limiting the generalizability of their preferences. Recruitment was particularly challenging, with several primary care clinicians citing their high workloads as barriers to participation. Despite efforts to engage a broad range of clinicians through multiple recruitment channels, including professional networks, email outreach, and direct invitations, the demands of clinical practice limited availability and may have contributed to a sample that does not fully capture the perspectives of all primary care settings. Additionally, while the study gathered insights through scenario-based feedback and mock-ups, real-world interactions may differ due to the pressures of time constraints, patient variability, and evolving clinical workflows. Another limitation is the absence of longitudinal data; the study primarily captured initial impressions of AI integration. However, clinician preferences and adoption patterns may change over time as they gain more experience with these tools. Furthermore, while clinicians expressed preferences for AI features such as optionality and transparency, the study did not extensively explore the technical feasibility of implementing these design elements within existing EMR systems, which may present challenges in real-world deployment.

Despite these limitations, this study provides valuable insights into how primary care clinicians envision AI as a supportive rather than directive tool in their daily workflows. The findings highlight the critical design principles for AI to seamlessly integrate into EMR encounters by emphasizing optionality, non-intrusiveness, and alignment with existing clinical processes. Understanding clinician preferences ensures that AI-enabled tools enhance decision-making without undermining autonomy or increasing cognitive burden. These contributions offer a foundation for future development efforts, ensuring that AI adoption in primary care remains clinician-driven and patient-centered.

In what scenarios do clinicians find AI most helpful, and under what circumstances might they perceive it as intrusive?

Clinicians found AI most useful and effective in structured, routine cases with clear decision pathways, where standardized guidelines could drive recommendations. Clinicians also noted that AI can serve as an effective second set of eyes, assisting them in keeping up with continuously evolving guidelines and ensuring adherence to best practices. One commonly cited example was uncomplicated UTIs, where AI could suggest appropriate medications and dosages based on clinical best practices and antibiotic stewardship guidelines. In these cases, AI was seen as an efficiency booster, reducing cognitive load by streamlining decision-making without disrupting the clinician's workflow.

Clinicians also saw value in AI for preventative care, particularly for reminders and alerts. AI could flag overdue screenings, prompt adherence to vaccination schedules, and ensure proper antibiotic stewardship; all tasks that align well with structured decision rules. In these contexts, AI served as a safety net, helping clinicians keep track of preventive measures that might otherwise be overlooked in a busy practice.

However, clinicians expressed significant concerns about AI's role in more complex, nuanced cases, particularly in mental health visits. In these encounters, treatment planning requires deeper patient-clinician dialogue, clinical intuition, and individualized care strategies, factors that they perceive AI is not well-equipped to address. AI-generated diagnostic suggestions for mental health conditions were met with skepticism, as clinicians worried about oversimplification and an over-reliance on quantitative screening tools like the GAD-7 and PHQ-9. While these tools can provide helpful data points, clinicians emphasized that patient history, lived experiences, and contextual nuances are critical to mental health care and cannot be effectively captured by AI-driven outputs alone.

Another major concern was decision autonomy. Clinicians wanted AI to assist, not override, their clinical judgment. They feared that overly assertive AI recommendations could create subtle pressure to follow suggestions, even when they might not align with the best care for an individual patient. Additionally, they raised concerns about over-documentation, where AI might clutter clinical notes with excessive or unnecessary suggestions, leading to documentation bloat rather than workflow efficiency. Instead of reducing workload, poorly designed AI interventions had the potential to increase cognitive burden by requiring clinicians to sort through irrelevant or redundant information.

These findings ultimately point to the need for context-sensitive AI design, ensuring that AI tools are helpful in structured cases while remaining discreet and non-intrusive in complex, judgment-driven scenarios. The key takeaway is that AI should function as a supportive aid when appropriate, and enhancing clinician decision-making rather than actively interfering with their expertise and purpose.

A key limitation of this study is that it focused on a subset of clinical scenarios, such as uncomplicated UTIs and mental health visits, without extensively exploring how AI might function in other contexts, such as chronic disease management. AI's perceived usefulness or intrusiveness may vary significantly depending on the case's complexity, the level of clinical uncertainty, and the need for patient-centered dialogue. Additionally, while the study examined AI preferences based on theoretical or prototype-based interactions, it did not account for the variability in AI capabilities currently available in real-world systems. More advanced AI may be perceived as more helpful, while less sophisticated models might be viewed as intrusive or unreliable. Another limitation is the absence of the patient's perspective. While clinicians' perceptions of AI intrusiveness were explored, patient comfort and expectations regarding AI involvement in their care could significantly influence clinicians' acceptance of these tools.

While this study did not capture every possible clinical scenario, it provides an enlightening starting point for understanding the situational nuances of AI integration in primary care. By identifying structured, routine cases as ideal for AI support while recognizing the complexities of nuanced patient-centered interactions, the study contributes to a growing body of research that prioritizes AI's role as an augmentative tool rather than a replacement for clinical judgment. These findings inform the design of AI systems that are contextually aware, ensuring they provide meaningful assistance where appropriate while minimizing the risk of disruption in more complex, nuanced encounters.

What design considerations are necessary to build trust, enhance usability, and ensure seamless integration of AI systems into primary care workflows?

This study has thematically identified the following key design principles that address the above research question.

Trust and Transparency Design Principle

AI must earn clinician trust through consistent high performance, transparency, and alignment with clinical decision-making. It should deliver accurate, explainable, and auditable recommendations while minimizing cognitive burden. Concepts such as confidence scores and training data, which are not typically covered in medical education, must be interpretable and optional, with assistive features like tooltips providing clear, accessible explanations as needed.

Clinician Authority and Decision-Making Design Principle

AI must be designed to reinforce clinician autonomy, offering evidence-based support without overriding decisions. Suggestions should be actionable, practical, and tailored to real-world clinical challenges, ensuring AI remains a safeguard and an intelligent assistant, not a directive authority.

Complexity Sensitivity Design Principle

AI should adapt to case complexity, providing direct support for simple cases while deferring to clinician judgment for nuanced scenarios. It must integrate past interactions and patient history to offer context-aware recommendations that align with continuity of care.

Workflow Efficiency and Integration Design Principle

AI must seamlessly integrate into clinical workflows, providing concise, structured, and immediately relevant insights. Features should enhance efficiency, minimize

unnecessary interactions, and align with natural reading flows. AI must function as a low-friction tool that complements existing EMR systems rather than complicating them.

Legal and Ethical Concerns Design Principle

AI must maintain clear role delineation, ensuring clinician-generated input and AI-suggested content are visually distinguishable and auditable. AI adoption should be supported by policies that protect clinicians from liability, ensuring that patient consent processes are clear, streamlined, and non-intrusive.

Human-Centered Design Principle

AI must enhance, not replace, the human connection in patient care. It should operate unobtrusively in the background, allowing clinicians to maintain cognitive engagement and patient-centered communication. AI should serve as a cognitive partner, reinforcing human expertise without fostering dependence.

These design principles collectively highlight the essential factors for fostering AI adoption in primary care by prioritizing trust, clinician autonomy, complexity sensitivity, workflow efficiency, legal clarity, and human-centered integration. These principles provide a structured foundation for designing AI-enabled EMR systems that enhance efficiency, preserve clinical judgment, and support patient-centered care.

While this study identified key design principles, it did not include direct usability testing with a fully functional AI system in a live EMR environment. Without real-time interaction, certain usability challenges' details, such as cognitive load, navigation difficulties, or unexpected workflow disruptions, may not have been fully captured.

Additionally, although the study explored how clinicians perceive trust in AI, deeper systemic issues related to accountability, liability, and regulatory oversight were not comprehensively addressed. Trust in AI extends beyond usability; it also requires clear guidelines on responsibility in cases where AI-generated recommendations may be incorrect or legally contentious. Furthermore, the study did not examine broader organizational and systemic barriers to AI adoption, such as the financial cost of implementation, the need for clinician training, or IT infrastructure limitations, all of which could impact the successful integration of AI into daily workflows.

Conclusion

Despite this research's constraints, the findings from the sub-questions provide a foundational understanding of the design principles that support the adoption of AI-enabled EMR encounters in Canadian primary care. This study demonstrated that a user-centered design approach effectively identifies key elements that build clinician trust and improve usability, ultimately contributing to the potential for greater adoption.

These insights offer actionable guidance for developers, policymakers, and healthcare organizations, ensuring AI implementation aligns with the practical realities of primary care. More broadly, this research contributes to the ongoing discourse on AI in healthcare, emphasizing the critical role of human-centered design in making AI an intuitive, trustworthy, and supportive tool in patient care.

Contributions

AI integration in healthcare is no longer a question of if but how. As industries increasingly adopt AI to improve efficiency and user experiences, healthcare is poised to follow. Given the critical role of the clinical encounter in primary care and ongoing challenges in the Canadian healthcare system, AI tools must enhance rather than disrupt a clinician's ability to provide care. The rapid expansion of AI research underscores the urgency of this effort. Bibliometric analyses indicate that publications in this field have doubled approximately every two years, reflecting growing attention from academia and industry (Guo et al., 2020). More recent analyses (2023-2024) confirm this trend, with research output surging from 153 publications in 2013 to over 4,500 in 2023 (Senthil et al., 2024). This growth highlights increased research activity and a shift toward real-world applications, such as predictive modeling and personalized medicine, reinforcing the need for structured, evidence-based AI integration strategies.

This dissertation contributes to the growing literature on AI adoption in healthcare by offering practical design principles that address key clinician concerns. These include strategies for building trust, pathways for adoption through less complex clinical scenarios, and approaches that position AI in a supportive role, reinforcing clinicians' authority in decision-making. Prior research has emphasized the need for effective AI-physician collaboration while identifying persistent barriers to adoption. For instance, a randomized controlled trial found that while large language models (LLMs) outperformed individual physicians in diagnostic reasoning, they did not significantly improve clinician decision-making compared to conventional resources (Goh et al., 2024). This suggests that AI's technical capabilities alone are insufficient for adoption unless thoughtfully integrated into clinical workflows. Similarly, a study on GPT-4's role in clinical management reasoning found that AI-assisted physicians performed significantly better on expert-developed rubrics, reinforcing AI's capacity

to enhance decision-making (Goh et al., 2025). However, increased time spent per case raised concerns about workflow inefficiencies and over-reliance on AI-generated recommendations. These findings align with this dissertation's emphasis on workflow optimization and the need for clinician-centered AI designs that streamline decision-making without adding cognitive burden.

Beyond performance concerns, the issues of trust and workflow integration have emerged as major barriers to AI adoption. Agarwal, Dugas, and Gao (2023) identify skepticism surrounding AI's transparency, ethical integrity, and impact on clinical judgment. They argue that AI must complement rather than replace human intelligence, requiring interpretability, clinician autonomy, and seamless integration (Agarwal et al., 2023). Ullah and Ali (2025) further highlight persistent challenges in AI implementation, including concerns over clinician skepticism and workflow disruptions. This dissertation applies a user-centered design approach that directly incorporates clinician feedback into AI interface development to address these barriers. The findings reinforce the importance of trust-building mechanisms, such as transparency in AI decision-making, non-intrusive support, and clear delineation between AI-generated content and clinician-entered data.

Bridging the AI Evaluation Gap

Recent research has called for AI evaluation frameworks that extend beyond technical accuracy to consider real-world usability and clinician experience (Al-Ansari et al., 2024). A systematic review of user-centered AI evaluation methods found that workflow disruptions, unclear success metrics, and challenges in user engagement often hinder adoption. These findings align with this study's thematic analysis, which underscores the importance of embedding AI tools in ways that support, rather than complicate, primary care delivery. While broader AI evaluation frameworks exist, the unique demands of primary care necessitate domain-specific

approaches that assess AI's impact on clinician workflows and patient interactions. By focusing on clinician perspectives, this research contributes to ongoing discussions on how AI should be evaluated in real-world healthcare settings.

Building on the gaps in AI evaluation frameworks highlighted by Kueper et al. (2020), this study advances practical methodologies for assessing AI usability, clinical effectiveness, and workflow integration in primary care. By adopting a user-centered design approach, this research ensures that usability and workflow considerations are embedded into AI development rather than treated as secondary concerns. The findings provide insights into the factors influencing clinician adoption, offering concrete design recommendations for AI-enabled EMRs.

Aligning with Primary Care Principles

The College of Family Physicians of Canada (CFPC, 2024) advocates for AI development that enhances primary care service delivery, clinician well-being, and health equity. This research reinforces the importance of ethical, sustainable, and user-centered AI implementation by examining the practical challenges of integrating AI into clinical workflows. The findings support CFPC's emphasis on rigorous AI evaluation, demonstrating that aligning AI design with primary care principles leads to improved clinician engagement, trust, and long-term adoption.

Emerging Frameworks and Future Directions

Since the completion of this research, a new evaluation framework for AI in healthcare, AI for IMPACTS, was introduced (Jacob et al., 2025). This framework shifts the focus of AI assessment beyond technical validation to broader clinical, social, and economic dimensions, addressing many of the limitations identified in previous AI evaluations (Jacob et al., 2025). While my research predates this framework, it underscores similar concerns regarding AI evaluation gaps, particularly in real-world implementation challenges. The emergence of AI for IMPACTS

highlights the timeliness of this dissertation, reinforcing the need for ongoing refinement in AI evaluation methodologies.

The results discussed in the IMPACTS framework align closely with key themes generated from this study. Both emphasize seamless AI integration, clinician trust, and human oversight in AI-assisted decision-making. This study extends the IMPACTS framework by directly grounding these evaluation criteria in qualitative insights from clinicians, revealing specific usability challenges and redesign implications. While IMPACTS provides a high-level assessment structure, this research contributes a more granular understanding of how these principles manifest in real-world AI adoption, reinforcing the need for structured yet clinician-informed AI design.

The key insights from this study establish a foundational understanding of clinicians' perspectives, shaping the thoughtful introduction of AI tools into practice.
This chapter highlights the contributions of this work across theoretical and practical domains and discusses this research's limitations and lessons learned.

Scientific Contributions

This research builds upon and extends existing theories related to AI in healthcare, user-centered design, and clinician workflow integration. Previous scoping reviews (Kueper et al., 2020) and studies have examined Canadian primary care clinicians' perceptions of AI readiness, highlighting both enthusiasm and hesitancy toward AI integration. One particularly relevant study used thematic analysis to reveal a fundamental disconnect between the envisioned uses of AI and the current reality in primary care settings (Terry et al., 2022). This misalignment highlights the significance of this dissertation in bridging the gap between user expectations and tangible design outcomes by ensuring AI tools are not only technically viable but also align with clinician workflows and expectations. Notably, this same study also found

that primary care clinicians identified the “co-creation” of AI-enabled tools as a necessary and foundational element for driving AI adoption (Terry et al., 2022)., reinforcing the need for AI solutions that are iteratively designed with direct clinician input rather than imposed as rigid, top-down implementations.

Another study conducted in Ontario primary care investigated AI’s potential for deriving patient social data but identified a critical prerequisite: before AI’s benefits can be realized, its design, deployment, and use must be co-developed with end users (Garies et al., 2024). This finding reinforces the central argument of this dissertation, while there is growing interest in AI’s role in primary care, its success depends on ensuring that AI tools are designed effectively, with clinician input at every stage. Without careful attention to usability, workflow integration, and clinician trust, AI adoption in primary care remains limited, not due to a lack of technological potential, but because of inadequate alignment with real-world clinical practice.

This sentiment is echoed in recent discussions within the Canadian primary care community. A family physician writing in the *Journal of the College of Family Physicians of Canada* emphasized the urgency of clinician involvement in shaping AI’s role in healthcare, stating: “We know AI is here to stay, and we need to be involved in ensuring it is applied in ways that benefit both patients and family physicians and in ways that enhance, rather than degrade, health equity.” (Green, 2024). This perspective reinforces the argument that AI adoption in primary care cannot be solely a technological initiative, it must be a collaborative effort that prioritizes usability, trust, and equitable healthcare delivery.

This dissertation also addresses several gaps identified in Canadian Family Physician’s report on AI in primary care, particularly the need for clinician-driven AI development and effective workflow integration. The report highlights AI’s potential to improve patient care but emphasizes its success depends on careful design,

deployment, and integration within real-world clinical environments (Kueper et al., 2024). Additionally, the report raises concerns about AI's impact on health equity and the importance of education and training for physicians adopting AI technologies. This dissertation contributes by exploring how AI can be designed to enhance, rather than disrupt, primary care workflows while also considering how clinicians can be supported in understanding and leveraging AI effectively.

Given the limited real-world testing of AI-based tools in primary care settings (Kueper et al., 2024) , this study makes a key contribution by highlighting complexity sensitivity as a critical AI design principle. By differentiating AI-assisted tasks based on case complexity, such as providing structured support for routine conditions like UTIs while limiting AI intervention in more nuanced cases like mental health, this research offers a foundational insight for future AI-enabled EMR design. These findings establish a starting point for further development, ensuring that AI integration is tailored to clinical complexity and decision-making needs rather than applying a one-size-fits-all approach.

Practical Contributions

This research produced tangible applications for software industry development, healthcare organizations, and policymakers.

A key strength of this study was its direct collaboration with an industry partner, which provided a real-world foundation for the research and ensured that the findings had practical relevance beyond theoretical exploration. Unlike many studies that investigate AI in clinical settings from a conceptual or experimental standpoint, this research was uniquely positioned to influence the redesign of a product actively used by clinicians in their daily workflows. This industry connection added depth and authenticity to the study, as participants were not simply reflecting on hypothetical AI

applications but engaging with an evolving system that could directly impact their practice. The ability to integrate clinician feedback into a tangible redesign process allowed for a more nuanced understanding of adoption barriers, trust-building factors, and workflow integration challenges, insights that might be less actionable in studies lacking direct ties to product development.

By grounding this work in both academic rigor and industry-driven application, the study contributes meaningful, real-world design considerations that can inform not only this specific product but broader discussions on AI implementation in primary care.

This study generated actionable design recommendations to enhance clinician autonomy and usability in AI-enabled EMR encounters. Key recommendations include introducing an AI Assist button, allowing clinicians to enable AI support on demand rather than imposing it by default. A confidence score toggle was also proposed, allowing clinicians to opt in or out based on their preferences, ensuring that AI remains unobtrusive in clinical workflows. Additionally, color-coded delineations were incorporated to improve readability and audibility, making it clear which content was AI-generated versus clinician-entered. Finally, context-aware AI suggestions were emphasized, ensuring that AI interventions align with case-specific appropriateness, reducing the risk of irrelevant or intrusive recommendations.

These findings highlighted a critical knowledge gap in clinicians' understanding and trust AI, underscoring the need for further research on AI decision-making in clinical practice. Through the design process, it became evident that clinicians expect AI to think in the same structured format, aligning with their established cognitive models and workflows. This expectation suggests that for AI to be truly trusted and seamlessly integrated, future research must explore how AI presents information and how clinicians interpret, evaluate, and act on AI-generated insights.

Moreover, if AI becomes an embedded part of primary care decision-making, there may be a need for curriculum shifts in medical education to equip clinicians with the skills to assess AI recommendations critically. Future research could examine how clinical reasoning models interact with AI decision support.

Ethical and Legal Contributions

This research underscores critical regulatory gaps in AI governance, particularly concerning liability, ethical use, and accountability in clinical decision-making (Darcel et al.). Clinicians expressed concerns about responsibility when AI-generated recommendations lead to errors, emphasizing the need for clear guidelines on oversight and legal protections. Without well-defined accountability structures, uncertainty over whether liability falls on the clinician, the healthcare institution, or the AI developer creates hesitation in adoption. To mitigate this, regulatory bodies should implement audit trails documenting AI-assisted decisions, ensuring transparency in clinical documentation and reducing legal ambiguity. Such mechanisms would allow for retrospective analysis of AI involvement in medical decision-making, helping clinicians defend their judgment while maintaining trust in AI's role as a supportive tool.

Beyond liability, ethical AI governance in primary care must prioritize bias mitigation, equitable AI recommendations, and informed consent frameworks. Policymakers must implement rigorous validation processes to ensure that AI-generated recommendations are clinically sound across diverse patient populations, reducing the risk of bias reinforcing health disparities. Additionally, informed consent policies must balance keeping patients informed and avoiding excessive administrative burdens for clinicians. Traditional consent models may need to evolve into streamlined, digital-first approaches that clearly communicate AI's role in care without disrupting workflow efficiency.

Interoperability also presents a major regulatory challenge, as AI tools must seamlessly integrate with existing electronic medical records (EMRs) to avoid workflow disruptions. Clinicians have voiced concerns that poorly integrated AI systems create additional cognitive load, ultimately slowing care rather than enhancing it. Standardized AI interfaces and data-sharing protocols should be established to create a consistent user experience across EMR systems. Healthcare IT developers must work alongside regulatory agencies to ensure that AI tools follow a uniform design standard, reducing clinicians' learning curves and ensuring AI enhances rather than complicates care delivery.

Furthermore, this research contributes to the broader discourse on AI's legal and ethical implications in primary care, particularly regarding documentation integrity and decision-making autonomy. Clinicians emphasized the need for a clear delineation between AI-generated content and clinician-modified input, reinforcing the importance of governance policies that maintain human oversight. Establishing visual and audit-friendly distinctions between AI-generated and clinician-entered data will ensure that AI remains a trusted, accountable tool in clinical practice. By addressing these regulatory gaps, AI adoption in primary care can move forward in a way that supports clinicians, protects patients, and upholds ethical medical standards.

Limitations and Future Research Directions

While this study provides valuable insights into AI integration in primary care EMRs, several areas remain unexplored or require further investigation to understand AI’s role in clinical decision-making fully.

One key limitation of this research is that while it captures clinician preferences and expectations, it does not assess longitudinal adoption patterns. Clinicians'

perspectives on AI integration may evolve over time as they gain more hands-on experience with these tools. Future research should explore how clinician attitudes, trust, and usage patterns shift with continued AI exposure, particularly as AI capabilities become more advanced. Longitudinal studies could assess whether AI-enabled EMRs improve efficiency, accuracy, and decision support in real-world practice beyond initial perceptions.

Additionally, while this study focused on user-centered design principles to enhance AI adoption, it did not investigate the cognitive decision-making processes that underlie clinician trust in AI recommendations. The findings suggest that clinicians expect AI to mirror their structured thinking patterns, which raises important questions about how AI should be designed to align with or complement clinician reasoning models. Future research should explore the intersection of AI-generated insights and clinical judgment, examining whether AI tools should adapt to existing cognitive frameworks or evolve clinician training to accommodate AI's analytical methods. If AI becomes an integral part of primary care, medical education curricula may need to shift to include AI literacy, critical evaluation of AI recommendations, and ethical decision-making in AI-supported environments.

Furthermore, while this study examined AI's perceived intrusiveness in certain clinical scenarios, such as mental health encounters, it did not investigate AI's potential role in more complex, high-risk decision-making contexts, such as emergency medicine, chronic disease management, or end-of-life care. Future research should assess whether AI can be effectively deployed in nuanced, ethically sensitive, or high-liability situations and what governance structures are required to mitigate risk, ensure patient safety, and maintain clinician autonomy.

Another limitation is that this study primarily focused on clinician perspectives without incorporating patient viewpoints on AI-assisted care. While clinicians play a

critical role in AI adoption, patient acceptance and comfort with AI-assisted decision-making are equally important. Further research should explore how patients perceive AI involvement in their care, whether they trust AI-generated recommendations, and how AI can be designed to support shared decision-making between clinicians and patients.

Finally, while the study's findings were framed through mock-ups, actual AI implementation in primary care will depend on technical constraints, regulatory frameworks, and vendor-specific integration challenges. Future research should examine the feasibility of embedding AI features within existing EMR platforms, considering factors such as data interoperability, information security, and the financial costs of AI adoption. Additionally, broader organizational and systemic barriers, such as institutional resistance, policy limitations, and IT infrastructure constraints, must be studied to develop practical strategies for scaling AI adoption across diverse healthcare settings.

Despite these limitations, this study establishes an essential foundation for clinician-driven AI design in primary care. Future research should build upon these findings to further refine AI usability, address cognitive and ethical considerations, and develop implementation frameworks that ensure AI enhances clinician decision-making without undermining autonomy or patient trust. By advancing these research directions, AI can evolve into a seamless, trusted, and effective clinical practice tool, supporting clinician expertise and patient-centered care.

Conclusion

This research represents a critical step toward integrating AI within primary care EMRs by addressing theoretical challenges and real-world implementation barriers. As large language models and AI-driven solutions advance at an unprecedented pace, the question is no longer if AI will transform healthcare but how to ensure its implementation is mindful, sustainable, and beneficial for clinicians and patients. This study moves beyond simply exploring AI's potential; it identifies practical pathways for adoption that allow primary care clinicians to engage with AI in intuitive, non-disruptive, and empowering ways rather than overwhelming.

Technology is only as effective as the humans who engage with it, and healthcare is no exception. The primary care sector faces mounting pressures, including growing patient loads, administrative burdens, and physician burnout. AI adoption is not just about introducing sophisticated tools but ensuring that these tools seamlessly integrate into clinician workflows, build trust over time, and enhance rather than complicate the clinical experience. This study underscores that successful AI integration demands more than technical capability, it requires careful attention to how clinicians perceive, trust, and interact with AI in daily practice. AI's full potential in primary care will only be realized if it adapts to individual clinician needs and enhances decision-making rather than imposing rigid processes.

Beyond individual adoption, this research provides a strategic roadmap for AI development and implementation, offering actionable insights into how AI can bridge the gap between healthcare and technology. By prioritizing usability, workflow alignment, clinician education, and policy frameworks, this study reinforces the need to design AI solutions around the end user rather than allowing technological advancement to dictate clinical practice. A core challenge in AI adoption is the disconnect between the languages healthcare professionals and technologists speak.

This research serves as a translator between these worlds, creating a shared framework that fosters collaboration and ensures AI tools drive meaningful outcomes rather than adding complexity.

Future research should build on these findings by focusing on real-world AI implementations, long-term usability testing, and the development of policy frameworks that clarify AI's role in clinical decision-making, liability, and interoperability. AI adoption in primary care must be approached as a continuous, iterative process that evolves alongside clinician needs, regulatory landscapes, and technological advancements. By bridging the gap between AI development and clinical practice, this study lays the groundwork for a more human-centered, scalable, and practical approach to AI adoption, one that prioritizes clinician autonomy, trust, and real-world usability while ensuring AI serves as a tool for enhancing, not replacing, clinical expertise.

References

Abdullah, A., Liew, S. M., Hanafi, N. S., Ng, C. J., Lai, P. S., Chia, Y. C., & Chu Kiong, L. (2016). What influences patients’ acceptance of a blood pressure telemonitoring service in primary care? A qualitative study. *Patient Preference and Adherence*, 99. https://doi.org/10.2147/ppa.s94687

Agarwal, R., Dugas, M., & Gao, G. (Gordon). (2023). Augmenting physicians with Artificial Intelligence to Transform Healthcare: Challenges and opportunities. *Journal of Economics & Management Strategy*, *33*(2), 360–374. https://doi.org/10.1111/jems.12555

Al-Ansari, N., Al-Thani, D., & Al-Mansoori, R. S. (2024). User-centered evaluation of Explainable Artificial Intelligence (XAI): A systematic literature review. *Human Behavior and Emerging Technologies*, *2024*(1). https://doi.org/10.1155/2024/4628855

Alroobaea, R., & Mayhew, P. J. (2014). How many participants are really enough for usability studies? *2014 Science and Information Conference*. https://doi.org/10.1109/sai.2014.6918171

Aronson, P. L., Politi, M. C., Schaeffer, P., Fleischer, E., Shapiro, E. D., Niccolai, L. M., Alpern, E. R., Bernstein, S. L., & Fraenkel, L. (2020). Development of an app to facilitate communication and shared decision-making with parents of febrile infants ≤ 60 days old. *Academic Emergency Medicine*, *28*(1), 46–59. https://doi.org/10.1111/acem.14082

Ayanso, A., Herath, T. C., & O'Brien, N. (2015). Understanding continuance intentions of physicians with Electronic Medical Records (EMR): An expectancy-confirmation perspective. *Decision Support Systems*, *77*, 112–122. https://doi.org/10.1016/j.dss.2015.06.003

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction*, *24*(6), 574–594. https://doi.org/10.1080/10447310802205776

Baron, R., Perrot, S., Guillemin, I., Alegre, C., Dias-Barbosa, C., Choy, E., Gilet, H., Cruccu, G., Desmeules, J., Margaux, J., Richards, S., Serra, E., Spaeth, M., & Arnould, B. (2014). Improving the primary care physicians' decision making for fibromyalgia in clinical practice: Development and validation of the fibromyalgia detection (fibrodetect®) screening tool. *Health and Quality of Life Outcomes*, *12*(1). https://doi.org/10.1186/s12955-014-0128-x

Brooke, J. (2013). SUS: a retrospective. *Journal of Usability Studies, 8*(2).

Cai, R. A., Beste, D., Chaplin, H., Varakliotis, S., Suffield, L., Josephs, F., Sen, D., Wedderburn, L. R., Ioannou, Y., Hailes, S., & Eleftheriou, D. (2017). Developing and evaluating JIApp: Acceptability and usability of a smartphone app system to improve self-management in young people with juvenile idiopathic arthritis. *JMIR mHealth and uHealth, 5*(8). https://doi.org/10.2196/mhealth.7229

Chong, W. W., Aslani, P., & Chen, T. F. (2013). Shared decision-making and interprofessional collaboration in Mental Healthcare: A qualitative study exploring perceptions of barriers and facilitators. *Journal of Interprofessional Care, 27*(5), 373–379. https://doi.org/10.3109/13561820.2013.785503

College of Family Physicians of Canada. (2024). CFPC statement on artificial intelligence for family medicine. https://www.cfpc.ca/CFPC/media/Resources/Health-Policy/CFPC-Statement-on-AI-for-FM-FINAL.pdf

Commonwealth Fund Survey, 2019. CIHI. (2020). https://www.cihi.ca/en/commonwealth-fund-survey-2019

Darcel, K., Upshaw, T., Craig-Neil, A., Macklin, J., Steele Gray, C., Chan, T. C., Gibson, J., & Pinto, A. D. (2023). Implementing Artificial Intelligence in canadian primary care: Barriers and strategies identified through a national deliberative dialogue. *PLOS ONE, 18*(2). https://doi.org/10.1371/journal.pone.0281733

Davis, F. D. (1993). User acceptance of information technology: System characteristics, user perceptions and behavioral impacts. *International Journal of Man-Machine Studies, 38*(3), 475–487. https://doi.org/10.1006/imms.1993.1022

Davis, F. D., & Granić, A. (2024). *The Technology Acceptance Model: 30 years of Tam*. Springer International Publishing AG.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—an ethical framework for a good AI Society: Opportunities, Risks, principles, and recommendations. *Minds and Machines*, *28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Ford, E. W., Menachemi, N., Huerta, T. R., & Yu, F. (2010). Hospital it adoption strategies associated with implementation success: Implications for achieving meaningful use. *Journal of Healthcare Management*, *55*(3), 175–189. https://doi.org/10.1097/00115514-201005000-00007

Francisco, K. M., & Burns, C. M. (2024). An approach to potentially increasing adoption of an artificial intelligence–enabled electronic medical record encounter in Canadian primary care: Protocol for a user-centered design. *JMIR Research Protocols*, *13*. https://doi.org/10.2196/54365

Garies, S., Liang, S., Weyman, K., Durant, S., Ramji, N., Alhaj, M., & Pinto, A. (2024a). Artificial Intelligence in primary care practice. *Canadian Family Physician*, *70*(7–8). https://doi.org/10.46747/cfp.700708e102

Goh, E., Gallo, R. J., Strong, E., Weng, Y., Kerman, H., Freed, J. A., Cool, J. A., Kanjee, Z., Lane, K. P., Parsons, A. S., Ahuja, N., Horvitz, E., Yang, D., Milstein, A., Olson, A. P., Hom, J., Chen, J. H., & Rodman, A. (2025). GPT-4 assistance for improvement of physician performance on patient care tasks: A randomized controlled trial. *Nature Medicine*. https://doi.org/10.1038/s41591-024-03456-y

Goh, E., Gallo, R., Hom, J., Strong, E., Weng, Y., Kerman, H., Cool, J. A., Kanjee, Z., Parsons, A. S., Ahuja, N., Horvitz, E., Yang, D., Milstein, A., Olson, A. P., Rodman, A., & Chen, J. H. (2024). Large language model influence on diagnostic reasoning. *JAMA Network Open*, *7*(10). https://doi.org/10.1001/jamanetworkopen.2024.40969

Gomez-Cabello, C. A., Borna, S., Pressman, S., Haider, S. A., Haider, C. R., & Forte, A. J. (2024). Artificial-intelligence-based clinical decision support systems in primary care: A scoping review of current clinical implementations. *European Journal of Investigation in Health, Psychology and Education*, *14*(3), 685–698. https://doi.org/10.3390/ejihpe14030045

Green, M. (2024). Could artificial intelligence improve patient care and physician workload? *Canadian Family Physicians of Canada*, *70*(3), 216–216. https://doi.org/10.46747/cfp.7003216

Guo, Y., Hao, Z., Zhao, S., Gong, J., & Yang, F. (2020). Artificial Intelligence in health care: Bibliometric analysis. *Journal of Medical Internet Research*, *22*(7). https://doi.org/10.2196/18228

Hassan, M., Kushniruk, A., & Borycki, E. (2024). Barriers to and facilitators of Artificial Intelligence Adoption in health care: Scoping review. *JMIR Human Factors*, *11*. https://doi.org/10.2196/48633

Hassan, N., Slight, R. D., Bimpong, K., Weiand, D., Vellinga, A., Morgan, G., & Slight, S. P. (2021). Clinicians' and patients' perceptions of the use of artificial intelligence decision aids to inform shared decision making: A systematic review. *The Lancet*, *398*. https://doi.org/10.1016/s0140-6736(21)02623-4

Hemmer, P., Schemmer, M., Riefle, L., Rosellen, N., Vössing, M., & Kühl, N. (2022, April 19). *Factors that influence the adoption of human-ai collaboration in clinical decision-making*. arXiv.org. https://arxiv.org/abs/2204.09082

Henselmans, I., Van Laarhoven, H. W., Van der Vloodt, J., De Haes, H. C., & Smets, E. M. (2016). Shared decision making about palliative chemotherapy: A qualitative observation of talk about patients' preferences. *Palliative Medicine, 31*(7), 625–633. https://doi.org/10.1177/0269216316676010

Holden, R. J., & Karsh, B.-T. (2010). The Technology Acceptance Model: Its Past and its future in Health Care. *Journal of Biomedical Informatics*, *43*(1), 159–172. https://doi.org/10.1016/j.jbi.2009.07.002

Hollangel, E., & Bye, A. (2000). Principles for modelling function allocation. *International Journal of Human-Computer Studies*, *52*(2), 253–265. https://doi.org/10.1006/ijhc.1999.0288

Holtzblatt, K., & Beyer, H. (2017). *Contextual design: Design for life*. Elsevier.

Hughes, T. M., Merath, K., Chen, Q., Sun, S., Palmer, E., Idrees, J. J., Okunrintemi, V., Squires, M., Beal, E. W., & Pawlik, T. M. (2018). Association of Shared Decision-making on patient-reported health outcomes and healthcare utilization. *The American Journal of Surgery*, *216*(1), 7–12. https://doi.org/10.1016/j.amjsurg.2018.01.011

Jacob, C., Brasier, N., Laurenzi, E., Heuss, S., Mougiakakou, S.-G., Cöltekin, A., & Peter, M. K. (2025). AI for impacts framework for evaluating the long-term real-world impacts of AI-powered clinician tools: Systematic review and narrative synthesis. *Journal of Medical Internet Research, 27*. https://doi.org/10.2196/67485

Jais, C., Hignett, S., Estupiñan, Z. G., & Hogervorst, E. (2018). Evidence based dementia personas: Human factors design for people living with dementia. *Ergonomics For People With Disabilities*, 215–226. https://doi.org/10.2478/9783110617832-018

Johnson, — Tracy. (2023, June 8). *International survey results reveal challenges experienced by family doctors*. CIHI. https://www.cihi.ca/en/news/international-survey-results-reveal-challenges-experienced-by-family-doctors

Jones, M., Koziel, C., Larsen, D., Berry, P., & Kubatka-Willms, E. (2017). Progress in the enhanced use of electronic medical records: Data from the Ontario Experience. *JMIR Medical Informatics, 5*(1). https://doi.org/10.2196/medinform.6928

Jongsma, K. R., Bekker, M. N., Haitjema, S., & Bredenoord, A. L. (2021). How digital health affects the patient-physician relationship: An empirical-ethics study into the perspectives and experiences in obstetric care. *Pregnancy Hypertension, 25*, 81–86. https://doi.org/10.1016/j.preghy.2021.05.017

Jundi, A. A. (2016). Protocol writing in clinical research. *JOURNAL OF CLINICAL AND DIAGNOSTIC RESEARCH*. https://doi.org/10.7860/jcdr/2016/21426.8865

King's Printer for Ontario, 2024. (2024, January 25). *Primary care payment models in Ontario*. ontario.ca. https://www.ontario.ca/page/primary-care-payment-models-ontario

Kohn, S. C., de Visser, E. J., Wiese, E., Lee, Y.-C., & Shaw, T. H. (2021). Measurement of trust in automation: A narrative review and reference guide. *Frontiers in Psychology, 12*. https://doi.org/10.3389/fpsyg.2021.604977

Kueper, J. K., Emu, M., Banbury, M., Bjerre, L. M., Choudhury, S., Green, M., Pimlott, N., Slade, S., Tsuei, S. H., & Sisler, J. (2024). Artificial Intelligence for Family Medicine Research in Canada: Current State and future directions. *Canadian Family Physician*, *70*(3), 161–168. https://doi.org/10.46747/cfp.7003161

Kueper, J. K., Terry, A. L., Zwarenstein, M., & Lizotte, D. J. (2020). Artificial Intelligence and Primary Care Research: A scoping review. *The Annals of Family Medicine*, *18*(3), 250–258. https://doi.org/10.1370/afm.2518

Liu, T.-L., Hetherington, T. C., Dharod, A., Carroll, T., Bundy, R., Nguyen, H., Bundy, H. E., Isreal, M., McWilliams, A., & Cleveland, J. A. (2024). Does AI-powered clinical documentation enhance clinician efficiency? A longitudinal study. *NEJM AI*, *1*(12). https://doi.org/10.1056/aioa2400659

London, A. J. (2019). Artificial Intelligence and black-box medical decisions: *accuracy versus explainability*. *Hastings Center Report*, *49*(1), 15–21. https://doi.org/10.1002/hast.973

Lowdermilk, T. (2013). *User-centered design: A developer's guide to building user-friendly applications*. O'Reilly Media, Inc.

Macfield, R. (2021, September 22). *How to specify the participant group size for Usability Studies: A practitioner's guide - jux*. JUX - The Journal of User Experience. https://uxpajournal.org/how-to-specify-the-participant-group-size-for-usability-studies-a-practitioners-guide/

McAlearney, A. S., Schweikhart, S. B., & Medow, M. A. (2004). Doctors' experience with handheld computers in Clinical Practice: Qualitative Study. *BMJ*, *328*(7449), 1162. https://doi.org/10.1136/bmj.328.7449.1162

Naik, N., Hameed, B. M., Shetty, D. K., Swain, D., Shah, M., Paul, R., Aggarwal, K., Ibrahim, S., Patil, V., Smriti, K., Shetty, S., Rai, B. P., Chlosta, P., & Somani, B. K. (2022). Legal and ethical consideration in artificial intelligence in Healthcare: Who takes responsibility? *Frontiers in Surgery*, *9*. https://doi.org/10.3389/fsurg.2022.862322

Nash, D. M., Thorpe, C., Brown, J. B., Kueper, J. K., Rayner, J., Lizotte, D. J., Terry, A. L., & Zwarenstein, M. (2023). Perceptions of artificial intelligence use in primary care: A qualitative study with providers and staff of Ontario Community Health Centres. *The Journal of the American Board of Family Medicine*, *36*(2), 221–228. https://doi.org/10.3122/jabfm.2022.220177r2

OntarioMD. (n.d.). *What is an EMR?* . EMR certification program FAQ. https://www.ontariomd.ca/emr-certification/faq#:~:text=An%20EMR%20is%20a%20secure,visits%2C%20prescriptions%2C%20and%20more.

Raymond, L., Paré, G., Ortiz de Guinea, A., Poba-Nzaou, P., Trudel, M.-C., Marsan, J., & Micheneau, T. (2015). Improving performance in medical practices through the extended use of Electronic Medical Record Systems: A survey of Canadian family physicians. *BMC Medical Informatics and Decision Making*, *15*(1). https://doi.org/10.1186/s12911-015-0152-8

Razai, M. S., Al-bedaery, R., Bowen, L., Yahia, R., Chandrasekaran, L., & Oakeshott, P. (2024a). Implementation Challenges of Artificial Intelligence (AI) in primary care: Perspectives of general practitioners in London UK. *PLOS ONE*, *19*(11). https://doi.org/10.1371/journal.pone.0314196

Razai, M. S., Al-bedaery, R., Bowen, L., Yahia, R., Chandrasekaran, L., & Oakeshott, P. (2024b). Implementation Challenges of Artificial Intelligence (AI) in primary care: Perspectives of general practitioners in London UK. *PLOS ONE*, *19*(11). https://doi.org/10.1371/journal.pone.0314196

Robin, J., Xu, M., Kaufman, L. D., & Simpson, W. (2021). Using digital speech assessments to detect early signs of cognitive impairment. *Frontiers in Digital Health*, *3*. https://doi.org/10.3389/fdgth.2021.749758

Rout, C. C., & Aldous, C. (2016). How to write a research protocol. *Southern African Journal of Anaesthesia and Analgesia*, *22*(4), 101–107. https://doi.org/10.1080/22201181.2016.1216664

Saldaña, J. (2021). *The coding manual for qualitative researchers*. SAGE Publications Ltd.

Schoorman, F. D., Mayer, R. C., & Davis, J. H. (2007). An integrative model of Organizational Trust: Past, Present, and future. *Academy of Management Review*, *32*(2), 344–354. https://doi.org/10.5465/amr.2007.24348410

Scott, I. A., van der Vegt, A., Lane, P., McPhail, S., & Magrabi, F. (2024). Achieving large-scale clinician adoption of AI-Enabled Decision Support. *BMJ Health & Care Informatics*, *31*(1). https://doi.org/10.1136/bmjhci-2023-100971

Senthil, R., Anand, T., Somala, C. S., & Saravanan, K. M. (2024). Bibliometric analysis of Artificial Intelligence in healthcare research: Trends and future directions. *Future Healthcare Journal*, *11*(3), 100182. https://doi.org/10.1016/j.fhj.2024.100182

Shanafelt, T. D., Dyrbye, L. N., Sinsky, C., Hasan, O., Satele, D., Sloan, J., & West, C. P. (2016). Relationship between clerical burden and characteristics of the electronic environment with physician burnout and professional satisfaction. *Mayo Clinic Proceedings*, *91*(7), 836–848. https://doi.org/10.1016/j.mayocp.2016.05.007

Sivaraman, V., Bukowski, L. A., Levin, J., Kahn, J. M., & Perer, A. (2023). Ignore, trust, or negotiate: Understanding clinician acceptance of AI-based treatment recommendations in health care. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–18. https://doi.org/10.1145/3544548.3581075

Terry, A. L., Kueper, J. K., Beleno, R., Brown, J. B., Cejic, S., Dang, J., Leger, D., McKay, S., Meredith, L., Pinto, A. D., Ryan, B. L., Stewart, M., Zwarenstein, M., & Lizotte, D. J. (2022). Is primary health care ready for Artificial Intelligence? what do primary health care stakeholders say? *BMC Medical Informatics and Decision Making*, *22*(1). https://doi.org/10.1186/s12911-022-01984-6

Tierney, W. M., Henning, J. M., Altillo, B. S., Rosenthal, M., Nordquist, E., Copelin, K., Li, J., Enriquez, C., Lange, J., Larson, D., & Burgermaster, M. (2022). User-centered design of a clinical tool for shared decision-making about diet in primary care. *Journal of General Internal Medicine*. https://doi.org/10.1007/s11606-022-07804-x

Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated Criteria for Reporting Qualitative Research (COREQ): A 32-item checklist for interviews and Focus Groups. *International Journal for Quality in Health Care*, *19*(6), 349–357. https://doi.org/10.1093/intqhc/mzm042

Tsai, C.-C., Liu, C.-F., Lin, H.-J., Lin, T.-C., Kuo, K.-M., Lin, J.-J., Chen, C.-J., & Lee, M.-C. (2022). Implementation of a patient-centered mobile shared decision making platform and Healthcare Workers' evaluation: A case in a medical center. *Informatics for Health and Social Care*, *48*(1), 68–79. https://doi.org/10.1080/17538157.2022.2054344

Turpin, R. (2022, November 21). *Machine learning AI in medical devices: Adapting regulatory frameworks and standards to ensure safety and performance*. Machine learning AI in medical devices. https://aistandardshub.org/research/machine-learning-ai-in-medical-devices-adapting-regulatory-frameworks-and-standards-to-ensure-safety-and-performance/

Ueno, T., Sawa, Y., Kim, Y., Urakami, J., Oura, H., & Seaborn, K. (2022). Trust in human-ai interaction: Scoping out models, measures, and methods. *CHI Conference on Human Factors in Computing Systems Extended Abstracts*, 1–7. https://doi.org/10.1145/3491101.3519772

Ullah, W., & Ali, Q. (2025). Role of artificial intelligence in healthcare settings: A systematic review. *Journal of Medical Artificial Intelligence*, *8*, 24–24. https://doi.org/10.21037/jmai-24-294

Van der Weijden, T., Pieterse, A. H., Koelewijn-van Loon, M. S., Knaapen, L., Légaré, F., Boivin, A., Burgers, J. S., Stiggelbout, A. M., Faber, M., & Elwyn, G. (2013). How can clinical practice guidelines be adapted to facilitate shared decision making? A qualitative key-informant study. *BMJ Quality & Safety*, *22*(10), 855–863. https://doi.org/10.1136/bmjqs-2012-001502 eper

Vereenooghe, L., Trussat, F., & Klose, K. (2020). *Applying the Technology Acceptance Model to Digital Mental Health Interventions: A Qualitative Exploration with Adults with Intellectual Disabilities*. https://doi.org/10.31234/osf.io/vtbr7

Wesche, J. S., & Sonderegger, A. (2019). When computers take the lead: The Automation of Leadership. *Computers in Human Behavior*, *101*, 197–209. https://doi.org/10.1016/j.chb.2019.07.027

Whaley, C., Bancsi, A., Ho, J. M.-W., Burns, C. M., & Grindrod, K. (2020). How do I keep myself safe? patient perspectives on including reason for use information on prescriptions and medication labels: A qualitative thematic analysis. *Journal of Pharmaceutical Policy and Practice*, *13*(1). https://doi.org/10.1186/s40545-020-00268-6

Wiens, J., Saria, S., Sendak, M., Ghassemi, M., Liu, V. X., Doshi-Velez, F., Jung, K., Heller, K., Kale, D., Saeed, M., Ossorio, P. N., Thadaney-Israni, S., & Goldenberg, A. (2019). Do no harm: A roadmap for responsible machine learning for health care. *Nature Medicine*, *25*(9), 1337–1340. https://doi.org/10.1038/s41591-019-0548-6

Winter, J. C. F., & Hancock, P. A. (2015). Reflections on the 1951 Fitts list: Do humans believe now that machines surpass them? *Procedia Manufacturing*, *3*, 5334–5341. https://doi.org/10.1016/j.promfg.2015.07.641

Yarbrough, A. K., & Smith, T. B. (2007). Technology acceptance among physicians. *Medical Care Research and Review*, *64*(6), 650–672. https://doi.org/10.1177/1077558707305942

Zhu, M., & Zhang, Y. (2021). Medical and Public Health Instructors' perceptions of online teaching: A qualitative study using the Technology Acceptance Model 2. *Education and Information Technologies*, *27*(2), 2385–2405. https://doi.org/10.1007/s10639-021-10681-2

Appendices

Appendix A Recruitment Materials
E-mail to Primary Care Physicians

Hello *insert name,*

I am in the process of my PhD in Systems Design Engineering at the University of Waterloo under the supervision of Dr. Catherine Burns. The specific research topic I am exploring is the potential impact of an AI-enabled electronic medical record (EMR) for primary care physicians in Ontario. As a part of my research, I am looking to conduct a two-part interview process that will take up to approximately 90 minutes total to complete. The first part will focus on presenting you with common scenarios seen in primary care where I can establish your typical workflow and decision-making process. The second part will focus on the same scenarios but with the introduction of an AI-enabled EMR. No patients will be required for the purposes of this study. The objective of this interview is to elicit feedback on the design of the AI-enabled EMR, and how the interface can be optimized to suit your needs.

Please see attached for one-pager describing some of the current research available in Canadian primary care and artificial intelligence.

This study has been reviewed and received ethics clearance through the University of Waterloo Research Ethics Board. I would require approximately 90 minutes of your time to discuss. A $75 Visa gift card will be provided in exchange for your valuable time and input. If you are able to participate, please respond to this e-mail and I will arrange a time with you.

Thank you for your consideration!

Social Media Post

Hello ***linkedin/twitter*** network, I need your help! My PhD research is focused on how the introduction of an AI-enabled "encounter module" in an electronic medical record can impact decision making for primary care physicians practicing in Ontario.

I have partnered with TELUS health, a key trailblazer in EMRs and digital health in Canada in the last decade. With expertise in primary care and physician engagement, they are an ideal partner to collaborate with on this innovative research to gain a full understanding of the potential of AI in the primary care domain.

This valuable research will inform how I can optimize the redesign of tools like this to improve the patient and provider experience in primary care. A $75 Visa gift card will be provided in exchange for your valuable time and input.

If you are or know of any interested people who are a practicing family physician in Ontario, Canada I would love to speak with you!

Appendix B Pre-Read Material

TELUS HEALTH AI-ENABLED ELECTRONIC MEDICAL RECORD RESEARCH
PhD Student: Krizia Francisco
Supervisor: Dr. Catherine Burns

Thank you for wanting to understand this research further. Your time and input is incredibly valuable and if you choose to participate, will help to inform important work on how electronic medical records can be equipped with AI-enabled functionality to improve the provider experience.

Background:

Artificial Intelligence (AI) arguably has endless possibilities for application, which is what makes it both incredibly exciting and difficult to implement. With varying expectations and levels of comprehension, it is important to identify various stakeholders in the healthcare system and understand their expectations and intentions regarding AI in healthcare. The area of focus for this research is primary care; primary care physicians have always acted as a "quarterback" in the patient's care and have played an important role in ensuring a patient's continuity of care in the Canadian healthcare system. The generalist approach and lack of specialization to a single system of the body makes this a more difficult sector to embed AI into, in contrast to specific specialties like oncology, or diagnostic imaging. These areas of specialty allow data to be more specific, which makes the exploration of the use of AI more narrowly scoped than if it were primary healthcare. There has yet to be a widespread exploration of the use of AI in primary healthcare and how patients and physicians envision AI use to encourage further uptake.

The data available in healthcare are considered sensitive and highly personal. The way AI presents information to a physician and patient in interaction has the potential to impact behaviors, including how decision-making is made. Understanding human acceptance and the potential impact of a clinical encounter with the introduction of AI is a key milestone in

enabling the industry to effectively predict the potential successes and barriers to AI-enabled products and services.

In a clinical interaction with a primary care provider, it's remained important to be inclusive and driven by an individual's values, and decisions. This can be facilitated through shared decision-making (SDM) which can be defined as "a process that respects the rights of patients to be fully involved in decisions about their care" (Hughes et al., 2018). SDM raises awareness in the clinician on how to best support the individual that stays consistent with the individual's requests.

The introduction of new technology in the patient and primary care interaction is not a new endeavor; in the past, having to integrate the use of Electronic Medical Records (EMRs) caused concern for the patient experience. A key example was the fear that, with the introduction of EMRs, patients may feel their physicians were too involved with their screens instead of listening to patient concerns. The concern for maintaining the human relationship in the patient and physician interaction speaks to one of the more important sides of healthcare, which is fostering and maintaining compassionate medicine and care. Despite concerns and challenges, EMRs have now become a standard part of care, and are further supported through professional colleges and regulatory policies. As EMRs have become a more integrated part of primary care delivery, it has also increasingly been recognized as a cause of physician burnout (Shanafelt et al., 2016).

Studies have demonstrated that physicians who are EMR users have actually reported being less satisfied with the amount of time allocated on tasks associated with their EMRs. This clearly demonstrates that EMRs have potential to be optimized as a tool that can support care delivery. The potential benefits of having searchable, digitally stored, and secured records that contribute to larger population health analytics are still significant factors of healthcare transformation driven by technology. With these learnings, it's important to highlight how we can optimize and design of AI in the primary care setting to ensure it's ultimately

supporting primary care delivery. Mirroring the example of EMR integration in a clinician's workflow, a similar discovery of benefits outweighing the potential perceived risks can apply to tools enabled with AI in primary care.

TELUS Health's Collaborative Health Record (CHR)

The TELUS Health CHR will be the EMR used for the purposes of this study. The primary module within the CHR that will be used for these interviews is the "Encounter". The Encounter module is where primary care clinicians spend majority of their time documenting, even during a clinical visit. It captures information on patient responses to questionnaires, health past history, examinations, assessment & planning, prescriptions, attachments, referrals, injections, follow-up items, and billing. Encounters require comprehensive information capture making it an ideal module to focus on for added artificial intelligence and it's potential to impact how a physician captures a clinical encounter.

What to expect during the interview:

A two-part interview process that will take up to approximately 90 minutes total to complete. The first part will focus on presenting you with a common scenario in primary care, so I can establish a workflow of how you interact and engage with your existing EMR. The second part will focus on same scenarios but with the introduction of an encounter module within the EMR built with artificial intelligence.

A potential subsequent interview of 30 minutes may be requested to validate findings. These interviews can take place either in-person, or video chat (e.g. Microsoft Teams) based on your preference.

Appendix C Oral Consent Script
Introduction:
Hello. Thank you for agreeing to participate in this audio call today. I'm *Krizia Francisco,* I am conducting interviews about Artificial Intelligence in the context of Canadian Healthcare. I'm conducting this as part of research for my PhD at University of Waterloo's Systems Design Engineering department in Waterloo, Ontario under the supervision of Dr. Catherine Burns.

Study procedures:
Physician Participants: I'm inviting you to do a one-on-one interview *on Microsoft Teams* that will take about 90 minutes. In the first 30 minutes I will present a common scenario in primary care and invite you to describe to me the process you typically go through based on that scenario in the TELUS CHR Encounter module. In the second part of the interview, I will present you with an AI-enabled version of the CHR Encounter module (pre-read material has been provided) and ask you various open-ended questions on the tool's design and its impact on the original scenarios.

With your permission, the interview will be audio recorded to facilitate collection of information, and later transcribed for analysis. Shortly after the interview has been completed, I will send you a copy of the transcript to give you an opportunity to confirm the accuracy of our conversation and to add or clarify any points that you wish.

Your identity will be confidential. Your name will not appear in any thesis or report resulting from this study, however, with your permission anonymous quotations may be used. Data collected during this study will be retained for at least 2 years to which only researchers associated with this project will have access. There are no known or anticipated risks to you as a participant in this study.

Risks:
The risks involved in participating in this study are minimal, and you might find some questions uncomfortable to answer. You do not need to answer questions that make you feel uncomfortable or that you do not want to answer. And you can withdraw (stop taking part) at any time

I describe below the steps I am taking to protect your privacy.
I will keep the information you tell me during the interview confidential. The use of quotations in any paper or publication resulting from this study will be anonymous. Any data from this research which will be shared or published will be the combined data of all participants. That means it will be reported for the whole group not for individual persons.

Voluntary participation:

- Your participation in this study is voluntary.
- You can decide to stop at any time, even part-way through the interview for whatever reason
- If you decide to stop participating, there will be no consequences to you.
- If you decide to stop we will ask you how you would like us to handle the data collected up to that point.
- This could include returning it to you, destroying it or using the data collected up to that point.

Once I've submitted any information related to this study for publication, you cannot withdraw then.

This study has been reviewed and received ethics clearance through a University of Waterloo Research Ethics Board (REB #43922). If you have questions for the Board, contact the Office of Research Ethics, toll-free at 1-833-643-2379 (Canada and USA), 1-519-888-4440, or reb@uwaterloo.ca.

I would be pleased to send you a short summary of the study results when I finish going over our results. Please let me know if you would like a summary and what would be the best way to get this to you.

Consent questions:

By providing your consent, you are not waiving your legal rights or releasing the investigator(s) or involved institution(s) from their legal and professional responsibilities.

 Do you have any questions or would like any additional details? *[Answer questions.]*

 Do you agree to participate in this study knowing that you can withdraw at any point with no consequences to you?

 Do you agree to the use of anonymous quotations in any paper or publication resulting from this research ?

 Do you agree to have your interview audio recorded to ensure an accurate recording of your responses?

[If yes, begin the interview.]

[If no, thank the participant for his/her time.

Appendix D Thematic Analysis

This is the thematic analysis table summary that includes the codes.

Theme	Sub-Themes	Codes
Trust & Transparency	Clinicians' trust is linked to high-performance expectations of AI.	• AI needs to be more advanced • Concern of physician image • Concern of physician replacement • Concern of patient safety • Based on biggest model possible • If I disagree with it, I'd turn summaries off. • Something like GPT is helpful • I'm not fearful • Needs to be a little bit more intelligent • No way to avoid AI • Performing worse than my medical student • Personalization over Time
	AI can build trust with clinicians by reinforcing their authority by providing relevant, nonintrusive support.	• Don't believe it, my weird human heuristics • Nervous • Potentially helpful potentially harmful • Pretty much what I would be doing • Likes the AI to be interactive • A flexible system is helpful

		• Doesn’t like the AI to generate repetitive or already existing information • AI introduces bias
	Transparency in how confidence scores were determined, and training data could improve trust.	• Black box problem • Don't trust big vendors • Specific towards me and how my practice • Where does the data go from the AI • Convey what the score means and how the score was calculated • Explanation is needed for numeric factors
	Clinicians felt uncertain about generic button functions, which led to hesitation.	• Button should always be edit • Don't know what I'm approving • What happens if I press it (4) • It’s not clear what happens if I press these buttons (5) • Doesn’t want decline and approve options • Approve and edit is time consuming
	AI can be helpful for supportive suggestions like reminders,	• AI can help reverse disease and improve health trajectories

	preventative care guidelines, or flagging allergies.	• AI needs wellness and wellbeing • Check for pregnancy reminder
	Suggestions should align with realistic, helpful actions that are evidence-based and practical.	• Comprehensive suggestions • Enable continuity of care • I like pathways • Needs realistic recommendations
Clinician Authority & Decision Making	AI should offer useful suggestions without appearing to make autonomous decisions. AI should not override clinician decision making as they always have final say.	• Devil's advocate AI • Shouldn't tell me what to do • Hard and complicated decision making • Lead to harm • OK fine, that's what it's telling me to do • Ultimate decision maker
	Clinicians should retain decision-making authority, AI's role should be to support, not dictate, clinical decisions.	• AI leads to less thinking for clinicians • Assistive like a calculator • GAD is not diagnostic • Use your own clinical judgment
Complexity Sensitivity	Context-sensitive suggestions: AI-suggested medications are acceptable for simple cases, while complex cases require heavier clinician discretion.	• Clinical decision-making appropriateness • Go further • Mental health is, whole different story • Misunderstand what MH means • UTI treatment is much more algorithmically treatable

	Contextual information provided by clinicians in prior interactions is important. Clinicians want AI that understands past interactions and can capture nuanced contexts.	• AI not good at capturing nuance • Love AI to suggest what I didn't think about • Nuances should be accounted for • Referral should be based on patient info
	Summaries need more comprehensive patient histories, with previous scores and relevant screening tools displayed.	• Make sections editable • Massive leap between score and assessment • Notes are a bit too simple • Patient history needs more
	Mixed preferences on medication suggestions, depending on the complexity of the case.	• Learn what I prescribe • No value-add if I'm not going to do it • Overly Pre-emptive • Suggest a medication for UTI
Workflow Efficiency and Integration	Clinicians desire actionable, patient-centered summaries with appropriate depth and focus.	• Growing fast it's coming • I want things to the point • Keep it simple • No one's gonna use it if there's too much stuff • What's actionable
	Clinicians expect prompts for preventative measures (e.g. mental health risk factors).	• AI needs wellness and wellbeing • Check for pregnancy reminder • Preventative care needs to be up to date • I want to make sure I talk about prev care

		• If patients are here for a UTI I would bring up all the other prev care items
	Clinicians are cautious of features that add workload or disrupt their workflow.	• Convenient • Decline button adds extra click • Don't need quick settings • Radical change of practice
	AI needs to fit smoothly into the existing workflow without adding steps.	• Consolidate what's relevant • Do not want to be providing prompts. • Don't give doc more things to do • How do we incorporate it • Make my life easier • More clicks • More work but faster over time • Think in the SOAP format • The approve or decline part in UI/UX is making clinicians more job to do
	Clinicians primarily saw the confidence score as a potential distraction but were curious about its utility, which may serve well in selective use.	• I care about the confidence score grade • Suggesting three level scaling for the confidence factor • Only likes probable suggestions with high or moderate confidence score • Confidence score must be interpretable • Optional on/off feature • Maybe confidence score is not needed

		• Exciting • Feed more info to increase confidence score • Importance of confidence score
Legal and Ethical Considerations	Clinicians require a visual representation that distinguishes what they say vs. what the AI generates before a note is finalized.	• I like color (3) • Need a distinction between clinician and AI generated (9) • Liability
	Concerns on how legal frameworks are going to support the introduction of AI in clinical practice.	• Clean as possible privacy and security • Nothing should be deleted • Indicate patient refused • Medical legal record • Risk appetite • What if my disagreement with AI is wrong • Consent aspects must be made clear • Unsure what safeguards are in place to protect me a physician if there are any negative impacts to patients/unintended consequences
	Ultimately viewed AI as a tool that does not require direct disclosure to patients of its use.	• Don't need explicit consent • Explain to patient • I don't tell patients • It's up to me to determine what goes in there, so it's either accepting

		something that a computer generated or doing it myself, but at the end of the day, I think I'm responsible for what's in there.
Human-Centered Design	AI must maintain a human element in patient interactions by not distracting clinicians and performing reliably.	• I would look foolish • Will people feel cared for • Wonderful it did what I did • Easy to understand
	AI should support human-centered care without creating dependence on AI.	• Is this going to replace me • Outcome driven • Worry it will further entrench bad patterns • Clinical cognitive activity considered

Appendix E

Detailed redesign requirement reflections generated from thematic analysis, ensuring alignment between qualitative findings and practical system modifications.

Ease of Use	**Feedback**
Thematic comment: It looks like the wants are related to the AI in a state of high-performance but doesn't sacrifice Personalizable aspects of work and not only decreases the workload, but definitely does not add work to the clinicians. Fear of increased workload drives them more than the potential benefit of decreased workload. All relates to how they will "accept" the technology	Wants: • AI to be trained • AI to be more intelligent • AI to learn over time • AI to make novel suggestions • Customizability • Decrease in workload • Easy to understand • Personalizable • Ability to modify notes • Searchable data that makes it easier to use Does Not Want: • To add workload on the clinician (4x) • Buttons add workload • Waste a user's time by including AI • Concerns: • Lack of human-to-human interaction
AI-Generated Summary Thematic comment: Clinicians stressed the need for the AI-generated summary to be succinct and readable via color coding. They also stressed the need for the CPP and the desire for flexibility in the overall use. Some suggestions included integrating labels or stamps, creating a library of common outputs, and pre-loading optionality and responses.	General desire of how it's presented • It to be succinct (3) • Color coding to match various functions (6) • CPP (4) • Empty state to not be overwhelming • Familiarity of how they already chart

The order and flow of how they read through the chart mattered. The suggested outcome should not be at the top but at the bottom. From a liability perspective, there were suggestions that if they disagreed with AI they want it to disappear. Many participants focused on the "suggestions" themselves (suggested outcome) • Design requirements include legislative requirements for charting (OMA, CPSO)	• Flexibility of overall use • Integration of labels or stamps • Library of common outputs • Pre-loaded optionality of responses • Having patient history accessible Formatting/Edits • Does not like the referral UI • Likes how it's visually sorted Order/Flow • Place suggestions at the bottom • Problem with the placement of suggested outcome • If there is a disagreement with the AI, AI suggestion disappears Suggestions • The suggestion did not have enough to it • Wants the AI to suggest next steps • Wants to be able to edit the suggestion • Wants them to be novel
Buttons The approve, edit, and decline buttons were not well received. The majority stated they didn't want more	• Adds mental load • Don't want more clicks • Time-consuming • Fewer clicks (3)

clicks, that it was time-consuming, and that it could potentially add mental load to the clinician. Some options they said were to use agree over approve. The buttons were unclear about what happened after being pressed, which caused clinicians to pause and express concern that they did not like that it was not obvious what the following prompt was going to be. **Notes:** Is it a learning curve that clinicians have to overcome to adjust to the use of buttons? Many of the comments point in that direction. • Some of the buttons felt unnecessary (e.g., it already exists in context)	• Approve and decline don't add value • Prefers agree over approve • Fewer clicks (4) • Unclear what the next steps are • Feeling uneasy about pressing a button
Confidence Score The majority of participants wanted the confidence score to be optional and displayed differently (whether it's distilled to 3-level scaling, probably suggestions like low, medium, or high, and making it easily interpretable). The confidence score is not a familiar term to them, which is why other participants said that they did not like it at all. Along with flexibility, 2 participants expressed the need for the ability to audit how the confidence score arrived at its current level. **Notes:** • Potentially, re-naming the concept might help • Confidence is also related to liability	• 1 Said confidence score was confusing • 3 Explicitly stated they did not like the confidence score • - 1 They explicitly stated they did like the confidence score • 7 Wanted the confidence score to be optional and presented differently if it had to be included • 4 explicitly stated they wanted the confidence score to be optional • 1 Wants to feed information to increase the confidence score • 2 Showed direct concern over being able to audit

• If we are transparent with how their LLM was trained, they may be less skeptical of where it came from	how the confidence score was determined
Decision Making	Wants clear delineation while editing clinician vs AI
General Usability Expectation	There will be more work involved in this at the beginning, but over time, we will learn it, and it will become a part of care. The risk from a time perspective primarily exists at the beginning of the learning curve.
Medications Will need to explore re-labelling in next design There was concern with the vendor pushing specific medications to give unfair advantage to specific pharma companies	• Medications should be viewable while selecting (and seeing previews of prescriptions) • Level of transparency
Other/ UX Preferences Explore if a floating window where clinicians can interact with the AI (providing prompts) could assist	• Automation is good for nuances • User likes to see information in one place • Does not like pop-ups • Viscerally likes the initial screen • Likes colored UX • Customizable Features for User Preferences • Wants the system to be more Personalizable • Limited tools used in interface
Quick Settings	Quick settings was noticed by 3 people – all stating that it is not helpful

Usefulness	Feedback
AI General It would be helpful for clinicians to understand how the AI is trained, for AI to capture nuance and want AI to remember the context clinicians have provided it in the past.	• How AI is trained • multi-turn interaction • Not good at nuance • Proactive assistance • Risk mitigated • Useful for early detection
AI-Generated Summary Clinicians desire a clear call to action, have varying degrees of error tolerance, and still strongly believe that it would be useful for them to understand the LLM's data source(s). The desire is for AI to be human-oriented for patients and not to create heavy dependence on AI for clinician decision-making. (for long-term clinician impact). Does not want AI involved in clinical decision-making (x6) The patient expects AI to provide proactive assistance (e.g., asking about risk factors for MH, context from past mental health history). This theme relates to "call to action" (e.g., sending patients for genuine preventative care like bloodwork) One clinician felt that the AI was overly pre-emptive, highlighting the varying degrees of expectation of "how far" AI goes with a suggestion. Clinicians expect AI to be trained by best practice guidelines (since they are difficult to keep up with). Consider the breadth of practice guidelines, where they come from, and how they're presented.	• Call to action • Error tolerance • Explainability interaction • How AI is trained • How AI is trained • Human oriented • Increased efficiency • Level of interaction/input • Long-term clinician impact • multi-turn interaction • Not good at nuance • Patient Summary • Proactive assistance • Recommendations from best practice • Risk mitigated • Summary structure

What is considered “enough” on diagnostic recommendations Summary structure: Many expressed that the patient history was not comprehensive enough; they also wanted comprehensive suggestions, notes were perceived as “too simple,” and two clinicians suggested that the diagnostic screening tools are not helpful and do not give much information. The previous score in the summary would be helpful.	
Confidence Score The confidence score should be higher for things more algorithmic, like a UTI; for more complex diagnoses, like MH, they don’t “trust” it. Selective confidence scores based on primary concern were suggested. 5 instances where participants viewed the confidence score as a distraction. The confidence score isn’t interpretable in its current state. One suggested hovering over the score to see the source. Otherwise, the majority expressed that it added distraction. If we add a confidence score, it matters where it is placed in the summary structure.	• Adaptable based on primary concern • Distraction • Explainability interaction • Summary structure

Decision Support
 	
Although not explicitly said by **all** participants, there was a general sentiment from all (including the data) that clinicians be the ultimate decision makers in a clinical interaction.
	• Appropriateness
	  ○ AI is suitable for algorithmically treatable diseases P1
	  ○ AI can be more confident with more general

Decision support when appropriate (e.g., pregnancy reminders for antibiotic prescriptions or allergies to medications)	disease outcomes like UTIs, compared to the mental health problems P6 o Complexity of Prescribing Decisions P8 o AI can add value to the diagnostic part P9 o AI tools can help chart P11 o Use of AI for less complicated cases P14 • Call to action • Distraction • Human oriented • Patient Summary • Personalization • Recommendations from best practice
General "Distraction": There's a resistance to accepting AI. The fear of AI adding an increased workload outweighs the perceived potential benefits. Some believe the AI was not adding value. "Error Tolerance" - Low error tolerance - if the AI summary makes bad suggestions they disagree with, it will not gain their trust, and they won't accept it. "How AI is Trained" - General sentiments on how the AI is trained	• Adaptable based on primary concern • Call to action - **Moved to decision support/AI AI-generated summary** • Distraction • Error tolerance • Explainability interaction - **Moved to AI AI-generated summary** • How AI is trained • How AI is AI trained • Human-oriented • Increased efficiency

"Human Oriented" - Clinicians want to maintain a sense of "humanness" in the interaction with patients (e.g., scores dictating an MH diagnosis could be seen as unprofessional). AI could be seen as a barrier between the clinician and the patient. "Increased Efficiency"—Some participants emphasize increased efficiency through the use of transcribing, calculations, charting, or a potential secretary for provider support. "Level of Interaction" - **Moved to Decision Support** "Multi-Turn" - **Moved to AI-Generated Summary** **"**Not good at nuance" - **Moved to AI General** Personalization exists for both aspects of designing for clinicians and patients. Some participants saw it as an opportunity to provide better patient personalization, and others saw a desire to maintain flexibility in their EMR. "Proactive Assistance" - Move to AI-Generated Summary/Decision Support "Recommendations from best practice" and "Risk Mitigated" - Move to Decision Support "Summary Structure" - moved to AI-Generated Summary.	• Level of interaction/input • multi-turn interaction • Not good at nuance • Optionality • Patient Summary • Personalization • Proactive assistance • Recommendations from best practice • Risk mitigated • Summary structure
Medications	• Call to action - Some want medication suggestions for straightforward treatments (e.g., UTI); some did not want medication suggestions for pathways

	with more nuance (e.g., MH) • Distraction - moved to AI General • Explainability interaction - changed to “TRUST” - Clinicians need help building trust with the AI suggesting specific medications (and why) • Patient Summary - Changed to decision support • Personalization - Clinicians want the AI to “learn what they subscribe” • Proactive assistance - Move to Decision Support
Suggested Outcome	• Appropriateness - AI can’t be making definitive decisions in the absence of the clinician • Call to action - Needs to provide realistic suggestions (that are helpful) • Distraction - Does not see suggestion feature as helpful • Error tolerance - Tied to confidence score (if used) don’t want to see low-quality suggestions • How AI is trained - Moved to”AI General”

	• Human oriented - Moved to AI Generated Summary/Decision Support • Long term clinician impact - There's concern on long-term impact of clinician decision making and over-dependence on AI • Proactive assistance - AI should work on pace with a clinician's workflow • Risk mitigated - Fear if skipping suggested outcome in addition to records retention of suggested outcome (should vs should not keep in chart) Some clinicians are interpreting the presence of an AI tool as monitoring/additional audit vs. assistive • Summary structure - Wants it to be comprehensive and presented in a way that they understand/accept (all want the suggested outcome but it needs to be changed how it is presented)
Suggestions	Move to "AI-Generated Summary"
Workflow	• Call to action - that involves the patient; what

	is next for them? (exists for both clinician and patient)
	• Distraction - Participants expressed the importance of not adding to the existing workload
	  ○ Same sentiment as "call to action"
	  ○ Don't want added workload BUT want to be able to audit
	• Explainability interaction - Same sentiment as trust in other areas; over time, the desire for explainability will lessen
	• Human oriented - Moved to AI-generated summary (highlighting comprehensive approaches to care, do not make diagnosis decisions based on quantitative screening tools alone, include information on person's whole health picture - e.g., MH should include convos about diet, exercise, sleep)
	• Increased efficiency - Participants sought AI to lessen their burden, speed things up, and make "their life easier."

	• Long-term clinician impact - Concerns over how clinicians approach the decision-making process long-term in addition to the ability to enable continuity of care (the core piece of primary care clinician's work) • Proactive assistance - Wants AI to make suggestions they do not think of (x4). **Move to "AI-Generated Summary/Decision Making"** • Risk mitigated - Risk mitigation processes that exist today (audit, consent, tracking patient decisions) should go unchanged with the introduction of AI • Summary structure - Aligns with how primary care clinicians are trained and work today

General Sentiments	**Feedback**
AI General	Many general sentiments on explainability, learning curve, etc. may need to be re-sorted
AI Trust Trust is not something clinicians were inherently feeling towards AI, it was seen as something that would be earned over time with the right AI system.	• Explainability of how AI arrives at an outcome (what data is used)

There was an overall tone of fear of the impact that AI could potentially have on them - this could be due to factors related to how policy has not caught up with the shift in technology. Skepticism and concern were voiced from the perspective of both clinicians and patient safety - in addition to the vendors that support this type of development work and potential connections to pharmaceuticals. Some maintained a neutral attitude towards what they were seeing, at times even amusement, likening its performance to that of a medical student or expressing and acknowledging that the introduction of this technology could be “really good, or bad” Others expressed a positive optimism towards AI and its role in the future. In all this, it is clear that clinicians are anticipating unintended consequences, whether positive or negative, with the introduction of AI.	• Skepticism of vendor market • Positive • Concern of physician image • Positive • Concern of physician replacement • Concern for patient safety • Who is ultimately liable for any mistakes the AI makes • Concern of repercussions • High trust in themselves in addition to understanding ownership of liability • Concern of patient safety, does not trust it "enough" for high-stakes decisions • Expectations of performance • Neutral attitude • Does not want to be "told by AI what to do", trust or ego? • Physicians are the clinical decision makers, not AI- ego • Concern of human connection between physician and patient - "will people feel cared for?" • Concern about how much people rely on AI that it

	impacts their decision-making • Data transparency - where does the information go? • Neutral • Receptive to new concepts and new technology • AI response strays from the focus, *clinical decision-making impact • Nervous about repercussions • Lack of trust in AI • Expected performance of AI • Lack of understanding of AI > lack of trust
Audit	• Opinions on record retention and how nothing should be deleted • Need to establish distinct roles between physician and AI
Clinical Context	• Clinicians reference that mental health is a different level of complexity for AI • Clinicians reference that UTIs are straightforward and, therefore appropriate enough for AI
Clinical Decision Making	• Physicians are the clinical decision makers, not AI • Quantifiable scores are only one dimension of

	assessment and decision making • Concern of AI biasing clinical decision making
Clinician Cognitive Activity	• One theme here was a fear of over-dependence on AI that clinicians think less - these comments came from 2 clinicians
Disclosure	• 7 instances where clinicians did not deem it necessary to disclose to patients they are using AI - they see it as synonymous with other tools
Legal	• Participants voice that physicians are the ultimate decision maker and because of there are fear of potentially repercussions • Unsure what safeguards are in place to protect me a a physician if there are any negative impacts to patients/unintended consequences • Some expressed neutral concerns from the legal perspective - doesn't think they need to be overly involved as they believe policy will catch up

Medications	• Optimistic about its ability to support medication management
Privacy, Security, and Consent	• The overarching theme clinicians felt consent should be clear but low touch
Workflow	• Acknowledgement that the introduction of AI needs to be carefully thought about - clinicians are eager but unsure how to incorporate especially with the supporting legal landscape behind

Appendix E: System Usability Scale (SUS)

SUS was presented to clinicians at the end of a validation interview.

AI Enabled EMR Encounter: System Usability Scale(2 Minutes)

Thank you for your valued time and participation in my research. As a final step, please complete the following quick usability scale based on the new mock-ups you reviewed in the interview today. For any questions, please reach out to krizia.mae.francisco@gmail.com

Answers will remain anonymous. This study has been reviewed and received ethics clearance through the University of Waterloo Research Ethics Board (43922).

I think that I would like to use this system frequently *

	1	2	3	4	5	
Strongly Disgree	○	○	○	○	○	Strongly Agree

I found the system unnecessarily complex *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I thought the system was easy to use *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I think that I would need the support of a technical person to be able to use this system *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I think that I would need the support of a technical person to be able to use this system *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I found the various functions in this system were well integrated *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I thought there was too much inconsistency in this system *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I would imagine that most people would learn to use this system very quickly *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I found the system very cumbersome *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

.

I felt very confident using the system *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I needed to learn a lot of things before I could get going with this system *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

Optional: Provide any additional feedback you'd like about today's mock-ups.

Your answer

Appendix F Part A: Tabulated answers from validation interviews in detail.

Participant	AI Assist Button	Color-Coded AI vs Clinician-Entered	Confidence Scores	Customization & Personalization	Volume of Clicks
P6	No additional feedback on the trigger button.	Accepted as presented; the clear delineation (green for manual, blue for AI) is appreciated.	Supports a non-numeric low/medium/high format; values transparency and quality through stratification.	No explicit comment provided.	Not commented on.
P7	Values the option to manually trigger AI assist rather than auto-population.	Finds the color coding useful for distinguishing AI-generated (blue) from clinician-entered (green) data, though he flags potential printing issues.	Prefers a tiered (low/medium/high) system over numeric scores; believes it better communicates certainty.	Advocates for streamlined, customizable fields (e.g. via API) to better align with workflow.	Emphasizes minimal clicks for efficiency (e.g. quick prescription ordering).
P10	No explicit comment regarding the trigger button.	Loves the clear color coding that distinguishes AI-generated (blue) from manual input (green) but	Appreciates the low/medium/high display; notes that low scores prompt extra review, though he finds the	Prefers personalization wanting choices among suggestions rather than one-size-fits-all defaults.	Not explicitly commented on.

		wonders about printed output.	calculation method vague.		
P01	Advocates for a manual prompt button to control when AI input appears—rejects constant background operation.	No explicit comment on color distinctions provided.	Explains that confidence scores should reflect AI certainty (based on similar cases) and prefers that only high-confidence suggestions appear; finds the calculation method vague.	Implies decision support should integrate seamlessly with his workflow (e.g. via personalized text chunks or stamps).	Emphasizes minimizing manual input; a single prompt should trigger AI support without extra keystrokes.
P13	Values a manual prompt that lets him review and decide on AI suggestions (e.g. referral for CBT) rather than auto-populated actions.	No explicit comment provided.	Comfortable with moderate certainty; prefers only high-confidence recommendations to appear, though he's unsure about the calculation.	Desires robust editing options, including the ability to print patient education materials for personalization.	Implies a preference for minimal extra clicks; prefers streamlined interactions.
P11	Supports the AI assist button for its control over when the encounter is populated; considers implementation.	Accepts the design overall with no explicit concerns regarding color coding.	Finds the non-numeric low/medium/high display effective; appreciates that the summary indicates moderate anxiety based on GAD scores.	No explicit comment on customization; appears neutral on personalization features.	Not explicitly commented on; finds the workflow acceptable.

P12	Accepts the concept of a manual AI assist button and seeks clarification on whether the generated output is intended as a summary from a scribed encounter.	No explicit comment provided on the color-coded delineation.	Supports a tiered confidence score system (e.g. yellow, green, red) and is comfortable with moderate certainty; requests clarity on the calculation and differentiation of confidence levels.	Suggests that information should be clearly tailored to the patient, with enhancements like mouse-over tooltips to differentiate generic lifestyle suggestions from patient-specific data.	Emphasizes the need for minimal extra clicks, prefers auto-population of prescriptions and streamlined interactions.
P4	Prefers a manual trigger rather than automatic AI suggestions.	Sees value in distinguishing AI-generated text but concerned about potential confusion if clinicians use similar colors.	Prefers low/medium/high format , understands the need for a toggle and likes optionality.	No explicit comment on customization; appears neutral on personalization features.	Not commented on.
P5	Prefers manual control over AI engagement, supporting the opt-in model.	Finds color differentiation helpful for distinguishing AI-generated content from clinician inputs.	Believes confidence scores should be clearly presented but non-intrusive.	Supports flexibility in AI recommendations to align with different clinical practices.	Does not see an issue with current click-through journey.

P3	Prefers a manual trigger rather than automatic AI suggestions.	Sees value in distinguishing AI-generated text but concerned about potential confusion if clinicians use similar colors.	Prefers low/medium/high format over numeric scores but questions the need for a toggle option.	Supports an API for clinicians to customize workflows and decision support tools.	Emphasizes minimizing clicks and maintaining workflow efficiency.

Appendix F Part B: Tabulated answers from validation interviews in detail.

Participant	AI Suggested Treatment Plans	Volume of Information	AI Generated Summary Decision Support	Medication Suggestion for UTI	Other Key Insights	Overall Summary
P6	Favors a suggestion-based approach that offers guidance rather than explicit prescriptions.	Believes more information upfront may increase accountability—but cautions that too much can add responsibility.	No specific feedback provided on the summary presentation.	Finds it helpful if a medication suggestion (e.g. nitrofurantoin) is offered as an option to accept or reject.	Notes that incorporating AI-generated data makes clinicians feel more accountable; sees potential quality enhancement but remains cautious about information volume.	Generally supportive of the design with a preference for a softer, suggestion-based interface that empowers clinical judgment.
P7	Believes treatment suggestions should come	Warns that an overly lengthy note can	Suggests that summaries should	Supports medication suggestions offered as	Emphasizes workflow efficiency and	Thoughtful and detailed; he values flexibility

	with rationale or options rather than being directive; stresses tailoring to clinical context.	overwhelm; concise and relevant information is key.	clearly separate positive and negative aspects to support decision-making.	selectable options (or reflecting his defaults) rather than forced orders.	integration with existing clinical practices; highlights the need for both positive and negative data.	and efficiency while raising practical usability concerns that align with real-world clinical workflow.
P10	Stresses that treatment suggestions should be clear and non-prescriptive, offering options instead of mandates.	Concerned about extraneous details; prefers information that is concise and relevant.	Prefers a point-form, audit-friendly summary that distinguishes AI suggestions from clinician input.	In the UTI context, he prefers multiple antibiotic options based on guidelines rather than auto-population.	Highlights the need for non-intrusive, adaptable decision support that aids new clinicians; values clarity and auditability.	Supportive overall; values the design's balance between offering useful guidance and preserving clinician control, while questioning some implementation details.

P01	Skeptical of auto-prescription for uncomplicated UTI; favors treatment suggestions presented as options due to preset preferences.	Expresses concern about information overload; only concise, relevant details should be shown during visits.	Prefers a floating panel for decision support that provides clear, actionable recommendations without permanent clutter.	Prefers medication suggestions as selectable options rather than auto-imposed orders.	Underlines the importance of efficiency and minimal disruption; highlights the need for clarity in how AI confidence is communicated.	Generally positive but cautious, supports an AI that enhances efficiency and decision support without overwhelming the clinician, stressing customization and clarity.
P13	Cautious about auto-generated treatment plans; prefers suggestions that he can toggle or modify, particularly for UTI	Open to comprehensive yet well-organized, actionable information while being mindful of overload.	Finds the summary useful for quickly getting up to speed before a visit, as long as it can be edited later.	Prefers medication suggestions as selectable options rather than forced auto-prescriptions, so he can rely on his	Stresses the importance of having control over AI suggestions and maintaining flexibility in decision-making.	Sees the design as a helpful balance between AI assistance and clinical judgment; appreciates the option to customize and refine

	management.			own preferences.		recommendations.
P11	Very positive about the AI-generated summary for mental health, noting that it effectively includes key lifestyle recommendations.	Describes the volume of information as succinct and well-balanced, not overwhelming.	Very positive; appreciates that the AI-generated summary provides clear, actionable decision support.	In the UTI scenario, he finds the medication suggestion (e.g. Nitrofurantoin) appropriate for uncomplicated cases.	Highlights the effective separation between AI-generated and manual input (blue vs. green) and values succinctness.	Enthusiastic about the redesign. Sees it as a practical and efficient tool that enhances clinical decision support, especially for mental health, while remaining appropriately balanced for other scenarios.
P12	Believes that for straightforward cases (e.g. uncomplicated UTI), a	Finds the overall volume acceptable if well-organized	Appreciates an efficient, quick overview before a patient	Supports auto-populated medication suggestions for	Raises detailed questions about the integration of AI scribe	Generally positive but inquisitive, values efficiency and minimal

	general treatment recommendation is acceptable if it auto-populates the prescription to reduce extra clicks; seeks clarity on distinguishing between detailed assessment/plan and a summary.	but stresses the importance of personalization to avoid generic clutter.	encounter; seeks clear differentiation between the detailed note and summary outcome.	uncomplicated UTI (e.g. nitrofurantoin) to save time, provided they reflect his typical practice.	output with manual editing and clarity in the presentation of confidence levels and treatment recommendations.	extra work, and while he supports the design's direction, he seeks further clarity on how AI-generated data (especially confidence scores and summary versus detailed notes) will be distinguished and integrated into his workflow.
P4	Prefers multiple treatment options rather than a	Positive with volume of information	Believes AI summaries should align with SOAP note	Supports medication suggestions, reminds important of	Advocates for AI to enhance workflow by reducing	Finds the redesign beneficial expresses optimism for

	single AI-recommended choice, puts emphasis on if AI can recall what has historically been done for a patient to avoid present treatment decisions	being presented.	structure and maintain real estate for patient data.	antibiotic stewardship.	duplication, integrating seamlessly with lab orders, and assisting in documentation.	future development. Re-states expected inevitability of AI-integration within EMR use.
P5	Values AI-generated suggestions but prefers multiple options rather than a single recommendation.	Supports AI-generated summaries but warns against excessive detail that could slow down workflow.	Finds AI summaries beneficial for structuring clinical notes, states it would be good for medical students learning and questions	Sees AI medication suggestions as useful, provided they are based on verified patient history.	Believes AI should act as an efficiency tool, assisting in documentation and reducing administrative workload.	Strongly supports the redesign, emphasizing AI's role as an assistive and time-saving feature.

			how more seasoned clinicians may react to it.			
P3	Prefers multiple treatment options rather than a single AI-recommended choice, with context provided on past treatments.	Cautions against overwhelming clinicians with too much AI-generated content, as it could extend patient visit times.	Believes AI summaries should align with SOAP note structure and maintain real estate for patient data.	Supports medication suggestions but stresses that recommendations should be based on patient history rather than self-reporting.	Advocates for AI to enhance workflow by reducing duplication, integrating seamlessly with lab orders, and assisting in documentation.	Finds the redesign beneficial but emphasizes the importance of AI remaining an assistive tool rather than a directive one.