*Enhancing*

# PUBLIC HEALTH DISEASE INVESTIGATIONS THROUGH AI SCRIBE TECHNOLOGY:

## A Pilot Study in Two Ontario Public Health Units

NOVEMBER 28

2025

# THANK YOU
## for your contributions

PublicHealth
WELLINGTON-DUFFERIN-GUELPH
Stay Well.

simcoe muskoka
DISTRICT HEALTH UNIT
Your Health Connection

TORONTO Public Health

UNIVERSITY OF WATERLOO

AI4PH AI FOR PUBLIC HEALTH

UNIVERSITY OF TORONTO
DALLA LANA SCHOOL OF PUBLIC HEALTH

# Acknowledgements


## Report Written by

**Dr. Helen Chen**, Professor of Practice, School of Public Health Sciences, University of Waterloo

**Casey Hirschfeld**, Evaluation Specialist, Simcoe Muskoka District Health Unit

**Ravi Shah**, Health Data Analyst, Wellington-Dufferin-Guelph Public Health

**Nauman Shakeel**, PhD Student, School of Public Health Sciences, University of Waterloo; AI4PH Intern, Dalla Lana School of Public Health, University of Toronto

**Dr. Steven Rebellato**, Vice President Environmental Health Department, Simcoe Muskoka District Health Unit

**Natalie Riewe**, Vice President Clinical Services Department, Simcoe Muskoka District Health Unit

## Report Reviewers and Contributors

**Dr. Michael Finkelstein**, Deputy Medical Officer of Health- Health Protection, Toronto Public Health

**Katie Goss**, Graphic Designer, Simcoe Muskoka District Health Unit

**Michael Whyte**, Manager Data and Analytics, Wellington-Dufferin-Guelph Public Health

## AI Scribe Locally Driven Collaborative Project Team

**Rose Aydlin**, BSc Student, School of Public Health Sciences, University of Waterloo

**Dr. Helen Chen**, Professor of Practice, School of Public Health Sciences, University of Waterloo

**Dr. Michael Finkelstein**, Deputy Medical Officer of Health- Health Protection, Toronto Public Health

**Casey Hirschfeld**, Evaluation Specialist, Simcoe Muskoka District Health Unit

**Rivalia Naidoo**, Health Promotion Specialist, Wellington-Dufferin-Guelph Public Health

**Dr. Shovita Padhi**, Associate Medical Officer of Health, Toronto Public Health

**Dr. Steven Rebellato**, Vice President Environmental Health Department, Simcoe Muskoka District Health Unit

**Natalie Riewe**, Vice President Clinical Services Department, Simcoe Muskoka District Health Unit

**Ravi Shah**, Health Data Analyst, Wellington-Dufferin-Guelph Public Health

**Nauman Shakeel**, PhD Student, School of Public Health Sciences, University of Waterloo; AI4PH Intern, Dalla Lana School of Public Health, University of Toronto

**Edson Takei**, BSc Student, School of Public Health Sciences, University of Waterloo

**Taylor Thorpe**, Program Assistant, Wellington-Dufferin-Guelph Public Health

**Michael Whyte**, Manager Data and Analytics, Wellington-Dufferin-Guelph Public Health

**Dr. Kyle Wilson**, Vice President Information Systems & Digital Innovation, Wellington-Dufferin-Guelph Public Health

## Other Supports

**Melanie Dupuis**, Research Analyst, Simcoe Muskoka District Health Unit

**Infectious Disease Investigators and Infectious Disease Managers** who piloted AI scribe technology at Simcoe Muskoka District Health Unit and Wellington-Dufferin-Guelph Public Health

**IT, Finance, and Procurement Staff** who supported purchasing, securing, and installing AI scribe at Simcoe Muskoka District Health Unit and Wellington-Dufferin-Guelph Public Health

*The AI Scribe Locally Driven Collaborative Project Team would like to thank Public Health Ontario (PHO) for its support of this project. The team gratefully acknowledges funding received from PHO through the Locally Driven Collaborative Projects program. The views expressed in this publication are the views of the project team, and do not necessarily reflect those of Public Health Ontario.*

**For additional information, please contact:**

- Dr. Helen Chen, University of Waterloo, Lead Academic Researcher, helen.chen@uwaterloo.ca
- Natalie Riewe, Simcoe Muskoka District Health Unit, LDCP Co-Lead, natalie.riewe@smdhu.org
- Dr. Kyle Wilson, Wellington-Dufferin-Guelph Public Health, LDCP Co-Lead, kyle.wilson@wdgpublichealth.ca

Suggested Citation:

Chen, H., Hirschfeld, C., Shah, R., Shakeel, N., Rebellato, S., Riewe, N., and AI Scribe LDCP Team. (2025). Enhancing public health disease investigations through AI scribe technology: A pilot study in two Ontario public health units. AI Scribe LDCP Team.

# Table of Contents

# Executive Summary

## Introduction

Public health disease investigations for Diseases of Public Health Significance (DOPHS) require extensive documentation, creating significant administrative burden for infectious disease (ID) investigators. This pilot study explored whether ambient AI scribe technology could reduce this burden and improve workflow efficiency without compromising documentation quality.

## Objectives

The project aimed to:

- Assess the effectiveness of AI scribe technology in reducing administrative workload.
- Evaluate the accuracy and completeness of AI-generated documentation.
- Identify benefits and challenges experienced by ID investigators.
- Provide recommendations for scaling AI scribe use in public health units (PHUs).

## Methods

The study was conducted by Simcoe Muskoka District Health Unit (SMDHU) and Wellington-Dufferin-Guelph Public Health (WDGPH) under the Locally Driven Collaborative Projects program. Fifteen AI scribe licenses were deployed between June 2025 and October 2025. Evaluation methods included simulation testing, time tracking, error tracking, user surveys, and focus groups. Research Ethics Board approval and a Privacy Impact Assessment were completed prior to implementation.

## Key Findings

- Documentation Time: SMDHU users saved an average of 12.9 minutes per interaction ($p<0.05$), while WDGPH users spent 10.3 minutes more ($p<0.05$). Differences may reflect variations in experience and case volume.
- Accuracy: There was an 80% agreement between the transcript and AI scribe generated note during simulation testing, but it occasionally hallucinated content when encounters were incomplete. Users developed custom templates to address gaps.
- User Experience: 86% of survey respondents found the tool easy to use; 71% reported accurate documentation; and most perceived reduced cognitive load. Focus groups

highlighted positive impacts on workflow and work-life balance at SMDHU, while WDGPH reported limited benefits due to fewer applicable cases.

- Challenges: Technical limitations included headset compatibility, lack of template standardization, and inability to integrate with documentation systems (e.g., iPHIS).

## Limitations

The study faced a small sample size (44 interactions using AI scribe), low survey response rate, and limited focus group participation, reducing generalizability.

## Conclusion

AI scribe technology shows promise for reducing administrative burden and improving documentation quality in public health investigations. However, successful implementation requires robust governance, technical adaptability, and integration with existing systems. Longer-term pilots and broader evaluations are recommended to confirm sustainability and scalability.

# Introduction and Background

Public health professionals engaged in disease investigations of Diseases of Public Health Significance (DOPHS) often experience high administrative burden due to extensive documentation requirements associated with these investigations. DOPHS such as salmonellosis, pertussis, chickenpox, invasive group A streptococcal disease (iGAS), and Lyme disease demand timely, accurate and comprehensive reporting. These infectious disease (ID) investigation processes involve extensive manual notetaking and often duplication of investigatory documentation across electronic information systems. Further, ID investigations are inherently dynamic and evolve in response to emerging information such as laboratory results, updated clinical presentations, or the identification of additional contacts. This evolving nature of ID investigation requires continuous documentation, further compounding the administrative workload, and increasing the potential for variability and inefficiencies. Addressing these challenges requires innovative solutions that maintain data quality while improving workflow efficiency and help reduce administrative burden.

Advanced automation and speech technologies, such as voice-to-text transcription and artificial intelligence (AI) scribes, have shown promise in reducing documentation burdens within primary care settings (Centre for Digital Health Evaluation, 2024) by improving efficiency without compromising data quality. These tools use speech recognition, large language models (LLMs) and natural language processing (NLP) to capture and organize clinical information, allowing providers to focus on patient care rather than administrative tasks. AI scribe technology has undergone multiple evaluations in primary care settings, including OntarioMD evaluation in 2024 (Centre for Digital Health Evaluation, 2024). OntarioMD evaluation demonstrated significant reductions in the time clinicians spend on administrative tasks, particularly the documentation of patient encounters without compromising the accuracy, quality, or completeness of clinical records (Centre for Digital Health Evaluation, 2024). Integrating this technology within public health units (PHUs) presents an opportunity to address critical inefficiencies in the investigation of DOPHS. This project seeks to assess the applicability of ambient AI scribe technology in context of DOPHS investigations to explore whether similar benefits can be realized in public health context.

The purpose of this pilot study was to determine if AI ambient scribe technology can reduce the administrative burden of skilled public health ID investigators, enabling them to dedicate more time to complex investigations and urgent public health responses. This is particularly vital under the Ontario Health Protection and Promotion Act, which mandates timely and accurate disease investigation as a cornerstone of effective public health intervention. By addressing existing inefficiencies, the pilot sought to determine if AI scribe may enhance the overall effectiveness and responsiveness of public health disease investigation processes.

To explore this potential, the AI scribe Locally Driven Collaborative Project (LDCP) launched as a joint initiative between Simcoe Muskoka District Health Unit (SMDHU) and Wellington-Dufferin-

Guelph Public Health (WDGPH), in partnership with Toronto Public Health (TPH), the University of Toronto (Ai4PH program), and the University of Waterloo (UWaterloo). The LDCP program unites PHUs with academic and community partners to collaboratively design and carry out applied research and program evaluation projects that address significant and shared public health interests. This report presents the findings of the pilot project, which evaluated the feasibility and impact of integrating AI scribe technology into public health ID investigation workflows.

Informed by findings of the 2024 OntarioMD AI Scribe evaluation on primary care settings (Centre for Digital Health Evaluation, 2024), the AI scribe LDCP project focused on assessing the implementation of this technology to improve efficiency and documentation quality in disease investigations of DOPHS. A total of 15 AI Scribe licenses were deployed between SMDHU and WDGPH ID investigators. The initiative was guided by a comprehensive evaluation framework, supported by a Privacy Impact Assessment (PIA) and Research Ethics Board (REB) approval. Due diligence to procure an appropriate AI scribe vendor was completed including comprehensive system testing, legal and privacy administration prior to pilot initiation. This was completed to ensure the project's security, privacy, and operational effectiveness across users in both PHUs.

To evaluate the feasibility and impact of this implementation, the project focused on the following central research questions:

- How does AI scribe technology impact the time spent on administrative tasks by ID investigators?
- What are the perceived benefits and challenges of integrating AI scribe technology in local PHUs?
- Does the use of AI scribe technology improve the accuracy and completeness of ID investigation records?

Based on these questions, the project identified four key objectives:

- Assess the effectiveness of AI scribe technology in reducing administrative workload.
- Evaluate the accuracy and completeness of documentation produced by AI scribe technology when compared to similar investigations not using an AI scribe.
- Identify the benefits and challenges experienced by ID investigators using AI scribe technology.
- Provide recommendations for scaling the use of AI scribe technology in PHUs.

To contextualize these objectives and inform the evaluation approach, a review of existing evidence on AI scribe technology and its impact on clinical and public health workflows was completed. The following literature review summarizes current findings in this area and highlights the gaps that this project aims to address.

# Literature Review

A literature scan was conducted using the following the search terms:  scribe AI evaluation in health, scribe AI in health, AI scribe evaluation (year range 2023-2025), AI scribe and public health, AI Scribe, AI Scribe evaluation, medical scribe, Scribe healthcare and AI healthcare documentation. The focus of the scan was to identify evaluation methods for assessing the role of scribes in health-related settings, given that the use of scribe technology in public health is a relatively recent development. Studies were excluded if they primarily addressed the design or development of scribe technology rather than its evaluation. No exclusions were made based on the study design.

There were 108 articles screened with 54 full text reviewed, resulting in 14 studies included for this literature review.

Seven studies directly discussed the evaluation of an AI scribe tool (Bhattacharyya et al., 2025; Bindal et al., 2025; Biro et al., 2025a, 2025b; Haniff et al., 2025; van Buchem et al., 2024; Moryousef et al., 2025). Bhattacharyya et al. (2025) examined an AI scribe in primary care to assess its ability to reduce administrative burden, support clinician well-being, and guide adoption across Canada. The study was funded by Ontario's Ministry of Health and conducted with OntarioMD and the eHealth Center of Excellence. The study used a two-phase design involving both simulation and real-world testing. In the simulation phase, provider documentation from simulated patient encounters was compared with and without the AI scribe to benchmark tool performance.  In the real-world phase, three AI scribe solutions were deployed among 152 clinicians and nurse practitioners in varied settings, and data were collected through surveys and interviews at baseline and three-month follow-up. The key metrics used in this study included documentation efficiency, accuracy of documentation, and reduction in after-hours administrative time.

The study also explored clinician satisfaction and the impact on work-life balance. The simulation results found a four-minute reduction per encounter, and all the AI scribes tested met acceptable quality standards, although performance varied. In the real-world phase, 83% of clinicians used the scribe daily and reported an average reduction of more than three hours per week on administrative tasks. Additionally, 55% of clinicians reported improved work-life balance (Bhattacharyya et al., 2025). Similarly, Bindal et al. (2025) evaluated an AI scribe system integrated with the electronic health record (EHR). The study found that the scribe produced high transcription accuracy and reliability across complex medical conversations. It effectively produced structured EHR reports, including patient history, diagnostics, and treatment plans. Omission was identified as the most critical and frequent error made by the scribe in simulation, which can hinder documentation safety and accuracy across multiple studies (Biro et al., 2025a, 2025b).

Four studies explored the concept of AI scribe through a LLM (Sezgin et al., 2023; Kernberg et al., 2024; Luo et al., 2025; Palm et al., 2025). These studies investigated the generation of subjective, objective, assessment, and plan (SOAP) notes by GPT-4 and other LLMs and compared them with checklist-based or human-authored notes. Kernberg et al. (2024) found that omission errors were the most frequent, and that longer transcripts and more complex encounters were inversely correlated with note accuracy. They found some additions to be fabricated hallucinations, while omissions included clinically relevant elements such as lab results or treatment recommendations. Luo et al. (2025) highlighted significant gaps in systematic evaluation, especially regarding patient outcomes and long-term organizational impact. Meanwhile, Palm et al. (2025) compared AI scribe-generated notes against human-authored notes, finding that the scribe was more thorough and organized but lower in accuracy and consistency.

Two systematic reviews focused on the evaluation of AI implementation in healthcare (Hassan et al., 2025; Ng et al., 2025). Ng et al. (2025) explored studies on AI-based transcription tools (automatic speech recognition, natural language processing, and LLM-enabled digital scribes) for clinical documentation, focusing on their accuracy, efficiency, usability, and integration into healthcare workflows. They examined word error rate (WER), F1 score, precision, recall, time savings, turnaround time, workflow speed, and costs compared with manual transcription. Findings on cost were inconsistent; some settings demonstrated cost savings, whereas others found automatic speech recognition more expensive than human transcription. In some instances, AI-assisted notes captured more clinically relevant details than physicians' manual notes, while in others, high error rates posed potential patient safety risks. Summarization performance improved but still required human editing; ROUGE F1 scores increased with human post-processing. Clinicians raised concerns about accuracy, especially with specialized terminology and accented speech (Ng et al., 2025).

Lastly, Wang et al. (2025) proposed an evaluation framework named SCRIBE for assessing ambient digital scribing tools in clinical settings. The metrics included efficiency, measured via time-motion studies, EHR log data, and chart completion times; quality, assessed through independent chart reviews, error rates, and completeness scores; clinician outcomes, measured using validated surveys (e.g., burnout inventories, workload scores); patient outcomes, evaluated through post-visit surveys, interviews, and satisfaction scales; and economic outcomes, including reported cost savings, staff time redistribution, and operational data. Overall, they concluded that AI scribes show promise for clinical integration, but more rigorous, long-term studies are needed to confirm their sustainability and broader impacts (Wang et al., 2025). Simulation testing was used consistently in most studies and was even advocated by Biro et al. (2025a) as an efficient way to evaluate AI scribes (Bhattacharyya et al., 2025; Biro et al., 2025b; Kernberg et al., 2024).

# Planning and Implementation of AI Scribe Technology

## Procurement and Vendor Selection

In alignment with established procurement policies, WDGPH led the procurement process for AI scribe technology on behalf of both participating PHUs (SMDHU and WDGPH). SMDHU ensured the process aligned with its own internal procurement protocols, maintaining a consistent and collaborative approach. A comprehensive evaluation matrix was developed to support the procurement of AI scribe technology, which was also informed by the OntarioMD AI Scribe Evaluation. This matrix incorporated criteria including adherence to privacy and security legislation, technical capability, and clinical usability (see Appendix 1). Several vendors were invited to provide demonstrations, which were assessed using the evaluation matrix to ensure a transparent and evidence-informed selection process.

Following a competitive procurement process in Q1 2025, and thorough legal and contractual review was completed in March 2025 followed by the selection of a preferred vendor. Key considerations in the decision included the vendor's compliance with Personal Health Information Protection Act (PHIPA), and SOC 2 standards; demonstrated effectiveness in clinical settings; and user-friendly interface. Integration capability with electronic medical records (EMRs) was also a critical factor, supporting scalability and future expansion. Contract negotiations extended beyond vendor selection to address important considerations related to data ownership, privacy, cybersecurity, and insurance.

## Legal, Contractual Terms, Privacy & Security

The negotiation process with the selected vendor for the AI scribe project led to several follow-up activities requiring an extended review and collaboration between the vendor, legal advisors, and representatives of both SMDHU and WDGPH. Activities focused on data ownership, compliance with PHIPA, data residency, breach response, and vendor obligations related to third-party access and subcontractors. Clarification was received on de-identified usage data, cyber security, cyber breach insurance coverage, in addition to clear provisions for data retention and destruction. The final agreement reflects a robust legal framework that aligns with Ontario's public sector expectations for digital health service providers. It also demonstrates the importance of detailed legal due diligence in technology-enabled public health partnerships, ensuring that innovation does not come at the expense of public trust or legal compliance.

## Privacy Impact Assessment

A Privacy Impact Assessment (PIA) was completed in May 2025. The PIA confirmed that the AI scribe technology adhered to PHIPA, Municipal Freedom of Information and Protection of Privacy Act (MFIPPA) and SOC 2 Type II standards. The PIA also incorporated additional privacy safeguards, including:

- Explicit verbal consent from clients before initiating any recording.
- Users setting an automatic deletion of transcripts after 48 hours.
- Secure user access controls and audit logs.
- No secondary use of personal health information.
- PIA mechanisms to ensure that privacy risks are mitigated while maintaining the integrity of the evaluation.

## Client Consent to Use AI Scribe Technology

The AI scribe project incorporated a consent process based on a script provided by the vendor and adapted for use for using AI scribe in voice call interactions. The model supports verbal consent, with clients informed that an AI scribe may be used to support documentation during the ID investigation. A script was provided to guide ID investigators in explaining the role of the AI scribe, its privacy protections, and the client's rights to opt out at any time (see Appendix 2). During this pilot, ID investigators were instructed that consent must be reaffirmed prior to each use of the AI scribe, and if declined, traditional documentation methods were used without change in disease investigation quality and approach.

The consent process highlighted that AI scribes will capture audio to generate draft notes, which were reviewed and edited by the ID investigator before becoming part of the record. All audio and transcripts are deleted in accordance with privacy policies and vendor commitments. ID investigators were trained to document that verbal consent was received, including use of standardized consent stamps or AI scribe-generated notations.

This approach prioritizes transparency, voluntary participation, and alignment with Ontario privacy legislation. Clients who requested more information were offered vendor documentation on the technology. This ensured that consent was informed, ongoing, and tailored to uphold individual privacy rights.

## Implementation of AI Scribe

Implementation of AI scribe technology at the two PHUs consisted of the following components:

### 1. Pilot Participant Recruitment, Selection, and Leadership

ID investigators were recruited and selected at both participating PHUs to pilot the AI scribe for ID investigations. Recruitment was initiated through targeted communication to the ID managers at

each PHU, requesting identification of appropriate staff to participate based on current roles, investigation experience, and interest in piloting innovation.

Following recruitment, a joint planning meeting was held with some of the ID team members from WDGPH and SMDHU, together with LCDP project team representatives. The purpose of this meeting was to jointly select DOPHS appropriate for piloting the AI scribe. The DOPHS were selected collaboratively, and the selection process was guided by historical investigation volume from previous years, manager expertise regarding local epidemiology, and practical feasibility within the pilot timeline.

While historical data informed the shortlist, managers emphasized that disease frequency and investigation influx vary seasonally. Therefore, DOPHS were selected with awareness that case volumes might fluctuate depending on time of year, and that the pilot should remain adaptable to those changes. Through consensus across both PHUs, the initial pilot DOPHS were Salmonella, Pertussis, Chickenpox at WDGPH, and iGAS at SMDHU. Later expansion of these DOPHS included Vector-Borne Diseases (VBD) at WDGPH and Lyme Disease at SMDHU. This approach allowed the pilot to adapt with real-world investigation volumes and increase the data collection over time.

### 2. Leadership and Scribe Champion Model

To support implementation, one staff member at each PHU was designated as the local pilot lead. These individuals acted as responsible for coordinating templates, testing software in real workflows, and serving as the primary contact for pilot participants. SMDHUs lead was an ID investigator and champion who also participated directly in the pilot and led most aspects of template development. WDGPHs lead was a project team member supporting template development and distribution. This ensured both PHUs adapted the same custom templates while starting the pilot and any changes to custom templates were addressed and incorporated at both PHUs to maintain the consistency of these templates.

This distributed leadership model ensured each PHU had an accessible on-site lead who could provide rapid support, reinforce consistent use, and communicate emerging issues to the project team.

### 3. Template Development for Selected DOPHS

Both PHUs agreed that adoption of AI scribe required standardized documentation templates tailored to the selected DOPHS. Learnings and findings from the simulation testing at UWaterloo informed the development of templates, including explicit prompts to ensure the scribe captured the relevant information for documentation. Templates were designed jointly to maintain cross-PHU consistency and comparability of pilot results. Custom templates were created by referencing the internal investigation tools used at WDGPH and SMDHU, and Ontario Investigation Tools to capture most of the information required for documentation.

Templates were aligned across both PHUs so that data capture, field structure, and documentation workflow remained consistent. Initial versions were developed with limited vendor support; however, as vendor involvement decreased after early setup, ongoing edits and refinements were completed independently by PHU staff. Template updates were tracked through regular cross-PHU check-ins to ensure both PHUs remained synchronized on revisions and implementation changes.

## 4. Training

Training was delivered in two layers:

- **PHU-led orientation sessions**: LDCP project team members at each PHU delivered training on privacy, legal considerations, and investigation workflow integration. This included practical onboarding support such as credential assignment, testing the AI scribe in advance of live use, and ensuring readiness for ID investigations.
- **Vendor-led software training**: The software vendor delivered tool-specific training for ID investigators and select LDCP project team members. Vendor demonstrations were complemented by PHU-developed supports to improve onboarding and confidence.

To reinforce training and support independent use, both PHUs created and distributed shared resources, including:

- An FAQ sheet covering project goals, expected use, and troubleshooting,
- An interactive user guide in a short printable format to support real-time investigations, and
- A client consent document to ensure investigators had ready access to compliant consent language during calls.

Additional demo videos and workflow-specific sessions clarified how the AI scribe would be used alongside soft-phone investigations, including how to manage recording start/stop, call flow, and simultaneous notetaking.

## 5. Technical Integration

Both PHUs collaborated with internal IT staff to ensure a secure and functional technical setup. IT teams supported:

- Onboarding of users to the AI scribe platform,
- Testing and resolving headset and audio configuration issues, and
- Confirming optimal access methods (web-based platform or browser extensions).

At WDGPH, ID investigators later transitioned to a browser extension. The IT team supported migration and configuration to ensure continued alignment with privacy and security requirements without hampering the project activity.

## 6. Ongoing Support and Collaboration

Regular support meetings were built into pilot operations to maintain momentum and address challenges early. These included:

- **PHU-level check-ins:** Recurring meetings were held with ID investigators, ID manager(s), and LDCP project members at each PHU. These sessions focused on trust-building, troubleshooting emerging issues, addressing technical barriers, and refining templates based on real-world use.
- **LDCP project oversight meetings:** Weekly cross-project meetings between both PHUs and the LDCP project team were held to review progress, share learnings, and coordinate refinements across sites and ensure project implementation is adaptable to emerging needs.

Insights gathered through implementation discussions were continuously fed back to the broader project team to support ongoing project re-evaluation and iterative improvement of both workflows and templates.

# Evaluation Methodology

## Research Ethics

Ethical oversight was obtained through the University of Waterloo's Research Ethics Board (REB #47284), with approval granted on March 11, 2025. An administrative review by Public Health Ontario's Ethics Research Board followed and was approved on March 28, 2025. The project adhered to Tri-Council Policy Statement (TCPS 2; 2022) and Ontario provincial privacy legislation.

## Simulation Testing

Simulation testing was conducted on twelve sample DOPHS interactions to evaluate the AI scribe under varying conditions. These conditions included stopping and restarting an encounter, incorporating pauses or periods of silence, introducing misinformation and subsequent corrections, adding small talk before the conversation begins, background noise, and diverse accents. Using the prepared use cases, the information generated by the scribe was evaluated against the scripted conversations. These measures were intended to simulate real world settings. In simulation, we evaluated the scribe by comparing the generated note against a reference checklist (see Appendix 3) that outlined the minimum required information for reporting. This checklist was developed based on the use cases and scenarios used as part of the evaluation, focusing on the quality of the documentation generated. UWaterloo developed an automated tool to extract text from the transcript and SOAP notes, then used a large language model with a prompt to determine whether each item on the reference checklist was present or absent. For example, if the transcript included client demographics, the SOAP note should also document them. Conversely, if a specific exposure was not mentioned in the transcript, it should not appear in the SOAP note. We measured the level of agreement between the transcript and the generated note based on the reference checklist.

## Time Tracking

To understand the impacts of time spent on administrative tasks by ID investigators using AI scribe, a time tracking method using Microsoft Excel® was employed. Completion of time tracking was a component of routine workflow for ID investigators while using AI scribe technology. This included investigators tracking time using both the traditional documentation method (i.e., without the use of AI scribe technology) and by using AI scribe technology for documentation.

Pilot data was analyzed using time tracking data, estimating the average reduction in documentation time with and without the AI scribe. The data was stratified by organization, Cliff's

delta was computed, and Welsch's t-test was conducted to understand if the mean reduction or potential increase was statistically significant.

## Error Tracking

To understand the accuracy and completeness of records using AI scribe technology, Excel® was used to document errors made by the AI scribe technology and any impacts these errors had on documentation time. Completion of error tracking was also a component of the routine workflow for ID investigators while using AI scribe technology. The ID investigator reviewed the documentation output provided by the AI scribe, corrected any errors or omissions, and tracked error-related information in the Excel® document. This data was analyzed for the total number of errors and mean editing times.

## User Survey

A user survey was employed to collect ID investigators' experiences using the AI scribe technology for ID investigations (see Appendix 4). The primary purpose of the survey was to understand user experience, benefits, and challenges with the technology, including functionality, performance, and compatibility in public health unit workflows. The survey, sent via email and collected using an online survey platform, was open for a 2-week period. In total, 7 of the 15 ID investigators participated in the survey. Close-ended survey questions were analyzed using counts and percentages. Open-ended survey questions were analyzed using thematic analysis.

## User Focus Groups

Semi-structured focus groups were conducted at each respective PHU to collect more in-depth qualitative information regarding ID investigators and managers experiences implementing the AI scribe technology for ID investigations, including benefits, challenges, and lessons learned/opportunities (see Appendix 5). The sample included 5 ID investigators and 1 manager from SMDHU and 10 ID investigators and 1 manager from WDGPH. In total, SMDHU had 5 of 6 individuals participate in a focus group and WDGPH had 5 of 11 participate. Focus group notes were validated using recordings and member-checking. After validation, the notes were compiled and analyzed using a qualitative thematic approach.

# Evaluation Findings

AI scribe was used by ID investigators at SMDHU and WDGPH between June and October 2025. During that time, there were 42 interactions with AI scribe and 20 without. Table 1 further breaks this down by DOPHS. In total, 10 of the 15 ID investigators had an opportunity to use the AI scribe, 5 from SMDHU and 5 from WDGPH.

*Table 1. Number of interactions with and without AI scribe by PHU and DOPHS.*

| PHU | DOPHS | Number of Interactions | | Number of Cases |
|---|---|---|---|---|
| | | With AI Scribe | Without AI Scribe | |
| **SMDHU** | Lyme | 4 | 5 | 9 |
| | Salmonellosis | 19 | 4 | 21 |
| | iGAS | 4 | 1 | 3 |
| **WDGPH** | Lyme | 9 | 8 | 17 |
| | Salmonellosis | 6 | 2 | 8 |
| | Chickenpox | 1 | 0 | 1 |
| | Anaplasmosis | 1 | 0 | 1 |
| **Total** | | **44** | **20** | **60** |

## Quality of AI Scribe Documentation

Simulation testing revealed that the AI scribe effectively captured relevant investigation information while filtering out irrelevant content, and its performance remained consistent across various testing conditions. By default, the template within the AI scribe did not include personal client information, so it was explicitly prompted to add it. Additionally, it was prompted not to populate sections when information was not provided. When an encounter was stopped and restarted, the AI scribe regenerated the note without access to the original version; moreover, if the conversation was incomplete and a structured note was selected, the system may hallucinate content such as inventing a treatment when treatment recommendations were not discussed. These risks were mitigated by prompting the AI scribe to avoid filling sections that were not covered during the encounter. On average, the AI scribe had an 80% agreement between the information from the transcript and the SOAP note generated based on the reference list. The

generated note followed a SOAP format, which does not cover the full scope needed for public health documentation, and therefore some content was missing (e.g., missing demographic information). To address this, users in the real-world testing created a custom template tailored to their organizational needs.

## Impact of AI Scribe on Documentation Time

On average, SMDHU users spent about 12.9 minutes less ($p<0.05$) documenting with AI scribe compared to those without it on each interaction. In contrast WDGPH users spent about 10.3 minutes more ($p<0.05$) documenting with AI scribe than those without it. Overall, a total of 148 errors were identified in documentation produced by the AI scribe, with 65 errors at SMDHU and 83 at WDGPH. This represents an average of 3.4 errors per interaction with the AI scribe. Mean editing times were 7.04 minutes for SMDHU and 12.94 minutes for WDGPH.

A further breakdown by DOPHS is shown in Table 2. For SMDHU, Lyme disease cases had a mean reduction in documentation time of 10.5 minutes per interaction; however, this difference was not statistically significant. In contrast, Salmonellosis showed a statistically significant mean reduction of 6.79 minutes. For WDGPH, Lyme disease had a statistically significant mean increase in documentation time of 14.31 minutes, whereas Salmonellosis had a mean increase of 2.5 minutes that was not statistically significant.

*Table 2. Mean documentation time by disease using Cliff delta test.*

| PHU | DOPHS | Mean Documentation Time (Min) | | Cliff's Delta | Confidence Interval (CI) | |
|---|---|---|---|---|---|---|
| | | With AI Scribe | Without AI Scribe | | Lower | Upper |
| **SMDHU** | Lyme | 22.50 | 33.00 | 0.550 | -0.36 | 0.92 |
| | Salmonellosis | 9.88 | 16.67 | 0.804 | 0.43 | 0.94 |
| **WDGPH** | Lyme | 25.56 | 11.25 | -0.875 | -0.98 | -0.36 |
| | Salmonellosis | 17.50 | 15.00 | -0.250 | -0.86 | 0.66 |

## Experiences Using AI Scribe

### Survey Findings

The majority of survey respondents reported their overall experience using AI scribe as good or very good (see Figure 1). Figure 2 further breaks down survey respondents' level of agreement on various components of their experience using the scribe, including that all respondents reported it was useful and 86% found it easy to use and navigate. Additionally, 71% of users reported the

documentation of the scribe as accurate and it made the documentation process more efficient. The majority also reported AI scribe decreased documentation time (see Figure 3). These results suggest AI scribes can enhance ID documentation without compromising data quality.

**"Positive overall impression of AI integration into everyday work. It was easy to use and has the potential to increase efficiency and accuracy of work." – User Survey Respondent**

*Figure 1. Survey respondents overall experience using AI scribe.*

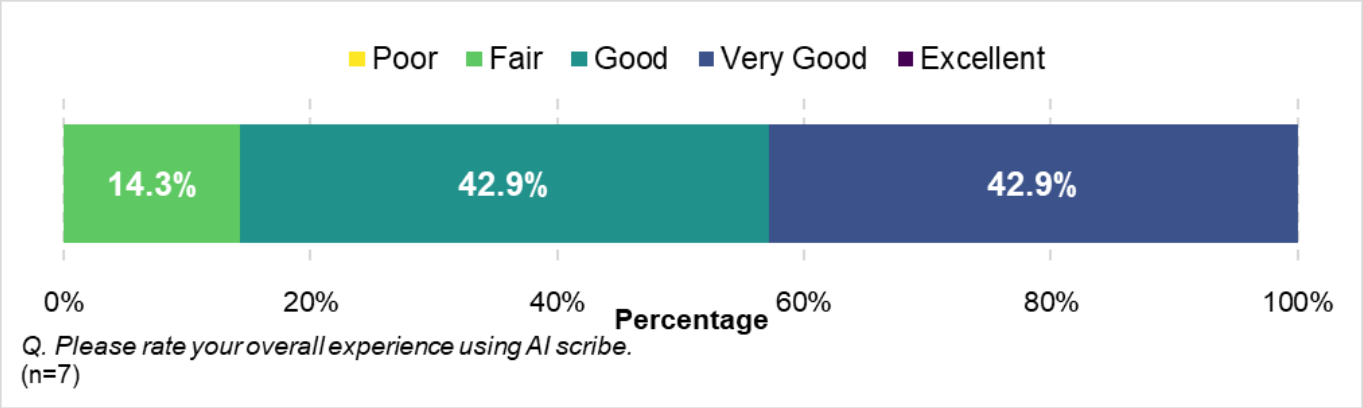



Figure 2. AI scribe user experience of survey respondents.

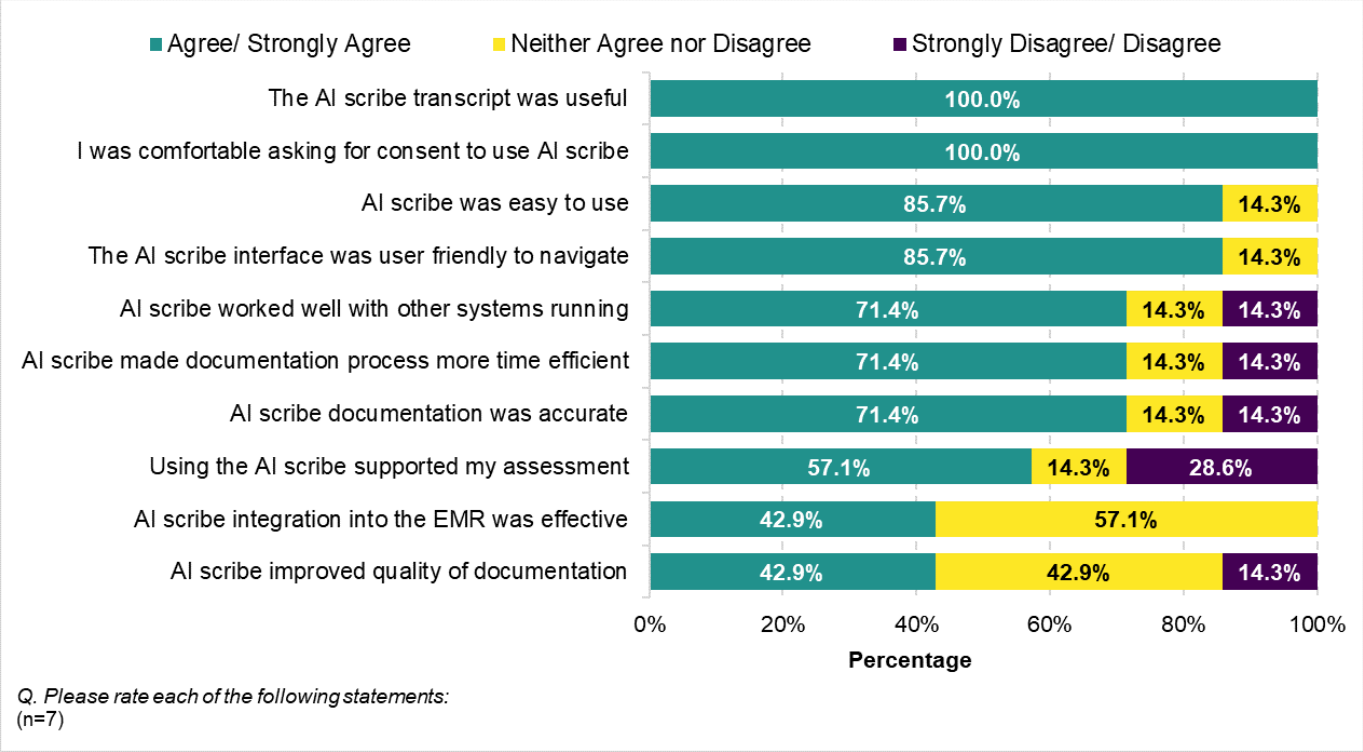
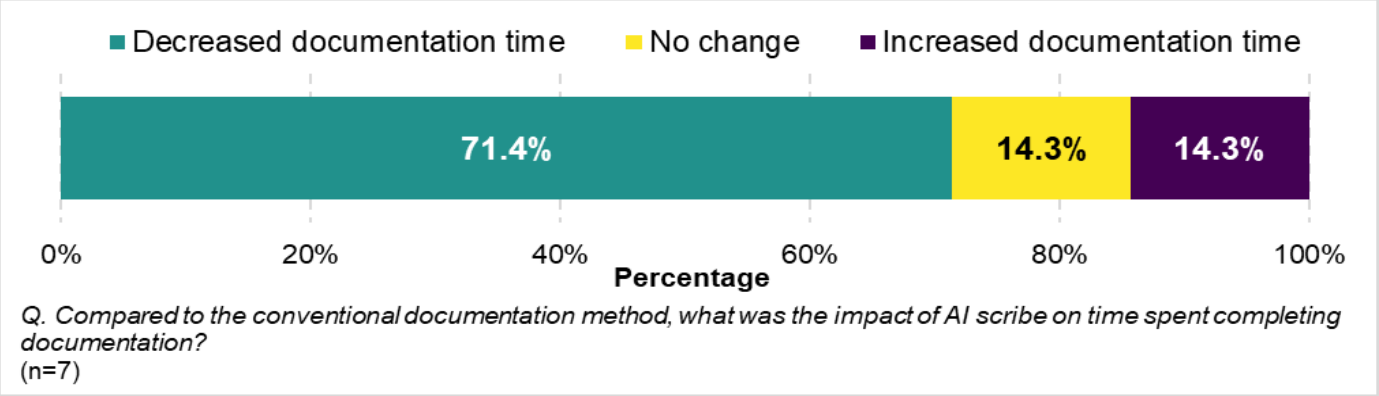


*Figure 3. Survey respondent perceived impact of AI scribe on time spent completing documentation.*

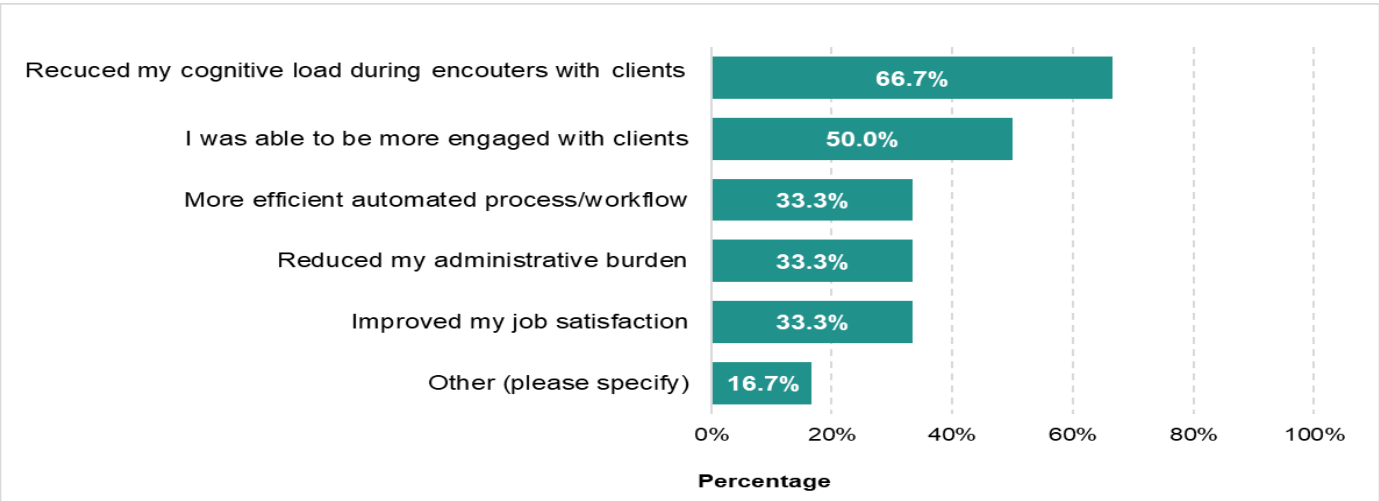



*Q. Compared to the conventional documentation method, what was the impact of AI scribe on time spent completing documentation?*
(n=7)

The most common benefit reported by survey respondents in implementing AI scribe was a reduction in cognitive load during encounters with clients (66.7%; see Figure 4). One limitation reported by survey respondents included the creation of concise investigation notes (40.0%; see Figure 5). This is important as ID investigation notes must include essential details (e.g., dates, locations) that were sometimes omitted by the AI scribe. Additionally, the AI scribe occasionally captured irrelevant content, such as casual conversations unrelated to the investigation. Another limitation reported by survey respondents was difficulty distinguishing between multiple speakers (40.0%). Other limitations noted by respondents included challenges with the technology not capturing public health ID investigation information (e.g., locations, animals, foods, dates, timelines) in the summarized note and the need for ID investigation specific templates.

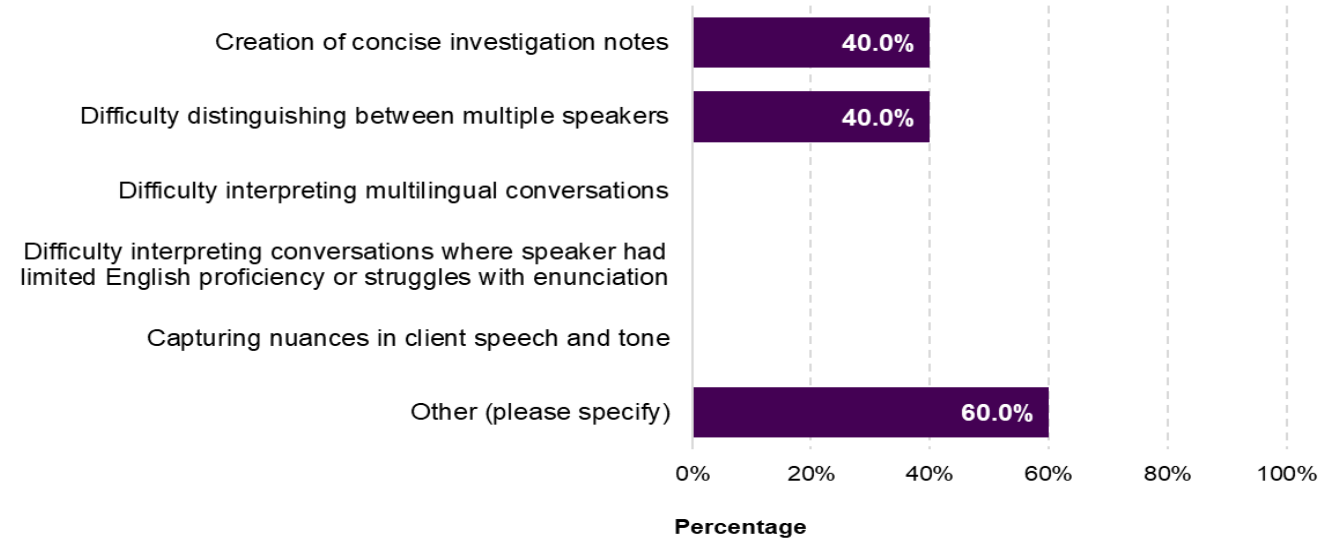
**"I continued to make notes during the phone call to ensure that the AI scribe recorded information properly. I don't see myself changing this." – User Survey Respondent**

*Figure 4. Benefits experienced by survey respondents in implementing AI scribe technology.*




*Q. Please identify any limitations of the AI scribe technology that you experienced. (Check all that apply)*
(n=5)

*Figure 5. Limitations of AI scribe technology experienced by survey respondents.*




*Q. Please identify any limitations of the AI scribe technology that you experienced. (Check all that apply)*
(n=5)

## Focus Group Findings

Focus groups conducted with staff from the two participating PHUs revealed distinct experiences with the AI scribe technology. Overall, 60% of participants reported positive user experiences, noting that clients were generally willing to provide consent for use of AI scribe. Most SMDHU participants expressed favourable perceptions, while WDGPH participants had fewer applicable cases, which limited their ability to meaningfully test the technology and contributed to greater hesitancy regarding its impact. Time-tracking results reflected a similar pattern across the two units.

Across discussions with both PHUs, four major themes emerged:

1. Positive and promising impact,
2. Flexibility and adaptability of the tool,
3. Building trust in technology, and
4. Changes in practice and culture.

**1. Positive and Promising Impact**

Most participants reported positive impacts after initial adjustment. SMDHU participants reported increased comfort and higher satisfaction with the technology. SMDHU participants cited reduced documentation time, improved work-life balance, and consistent documentation quality. For example, some users noted being able to complete documentation the following morning when working near the end of a shift because the tool retained the transcript for later use.

> **“With the scribe, I could finish a phone call at 4:00 or 4:30, close it all down because I knew the scribe had the transcript of what everybody said and I had my minimal notes, and I could chart the next morning. So I didn’t have to stay extra late to finish off a late phone call, which was different for me.” – Focus Group Participant**

WDGPH participants recognized the tool’s potential, but the limited number of applicable cases prevented them from observing measurable efficiency gains.

> **"I had one case who spoke with a heavy accent and at times it was difficult for me to understand what was being said and I was surprised that the AI scribe picked it up accurately." – Focus Group Participant**

**2. Flexibility and Adaptability of the Tool**

The AI scribe was originally developed for in person clinical settings and therefore required adaptations to meet public health needs. Key challenges included:

- Capturing two-way audio from both investigators and clients during voice calls,
- Ensuring the tool captured both clinical and non-clinical information, and
- Developing templates appropriate for public health documentation.

The two PHUs created custom templates, but this required significant time and coordination. SMDHU users invested heavily in refining templates; however, the technology could not import templates directly. Instead, users deleted the autogenerated template content and pasted their own text. Templates could not be deleted once created, leading to clutter during testing, and templates were not shareable within the platform. Each user needed to manually copy and paste updated templates into their own environment.

Both organizations identified the importance of technical support from the vendor to ensure success in deployment. The AI scribe’s challenges in some instances to capture two-way audio reliably caused substantial delays at the beginning of the pilot. Only a single, headset model was found to be compatible with the system, and procurement delays further reduced the sample size for evaluation. To support consistency across both PHUs, WDGPH and SMDHU shared their template scripts to help standardize documentation practices. WDGPH relied on several workarounds, initially using a speaker-based method to allow participants to operate without a headset. They later transitioned to a web browser extension, which offered improved headset compatibility but introduced new limitations related to template creation and management.

**"I found the smart edit tool on the Scribe to be helpful…it could easily [edit notes to] my personal charting style within the template. It was easy to reorganize information."** – Focus Group Participant

**3. Building Trust in Technology**

Trust in the AI scribe evolved differently between PHUs. SMDHU users were initially hesitant but grew more comfortable as they gained experience with the technology. WDGPH users did not comment extensively on trust, largely because they had limited opportunities to use the technology.

**"I think it was a little bit of learning to trust technology. I know the first couple of investigations I did, I was just keeping my regular notes, writing down every single thing I could, and as I started to trust that it would actually pick up relevant information, I decreased some of the physical notes I was making. So definitely it was a little bit of learning to trust the technology."** – Focus Group Participant

**4. Changes in Practice and Culture**

Participants noted a shift in practice and culture, from manually documenting everything to relying on an AI scribe to capture information. Participants valued the AI scribe's transcript feature, which allowed them to review the conversation after each session. This enabled them to identify any omissions or errors that could have occurred in LLM-generated notes. These types of errors led to users from both PHUs being more rigorous in reviewing generated notes. Most participants also noted that important dates and names of places, such as restaurants, were often transcribed incorrectly, as the AI scribe had difficulty capturing proper nouns accurately.

**"The scribe was very easy to use [and] to build it into the conversation with clients. It was nice to have the transcript so that if you missed any details you could go back and see the full transcript notes."** – Focus Group Participant

Despite differences in their experiences, both PHUs recognized the significant potential and impact of the AI scribe. SMDHU experienced greater benefits, including reduced documentation time and improvements to work–life balance, while WDGPH did not observe similar efficiencies due to limited cases. Both PHUs emphasized the need for better integration, ongoing and initial technical support, and a tool more adaptable to public health workflows.

**"If I had 2 or 3 phone calls back-to-back, I didn't have to worry about documenting this phone call because I would forget what I said in this phone call when I would move onto another phone call. It enabled me to do back-to-back phone calls a little bit more quickly. I know myself, my brain won't remember things if I don't write it down."** **– Focus Group Participant**

# Learnings and Recommendations

As part of the evaluation, several factors were identified that impact the effectiveness, scalability, and long-term sustainability of AI scribe within public health settings. These factors highlight areas requiring attention for prospective users in public health to ensure safe, equitable, and consistent implementation. The following sections summarize key factors observed across governance, operational reliability, technical functionality, and equity considerations, and offers recommendations.

## 1. Governance and Standardization

- The AI scribe technology lacked centralized template version control resulting in each user requiring development of a template within their profile, and any global changes required individual users to recreate and update their templates, often needing the support from a lead at each PHU.
- Fragmentation occurs across disease types due to the absence of standardized templates.
- ID investigations require highly structured, disease-specific information, meaning a single universal template is not feasible. For example, a Salmonellosis template cannot be used for other enteric diseases; each disease requires its own customized template.
- Collectively, these challenges highlight the importance of a more robust template lifecycle governance and standardized practices in future use.

## 2. Integration into PHU Workflows

- The AI scribe can supplement documentation but does not replace existing mandatory documentation systems (e.g., InputHealth, iPHIS, Ontario Investigation Tools). Users must toggle between systems, transferring information from AI-generated notes into required documentation systems.
- As with all generative AI tools, risks such as hallucinations, omissions, and inaccuracies were identified throughout the pilot at different stages of documentation. When using AI scribe technology to generate documentation, ID investigators need to thoroughly review any output prior to transferring the AI-generated notes into required documentation systems. This is in alignment with expectations set out by professional regulatory standards.
- The adoption of the AI scribe for ID investigations required time and adjustment. As this was a new technology for ID investigators, there was an initial learning curve. Building trust in AI scribe, shifting away from traditional note-taking practices, and realizing time savings requires sufficient experience and familiarity with the technology.

- Both PHUs emphasized the need for better integration of the scribed notes with the required documentation systems, and reducing duplicate work. If the AI scribe could populate information directly into the documentation system, users believed that the time savings would be far more substantial. While the AI scribe technology did not fully adapt to the public health context, users developed workarounds that allowed them to benefit from its core functions.

## 3. Technical and Platform Limitations

- There are notable functionality differences between the AI scribes web browser extension and the web application, which vary by vendor, and reduce consistency across users. Audio compatibility was also limited, either requiring the web browser extension or relying on specific headset models that are outdated and no longer available on the market.
- The lack of self-service access controls including role access and account reassignments reduce flexibility for IT and leads when onboarding, offboarding, or reassigning staff.

## 4. Equity, Accessibility, and Change Management

- Users with lower digital literacy may experience challenges adopting and effectively using the AI scribe tool.
- To support change management, having an ID investigator lead the implementation of AI scribe technology is important. This lead role can support template creation as an ID investigation subject matter expert and act as a champion for AI scribe, fostering a culture of innovation and possibly contributing to greater time savings.
- This may present equity, accessibility, and change management considerations that must be addressed in future planning and implementation.

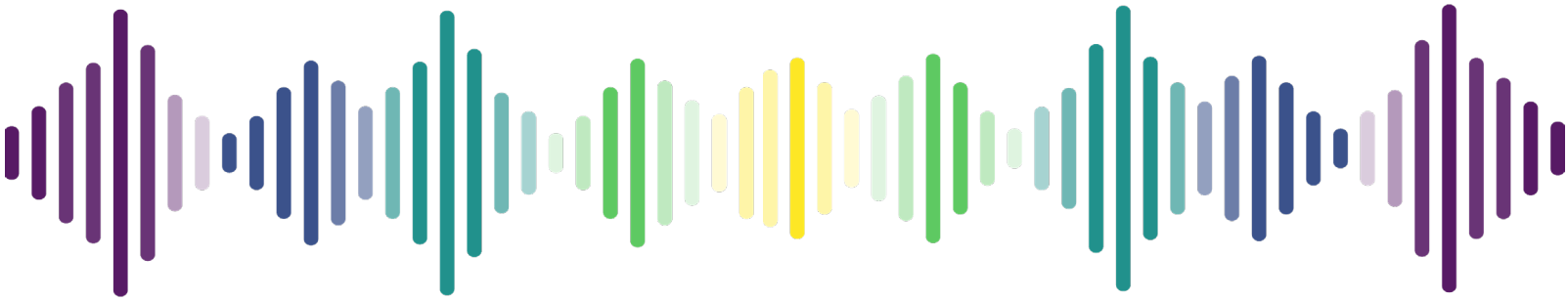
# Study Limitations

The study encountered several limitations. The use of an LLM in the simulation testing could introduce a risk of hallucination errors. The number of cases using AI scribe was limited due to deployment delays attributed to upfront technical, legal, and operational adaptations required. Only 44 interactions with 41 cases involved the AI scribe, which may also reflect seasonal trends and a low volume of eligible cases within selected DOPHS. This led to reduced capacity to fully assess the true impact of the AI scribe. In addition, some DOPHS had no cases or only one case, leading to their exclusion from time-related analysis as time comparisons with and without AI scribe were not available.

Furthermore, the study experienced a low survey response rate and limited participation in focus groups. This may be attributable to the low number of ID investigations, limited experience with the AI scribe, and ongoing interdepartmental movement and staff turnover.

# Conclusion

While initial findings suggest that the AI scribe tool is supportive, the short duration of tool use introduces the need for ongoing pilot testing and evaluation both within Infectious Disease programs and within other public health program areas. Focus group participants noted that they had limited time to work with AI scribe and were only beginning to feel familiar and comfortable with the solution. Participants highlighted that additional time would allow for a more accurate evaluation and clearer understanding of both benefits and limitations.



# References


Bhattacharyya, O., Agarwal, P., Ha, E., Yong, J., & Montague, E. (2025). *Accelerating AI adoption for reducing administrative burden in primary care: Insights from evaluating AI scribes*. HealthcarePapers, 22(4), 63-68.

Bindal, R., Dewangan, P., Walambe, R., Ramanna, S., & Kotecha, K. (2025). *Framework for digital medical scribe technology for healthcare documentation*. Engineered Science, 33, 1377.

Biro, J., Handley, J. L., Cobb, N. K., Kottamasu, V., Collins, J., Krevat, S., & Ratwani, R. M. (2025b). *Accuracy and safety of AI-enabled scribe technology: Instrument validation study*. Journal of Medical Internet Research, 27, e64993.

Biro, J. M., Handley, J. L., Mickler, J., Reddy, S., Kottamasu, V., Ratwani, R. M., & Cobb, N. K. (2025a). *The value of simulation testing for the evaluation of ambient digital scribes: A case report*. Journal of the American Medical Informatics Association, 32(5), 928-931.

Centre for Digital Health Evaluation. (2024). *Clinical evaluation of artificial intelligence and automation technology to reduce administrative burden in primary care*. Women's College Hospital Institute for Health System Solutions and Virtual Care. https://www.ontariomd.ca/documents/ai%20scribe/ai%20scribe%20evaluation_final%20report_vf.pdf

Haniff, Q., Meng, Z., Pongkemmanun, T., Sia, Z. C., Newport, H., Ooi, Y., & Jani, B. D. (2025). *Use of artificial intelligence to transcribe and summarise general practice consultations*. Journal of Medical Artificial Intelligence, 8.

Hassan, H., Zipursky, A. R., Rabbani, N., You, J. G. T., Tse, G., Orenstein, E., ... & Yan, A. P. (2025). *Special topic on burnout: Clinical implementation of artificial intelligence scribes in healthcare: A systematic review*. Applied Clinical Informatics.

Kernberg, A., Gold, J. A., & Mohan, V. (2024). *Using ChatGPT-4 to create structured medical notes from audio recordings of physician–patient encounters: Comparative study*. Journal of Medical Internet Research, 26, e54419.

Luo, X., Zhou, L., Adelgais, K., & Zhang, Z. (2025). *Assessing the effectiveness of automatic speech recognition technology in emergency medicine settings: A comparative study of four AI-powered engines*. Journal of Healthcare Informatics Research, 1-19.

Moryousef, J., Nadesan, P., Uy, M., Matti, D., & Guo, Y. (2025). *Assessing the efficacy and clinical utility of artificial intelligence scribes in urology*. Urology, 196, 12–17.

Ng, J. J. W., Wang, E., Zhou, X., Zhou, K. X., Goh, C. X. L., Sim, G. Z. N., … & Ng, Q. X. (2025). *Evaluating the performance of artificial intelligence-based speech recognition for clinical documentation: A systematic review*. BMC Medical Informatics and Decision Making, 25(1), 236.

Palm, E., Manikantan, A., Pepin, M. E., Mahal, H., & Belwadi, S. S. (2025). *Assessing the quality of AI-generated clinical notes: A validated evaluation of a large language model scribe*. arXiv preprint arXiv:2505.17047.

Sezgin, E., Sirrianni, J., & Kranz, K. (2023). *Development and evaluation of a digital scribe: Conversation summarization pipeline for emergency department counseling sessions towards reducing documentation burden*. medRxiv.

van Buchem, M. M., Kant, I. M., King, L., Kazmaier, J., Steyerberg, E. W., & Bauer, M. P. (2024). *Impact of a digital scribe system on clinical documentation time and quality: Usability study*. JMIR AI, 3(1), e60020.

Wang, H., Yang, R., Alwakeel, M., Kayastha, A., Chowdhury, A., Biro, J. M., … & Hong, C. (2025). *An evaluation framework for ambient digital scribing tools in clinical applications*. npj Digital Medicine, 8(1), 358.

# Acronyms

**AI** – Artificial Intelligence
**CER** – Character Error Rate
**DOPHS** – Diseases of Public Health Significance
**EHR** – Electronic Health Record
**EMR** – Electronic Medical Record
**ID** – Infectious Disease
**iGAS** – Invasive Group A Streptococcal Disease
**LDCP** – Locally Driven Collaborative Project
**LLM** – Large Language Model
**MFIPPA** – Municipal Freedom of Information and Protection of Privacy Act
**PHIPA** – Personal Health Information Protection Act
**PIA** – Privacy Impact Assessment
**PHUs** – Public Health Units
**REB** – Research Ethics Board
**SOAP** – Subjective, Objective, Assessment, Plan (note format)
**SMDHU** – Simcoe Muskoka District Health Unit
**TCPS 2** – Tri-Council Policy Statement (version 2)
**TPH** – Toronto Public Health
**WDGPH** – Wellington-Dufferin-Guelph Public Health
**UWaterloo** – University of Waterloo
**VBD** – Vector Borne Diseases
**WER** – Word Error Rate

# Appendix 1: AI Scribe Technology Procurement Evaluation Matrix (Demonstration Stage)

| Evaluation Criterion | | Weighting | Vendor 1 | Vendor 1 Comments | Vendor 2 | Vendor 2 Comments |
|---|---|---|---|---|---|---|
| Functionality | Generated notes | *15* | | | | |
| | Note types (e.g., SOAP, referral, patient handout, insurance letter, specialties, IHPs) | | | | | |
| | Separation of multiple problems in same visit | | | | | |
| | Generate multiple types of notes from the same transcript | | | | | |
| | Languages supported | | | | | |
| | Dictate edits or additions | | | | | |
| | Note customization (e.g., provide instructions on the note sections, structure, writing style, pronouns) | | | | | |
| | Custom templates | | | | | |
| | Custom dictionary (e.g., user or clinic specific words, names, phrases) | | | | | |
| | Use with video and phone visits | | | | | |
| | Record visit while offline then process when internet access available | | | | | |
| | Able to perform functions in the EMR (e.g., messages, reminders, open forms, referrals, billings, populate CPP)* | | | | | |
| | Pre-appointment patient interview | | | | | |
| Usability | Overall usability (e.g., intuitive, efficient) | *10* | | | | |
| | Window size for editing notes | | | | | |
| | Retrieving notes and transcripts from previous visits | | | | | |
| | Able to open scribe from patient chart * | | | | | |
| | Automatic login (i.e., Don't need to enter username and password) * | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| | Does not require manual entry of patient info (e.g., name, dob) * | | | | | |
| | Able to post note directly to chart in EMR * | | | | | |
| **References** | | *10* | | | | |
| **Privacy and Security** | Onboarding session | *30* | | | | |
| | Phone support | | | | | |
| | Ongoing support and Maintenance | | | | | |
| **Training and Support** | Onboarding session | *10* | | | | |
| | Phone support | | | | | |
| | Ongoing support and Maintenance | | | | | |
| **Cost** | *Annual cost for 15 licences (10 WDG + 5 SMDHU) in top-tier category (pro category) in CAD including any applicable discounts* | *25* | | | | |
| | *Additional fees for each of the following*<br>• *Custom templates*<br>• *Template sharing*<br>• *Integration (EMR)*<br>• *Priority support*<br>• *Other* | | | | | |
| **Total** | | *100* | | | | |

**Associated with the level of integration with the EMR.*

# Appendix 2: Verbal Client Consent Script to Use AI Scribe

I am requesting your consent for the use of a software tool called *[insert AI scribe software name]* to assist me to create notes for this record. Instead of me typing or writing notes as we speak, *[insert AI scribe software name]* records what you and I say in our conversation and uses AI to write a summary. I will review the summary notes created by *[insert AI scribe software name]* and make edits prior to me adding these to your record.

*[insert AI scribe software name]* does not keep the recording of our conversation and *[insert AI scribe software name]* encrypts what they store very carefully to protect your privacy. *[insert AI scribe software name]* will not use any of your personal health information, or the recording of our conversation to train their AI.

There is more detailed information available about how *[insert AI scribe software name]* handles information that I can share if you have questions.

Do I have your consent to use *[insert AI scribe software name]* for this interaction?

**Repeat call to with same client can used a shortened consent**

Based on your consent for me to use an AI scribe to help take notes during our previous conversation, I would like to confirm your consent to use it again this time. Do consent for me to use the AI scribe?

# Appendix 3: Reference Checklist

The following checklist was used to evaluate the quality of documentation generated by the AI scribe. The checklist is disease specific for pertussis and salmonella. Note: this reference checklist is not inclusive of all required information but specific for the scenarios used as part of the simulation testing.

**Pertussis**

Category 1: Exposure

- Exposure Level
- Exposure Name
- Health Unit Responsible
- Earliest Exposure Date
- Most recent Exposure Date (if known)
- Exposure type
- Exposure address
- Exposure setting

Category 2: Client Demographics:

- Patient's Name
- Patient's Date of Birth
- Phone Number

Category 3: Lab

- Specimen Collected (Y/N)
- Specimen Collection Date (If Applicable)
- Specimen results
- Complications (if Applicable)

Category 4: Interventions

- Hospitalization and Treatment Info (if Applicable)
- ER Visit
- Phone call
- Education and Counselling Infor provided by caller

Category 5: Risks

- Immunization History

- Pregnant (Y/N)
- Immunocompromised
- Travel
- Working/Living/Volunteering in High-Risk Environment

Category 6: Symptoms

- Symptoms
- Complications

Category 7: Outcome

- Patient Outcome (Recovered/Residual Effects (specify)/ Fatal (include date of death))

Category 8: Contacts

- Contact tracing necessary (Y/N)
- Name of individuals that need or were contacted (other than callee)
- Intervention of contacts - PEP, Vaccine

**Salmonella**

Category 1: Exposure

- Exposure Level
- Exposure Name
- Health Unit Responsible
- Earliest Exposure Date
- Most Recent Exposure Date (if known)
- Exposure Type
- Exposure Address
- Exposure Setting

Category 2: Client Demographics:

- Patient's Name
- Patient's Date of Birth
- Phone Number

Category 3: Lab

- Specimen Collected (Y/N)
- Specimen Collection Date (If Applicable)
- Specimen results
- Complications (if Applicable)

Category 4: Interventions

- Hospitalization and Treatment Info (if Applicable)
- ER Visit
- Phone call
- Education and Counselling Infor provided by caller

Category 5: Risks

- Pregnant (Y/N)
- Working/Living/Volunteering in High-Risk Environment (Y/N)
- Immunocompromised (Y/N)
- Travel outside of province 7 days prior to onset (Y/N) - if Yes specify place and departure and arrival date
- Consumption of eggs or food containing eggs 7 days prior to onset (Y/N)
- Consumption of shelled eggs 7 days prior to onset (Y/N)
- Consumption of liquid eggs 7 days prior to onset (Y/N)
- Consumption of farm-gate/ungraded eggs 7 days prior to onset (Y/N)
- Consumption of raw or undercooked eggs (runny eggs, over-easy, or in raw cookie dough) 7 days prior to onset (Y/N)
- Consumption of chicken/chicken products 7 days prior to onset (Y/N)
- Consumption of whole chicken/cuts 7 days prior to onset (Y/N)
- Consumption of ground chicken 7 days prior to onset (Y/N)
- Consumption of frozen processed chicken products cooked at home (nuggets, chicken burgers, chicken strips) 7 days prior to onset (Y/N)
- Consumption of shawarma or donair 7 days prior to onset (Y/N)
- Consumption of beef 7 days prior to onset (Y/N)
- Consumption of bird meat - other than chicken (e.g., turkey, ground turkey, Corsish hen, duck) 7 days prior to onset (Y/N)
- Consumption of ground chicken 7 days prior to onset (Y/N)
- Consumption of ready-to-eat pre-washed or pre-made salads 7 days prior to onset (Y/N)
- Consumption of ice cream, gelato and other frozen dairy-based desserts 7 days prior to onset (Y/N)
- Consumption of ground beef 7 days prior to onset (Y/N)
- Consumption of chocolate 7 days prior to onset (Y/N)
- Consumption of fish 7 days prior to onset (Y/N)
- Consumption of raw/unpasteurized milk or milk products 7 days prior to onset (Y/N)
- Consumption of raw vegetables (e.g., rutabaga, cucumber) 7 days prior to onset (Y/N)

- Consumption of sprouts (e.g., bean, alfafa, or other kinds including in sandwich or salad) 7 days prior to onset (Y/N)
- Consumption of other seafood 7 days prior to onset (Y/N)
- Consumption of pork 7 days prior to onset (Y/N)
- Consumption of raw fruits (any melons, pre-cut fruit tray, papayas, mangoes, tomatoes) 7 days prior to onset (Y/N)
- Consumption of seeds, tahini, nuts or nut beer 7 days prior to onset (Y/N)
- Consumption of raw/unpasteurized juice/cider 7 days prior to onset (Y/N)
- Consumption of fresh herbs (e.g., cilantro, basil and parsley) 7 days prior to onset (Y/N)
- Consumption of deli meats (e.g., cold cuts, bologna, salami, pepperoni, and kielbasa) 7 days prior to onset (Y/N)
- Contact with animals (including pets) 7 days prior to onset (Y/N)
- Contact with pet treats/foods 7 days prior to onset (Y/N)
- Contact with reptiles/amphibians or their environment 7 days prior to onset (Y/N)
- Contact with backyard poultry (including chickens, ducks and their environment) 7 days prior to onset (Y/N)
- Swim or contact with water from swimming pools, hot tubs, wading pools or water parks in the province of Ontario, Canada 7 days prior to onset (Y/N)
- Close contact with positive case 7 days prior to onset (Y/N)
- Cross contamination of ready-to-eat foods with raw poultry or meat 7 days prior to onset (Y/N)
- Failure to wash hands properly after handling raw eggs or food containing raw eggs 7 days prior to onset (Y/N)
- Poor hand hygiene (Y/N)

Category 6: Additional Risks

- Is Patient aware of how they became sick? (Y/N)
- Was patient in any specific diet 7 days prior to onset (Y/N)
- Attendance of any special functions (e.g., weddings, parties, showers, family gatherings, child care) 7 days prior to onset (Y/N)

Category 7: Outcome

- Patient Outcome (Recovered/Residual Effects (specify)/ Fatal (include date of death))

Category 8: R/X Treatments

- Antibiotic Prescription - med, duration, route

Category 9: Additional Notes

- Over-the-counter medications (specify if Yes)

# Appendix 4: User Survey

This online survey is a part of the Locally Driven Collaborative Project (LDCP), *Enhancing Public Health Efficiency through AI Scribe Technology: a Pilot Study in Ontario Public Health Units*, being led by the Simcoe Muskoka District Health Unit, Wellington Dufferin Guelph Public Health, Toronto Public Health, and University of Waterloo.

We are inviting you to participate in a survey which aims to gather information on your experience with piloting the AI Scribe technology in case management for Diseases of Public Health Significance (DOPHS), including the identification of any benefits and challenges experienced.

The survey will take approximately 10 to 15 minutes of your time.

The benefit of participating in this survey is that the information shared will help identify best practices and potential barriers to implementation, providing valuable information for scaling the use of AI scribe technology across other public health units. There are no anticipated risks to participating. You will be completing the study by an online survey operated by Medallia Agile Research. When information is transmitted or stored on the internet, privacy cannot be guaranteed. There is always a risk your responses may be intercepted by a third party (e.g., hackers).

Participation in this survey is voluntary. You can choose to end the survey at any time, and you can also choose to skip any question that you do not want to answer. Personal health information cannot be disclosed in the open-ended questions. If PH is inadvertently disclosed, this information will be immediately deleted from the dataset before analysis. Your answers are recorded anonymously and cannot be deleted after you have submitted your responses. The data will be stored in a secure drive, for up to 10 years, with access limited to the LDCP team. Anonymous individual responses will only be seen by LDCP team members assigned to collection and analysis of the data. When we report the results of the survey, answers will be grouped together, and **we will strive to keep your individual responses anonymous; however, given the small number of people being surveyed, we cannot guarantee anonymity. Your identity will be confidential.**

The information you provide, as well as information collected from other survey responses and evaluation methods (i.e., focus groups and program metrics), will be compiled into an evaluation report that will be reviewed by those supporting the LDCP, prior to it being shared with the broader public health community and other relevant interest-holders. Other knowledge translation avenues may also be used to share the evaluation findings, such as journal publications, webinars, and conference presentations.

This study has been reviewed and received ethics clearance through the University of Waterloo Research Ethics Board (REB #47284). If you have questions for the REB, contact the Office of Research Ethics, toll-free at 1-833-643-2379 (Canada and USA), 1-519-888-4440, or reb@uwaterloo.ca.

If you have any questions about this survey, please contact Natalie Riewe Natalie.Reiwe@smdhu.org.

The survey will be open until 5 pm on **[insert date]**.

By agreeing to participate in the study you are not waiving your legal rights or releasing the research investigator(s) or involved institution(s) from their legal and professional responsibilities.

I agree to participate in this study.

- o Yes
- o No (skip to thank you page)

I agree to the use of de-identified quotations in published findings.

- o Yes
- o No

## Demographics

1. **Which public health unit are you employed with?**
   - o Simcoe Muskoka District Health Unit
   - o Wellington Dufferin Guelph Public Health

## Experiences using AI Scribe

2. **Please rate each of the following statements.**

| | Strongly Disagree | Disagree | Neither Disagree nor Agree | Agree | Strongly Agree | Please explain |
|---|---|---|---|---|---|---|
| **AI scribe was easy to use.** | | | | | | |
| **The AI scribe interface was user friendly to navigate.** | | | | | | |
| **AI scribe worked well with other systems running (e.g., iPHIS, InputHealth).** | | | | | | |
| **AI scribe made documentation process more time efficient.** | | | | | | |
| **The AI scribe transcript was useful.** | | | | | | |
| **AI scribe documentation was accurate.** | | | | | | |
| **AI scribe improved quality of documentation.** | | | | | | |
| **Using the AI scribe supported my assessment.** | | | | | | |
| **AI scribe integration into the EMR was effective (e.g., iPHIS, InputHealth).** | | | | | | |
| **I was comfortable asking for consent to use AI scribe.** | | | | | | |

3. Please rate your overall experience using AI scribe.
   - o Poor (please explain): _____
   - o Fair (please explain): _____
   - o Good
   - o Very good
   - o Excellent

## Documentation Time

4. **Compared to the conventional documentation method, what was the impact of AI scribe on time spent completing documentation?**
   - o AI scribe decreased documentation time
   - o No change in documentation time
   - o AI scribe increased documentation time

## Benefits

5. **Please identify any benefits you experienced in implementing AI scribe technology. (Check all that apply)**
   - □ More efficient automated process/workflow
   - □ Reduced my administrative burden
   - □ Improved my job satisfaction
   - □ Reduced my cognitive load during encounters with clients
   - □ I was able to be more engaged with clients
   - □ Other (please specify): _____

## Limitations and Challenges

6. **Please identify any limitations of the AI scribe technology that you experienced. (Check all that apply)**
   - □ Creation of concise investigation notes
   - □ Difficulty distinguishing between multiple speakers
   - □ Difficulty interpreting multilingual conversations
   - □ Difficulty interpreting conversations where speaker had limited English proficiency or struggles with enunciation
   - □ Capturing nuances in client speech and tone
   - □ Other (please specify): _____

7. **Please identify any barriers and/or challenges you experienced in implementing AI scribe technology. (e.g., workflow challenges)**

## Additional Feedback

8. **Please provide any additional feedback regarding your experience implementing AI scribe for case management.**

**Thank you for your participation**

# Appendix 5: Focus Group Facilitation Guide

**Information Letter/Consent**

The focus group will be facilitated by Nauman Shakeel, a graduate student in the Master of Public Health Science Program at the University of Waterloo, Casey Hirschfeld, an Evaluation Specialist at Simcoe Muskoka District Health Unit, and Ravi Shah, a Health Data Analyst [or designate] at Wellington-Dufferin-Guelph Public Health.

This focus group is part of the Locally Driven Collaborative Project (LDCP), *Enhancing Public Health Efficiency through AI Scribe Technology: a Pilot Study in Ontario Public Health Units*, being led by the Simcoe Muskoka District Health Unit, Wellington Dufferin Guelph Public Health, Toronto Public Health, and University of Waterloo.

The purpose of this focus group is to collect information on your experience with piloting the AI Scribe technology in case management for Diseases of Public Health Significance (DOPHS), including the identification of any benefits and challenges experienced, as part of a research project.

The focus group will take approximately 60 to 90 minutes of your time. Participation is completely voluntary.

The focus group will be conducted over an online platform, Microsoft Teams. Microsoft Teams has implemented technical, administrative, and physical safeguards to protect the information provided via the services from loss, misuse, and unauthorized access, disclosure, alteration, or destruction. However, no Internet transmission is ever fully secure or error free. With your permission, I would like to record this focus group. Although an LDCP team member will be taking notes throughout our focus group, having a recording allows us to ensure that the notes are complete. Personal health information cannot be disclosed during a recorded session. If any personal health information is inadvertently disclosed the recording will be stopped and deleted. The recording will be transferred from OneDrive to a secure drive where a member of the LDCP team will play it back and complete the interview notes; only the research team members analyzing the data will have access to the recording. Following finalization of the report, the recordings will be deleted from the secure network folder. We will give you the opportunity to review the notes before we analyze them. After the report has been completed, the recording will be deleted from the secure drive. All notes and analysis files will be stored in a secure network drive, for up to 10 years, with access limited to the LDCP team.

The benefit of participating in this focus group is that the information shared will help identify best practices and potential barriers to implementation, providing valuable information for scaling the use of AI scribe technology for other Diseases of Public Health Significance and across other public health units. I do not anticipate any major risks to participating; if, however, there are questions which you do not wish to answer, you may let me know.

We will not ask you any personal or identifying questions and no responses will be attributed to any one particular individual. Any direct quotes used in the report will be done so anonymously and with your permission. **We will strive to keep your individual responses anonymous; however, given the small number of people participating in focus groups, we cannot guarantee anonymity. Your identity will be confidential.** Given the group format of this session, the research team and other participants in the focus group will know what you

said. We will ask you to keep in confidence information that identifies or could potentially identify a participant and/or their comments.

The information you provide, as well as information collected from other staff participating in focus groups and other evaluation methods (i.e., survey and program metrics), will be compiled into an evaluation report that will be reviewed by those supporting the LDCP, prior to its being shared with the broader public health community and other relevant interest-holders. Other knowledge translation avenues may also be used to share the evaluation findings, such as journal publications, webinars, and conference presentations.

If you withdraw your consent after the focus group has been completed, we cannot delete your record since the notes will be recorded anonymously and may represent the opinion of more than one individual participating in the focus group.

This study has been reviewed and received ethics clearance through the University of Waterloo Research Ethics Board (REB #47284). If you have questions for the REB, contact the Office of Research Ethics, toll-free at 1-833-643-2379 (Canada and USA), 1-519-888-4440, or reb@uwaterloo.ca.

If you have any questions about this interview, please contact casey.hirschfeld@smdhu.org [for SMDHU participants] or Ravi.Shah@wdgpublichealth.ca [for WDGPH participants].

[Individual consents will be obtained via email in advance of the focus group taking place]

1. **Do you agree to participate in the focus group based on the information that I have just shared with you?**
   ☐ Yes ☐ No
2. **Do I have your permission to record the focus group?**
   ☐ Yes ☐ No

*Note: If an individual chooses to decline being recorded, they will be unable to participate in the focus group. An interview will be offered to support participation.*

3. **Do I have your permission to use de-identified quotations in the evaluation report and journal paper?**
   ☐ Yes ☐ No
4. **Do you have any questions for me before we get started?**

**Focus Group Questions and Probes**

**Context/Experience**

- Can you please describe your experience using the AI scribe technology for case management?

**Benefits**

- What benefits did you experience in implementing the AI scribe technology for case management?
    - Probe: What were the positive impacts of the technology on workflow, time spent completing documentation, job satisfaction, client engagement, technology features, etc.?
- What supports do you feel were helpful in implementing the AI scribe technology?

**Challenges**

- What challenges did you experience in implementing the AI scribe technology for case management?
    - Probe: What were the negative impacts of the technology on workflow, time spent completing documentation, job satisfaction, client engagement, technology features, etc.?
    - Probe: What barriers did you experience?
- What supports, if any, could have been better for implementing the AI scribe technology?
    - Probe: Were any supports missing that would have been helpful?

**Lessons Learned/Opportunities**

- Reflecting on the benefits and challenges, what were the lessons learned?
    - Probes: review benefits and challenges noted by participants.
- What additional features or capabilities would you like to see in AI scribe technology to work efficiently?
- How can we leverage what we have learned to enhance the scalability of AI scribe technology for case management?
- How can AI scribe technology's scalability be enhanced for case management? (e.g., use for other DOPHS and/or at other PHUs)?

**Closing**

Once the notes have been validated, you will receive a copy to review to ensure all applicable information was captured from an LDCP team member. Although not mandatory, you will be given one week to review the notes if you choose.

**Thank you very much for your time.**