OntarioMD

Empowered Practices. Enhanced Care.

AI Scribe Evaluation Report

based on WELL AI Voice

March 2024

Contents

EXECUTIVE SUMMARY

Introduction

OntarioMD conducted an evaluation study of Artificial Intelligence (AI) medical scribe to explore the potential benefits and challenges of implementing AI scribe technology in clinical settings. Using the WELL AI Voice solution, this initiative sought to assess the impact of this representative AI scribe on improving physician efficiency, reducing administrative burdens, and enhancing patient care experiences.

Objectives

The study examined the AI Scribe solution in three main areas:

- User Experience and Physician Burnout: Investigate how the AI Scribe integrates into clinical workflows, its ease of use, accuracy in documenting clinical conversations, and its impact on reducing documentation-related physician burnout.
- Technical Performance: Assess the AI Scribe's documentation accuracy, processing speed, and adherence to data security, privacy, and confidentiality standards.
- Patient Perspective: Analyze patient consent, comfort levels, satisfaction, and the quality of interactions during consultations involving the AI Scribe.

Key Findings

The AI Scribe implementation significantly impacted physician workflow, notably reducing documentation time by an average of 3.4 minutes per visit and alleviating cognitive load by 85%. Physician satisfaction with the AI scribe was high, with 81% satisfaction reported, 88% recognizing its role in workload redistribution, and 50% observing a decrease in after-hours documentation. In terms of technical performance and clinical accuracy, the AI scribe achieved a 95% accuracy rate in transcribing complex medical terminology. This precision was acknowledged by 75% of physicians, who appreciated the AI scribe's capability to capture key elements of consultations, thereby enhancing patient care. From the patient's perspective, the use of AI Scribe during consultations led to a significant increase in comfort (92%) and satisfaction (85%). Furthermore, 78% of interactions were perceived to strengthen the physician-patient relationship.

Limitations

The study's limitations, such as its small, non-random sample size, potential biases (including selection and response biases), and the incentivization method for participation, challenge the generalizability of its findings. Future phases plan to recruit a broader, more diverse group of physicians to address these issues using more stringent research designs and methods. Additionally, employing qualitative research methods will offer deeper insights into the AI scribe's impact on clinical efficiency and time savings.

Conclusion

The OntarioMD evaluation of WELL AI Voice represents an initial exploration into the impact of AI scribe technologies in modern clinical practice. This evaluation study's findings are promising, demonstrating the AI scribe's strong capability to accurately transcribe clinical conversations, particularly in capturing complex medical terminology. Additionally, the technology has shown potential in reducing the time and cognitive load associated with documentation, suggesting it could play a role in mitigating physician burnout. Patients, too, have reported high comfort and satisfaction during consultations involving AI scribes.

Despite these positive outcomes, the results should be approached with caution. They underscore the importance of further research to validate and expand upon these preliminary findings. The benefits observed in clinical practice, patient experience, and physician well-being signal a need for studies with larger, more diverse samples and rigorous methodological approaches. The future of AI scribes in healthcare holds promise, yet realizing their full potential and ensuring their successful integration into clinical settings necessitates a balanced consideration of the advancements in technology and the limitations inherent in this study.

INTRODUCTION/BACKGROUND

OntarioMD (OMD) is leading the efficient and scalable use of digital health solutions for primary care. We are committed to improving patient care and reducing physicians' administrative burden by effectively supporting the implementation of innovative technologies. A key area of our focus is understanding the use of artificial intelligence (AI) in clinical settings, particularly clinical ambient AI scribes ("AI scribes"). As an emerging technology, AI scribes have the potential to improve physician workflow, enhance patient-physician interactions, and mitigate physician burnout.

The Ontario Medical Association (OMA) has identified physician burnout as a significant issue within the health-care sector, characterized by exhaustion, depersonalization, and a perceived lack of personal accomplishment. It not only adversely impacts the well-being of physicians but also potentially impacts patient care, safety, and overall health-care delivery. Technological advancements, particularly in the realm of AI, present promising solutions to this pervasive problem. AI scribes, designed to alleviate the administrative burden and enhance the efficiency of physicians, are in the vanguard of these solutions and are one of the recommended solutions in the OMA's *Prescription for Ontario*.

AI scribes function by digitally transcribing patient-physician interactions, reducing the time physicians spend on data entry and documentation - a key contributor to burnout. By automating this process, physicians can focus more on patient care, improving their work satisfaction and well-being. However, successful AI scribes integration and use depend on user experience, clinical relevance, and technical performance. It is crucial to evaluate these factors to ensure the technology aligns with the needs and workflows of physicians.

OMD and the OMA recognize the opportunity AI scribes offer in helping transform health-care delivery and improve physician well-being. To this end, OMD designed and conducted an evaluation of the effectiveness of clinical AI scribes in reducing physician burnout from various perspectives with support from the OMA. Our primary objective in this preliminary pilot was to assess the impact of 'WELL AI Voice' (developed by Tali AI) on physician efficiency. Participation by WELL Health allowed OMD to conduct an initial examination of the potential benefits and pinpoint areas for refinement using AI scribes. This collaboration marked a significant step in our evaluation process, illuminating the potential advantages and areas for improvement in using AI scribes.

Furthermore, OMD is strategically collaborating with other public and private entities with the goal of conducting a more extensive analysis of AI scribe solutions used in primary health-care settings in Canada. The groundwork laid by the preliminary evaluation is expected to contribute invaluable insights that will shape the design and implementation of the pilot study to follow. Our ultimate aim remains constant - to enhance the well-being of physicians and the quality of patient care.

OBJECTIVES OF THE EVALUATION

The evaluation of AI Scribe, leveraging the WELL AI Voice solution, was a strategically planned initiative aimed at quantifying the impact of AI technology on alleviating physician burnout and augmenting clinical efficiency. The objective is to assess the AI scribe's capabilities and effects within these three areas:

- **User Experience and Impact on Physician Burnout:**
  - **Integration and Usability**: Examining how seamlessly 'WELL AI Voice' integrates into physicians' workflows, focusing on its user interface and overall design to determine if it aligns with the ease of use required in fast-paced clinical environments.
  - **Clinical Relevance**: Investigating the ability of 'WELL AI Voice' to accurately interpret and document clinical conversations and whether it adequately supports the diverse lexicon used in medical practices.
  - **Impact on Burnout**: Evaluating how 'WELL AI Voice' reduces physicians' time and effort on documentation tasks, potentially lessening factors contributing to burnout.
- **Evaluation of Technical Performance:**
  - **Performance Metrics**: Assessing the accuracy and speed of 'WELL AI Voice' in creating clinical documentation and its ability to effectively transcribe across various accents, speech patterns, and clinical environments.
  - **Data Security**: Ensuring that 'WELL AI Voice' adheres to the stringent requirements for patient data privacy and security, maintaining confidentiality and data residency.
- **Patient Perspective Evaluation:**
  - **Patient Consent and Comfort**: Analyzing patients' comfort with the use of 'WELL AI Voice' in clinical consultations and their willingness to consent to its operation during medical visits.
  - **Satisfaction and Interaction Quality**: Gauging patient satisfaction with the quality of care and the clarity of communication in consultations where 'WELL AI Voice' is used.
  - **Effect on Physician-Patient Relationship**: Assessing the influence of 'WELL AI Voice' on the dynamics of the physician-patient relationship.

ROLES & RESPONSIBILITIES

OntarioMD (OMD)

OMD took the lead role in conducting evaluation activities and was responsible for:

- **Assessment Metric Development**: OMD created an inclusive assessment metric that captures various aspects of the AI scribe solution, such as user experience, clinical relevance, and technical performance.

- **Physician Recruitment**: WELL Health contacted potential candidates expressing their willingness to participate in the evaluation process, and OMD selected physicians who met the inclusion criteria. The aim was to ensure a diverse and representative sample.
- **Supervised Onboarding:** OMD closely supervised the process, ensuring the solution was installed correctly and configured for each participant's needs. Inclusive training was provided to pilot participants for effective use of the AI scribe technology.
- **Legal review, contract development:** OMD negotiated the pilot agreement with WELL Health and conducted a legal review of documents being presented to physicians (including the terms of use between the physicians and WELL Health). OMD considered privacy issues and risks and provided feedback on the privacy policy and other communications, including the 'Frequently Asked Questions' (FAQs) document to help guide the participants in the evaluation. OntarioMD also provided input on the consent process and forms used for purposes of the Pilot. OMD met with vendors to better understand the legal and privacy considerations.
- **Pilot Evaluation**: OMD managed and monitored the evaluation process in the selected clinics, collecting data and insights to assess the AI scribe solution's effectiveness in mitigating physician burnout.

Ontario Medical Association (OMA)
Sponsor and partner in this evaluation, providing financial support and offering its insights and expertise to ensure a well-rounded evaluation of the AI scribe solution.

WELL Health (AI solution vendor)
WELL Health, the AI solution vendor, played a vital role in this pilot to ensure the successful implementation and use of the AI scribe solution ('WELL AI Voice') during the evaluation process. It was responsible for:

- **Deployment**: The vendor was responsible for ensuring that the solution was appropriately installed and configured to meet the specific needs of the clinics.
- **Training**: The vendor provided pilot participants with thorough training on the AI scribe solution, ensuring they had the necessary knowledge and skills to use the technology effectively.
- **Ongoing support and troubleshooting**: WELL Health provided continuous support and troubleshooting for pilot participants throughout the evaluation process, addressing any technical issues or challenges and ensuring the smooth operation of the AI scribe solution in the participating clinics.

EVALUATION FRAMEWORK AND METHODOLOGY

Evaluation Criteria

- **Clinical Relevance**

- Accuracy in Clinical Conversations: Determine the AI scribe's ability to accurately transcribe and interpret patient-physician dialogues, including comprehending clinical terminology and context.
- Comprehensiveness: Assess whether the AI scribe captures the full scope of clinical conversations, including details like symptoms, diagnoses, medications, and treatment plans.
- Legal and privacy compliance and risk. Clear outline of roles and responsibilities.

- **Usability**
  - Usability: Evaluate the user interface and overall design of the AI scribe for ease of use, especially in fast-paced clinical environments.
- **Impact on Physician Burnout**
  - Reduction in Documentation Tasks: Evaluate how the AI scribe influences the time and effort physicians spend on documentation, focusing on its potential to alleviate factors contributing to burnout.
  - Cognitive Load: Measure the AI scribe's effect on reducing the cognitive burden of recalling and documenting patient information.
- **Technical Performance**
  - Transcription Accuracy and Speed: Assess the accuracy and timeliness of the AI scribe in creating clinical documentation, including its adaptability to different accents, speech patterns, and clinical environments.
  - Data Security and Privacy: Verify compliance with privacy laws, ensuring patient data confidentiality and residency.
- **Patient Perspective**
  - Consent and Comfort: Evaluate patients' comfort with and consent for the use of 'WELL AI Voice' during clinical consultations.
  - Satisfaction and Interaction Quality: Gauge patient satisfaction regarding the quality of care and clarity of communication when the AI scribe is used.
  - Impact on Physician-Patient Relationship: Assess the influence of the AI scribe on the dynamics of the physician-patient relationship.
- **Scalability and Support**
  - Scalability: Determine the system's ability to handle increasing volumes of patient data and conversations.
  - Vendor Support: Evaluate the quality of vendor support, including problem-solving efficiency, training provision, and ongoing assistance.
  - User Feedback: Collect and analyze user feedback regarding ease of use and satisfaction with the AI scribe's features and performance.

Evaluation Methodology

Data were gathered through a combination of stakeholders' surveys and forms:

- Technical questionnaire for the AI scribe vendor
- Physicians' survey (online)

- Patient survey (paper)
- Time trackers (online and paper formats)

- **Technical Assessment - Product Information (Appendix 1):** This questionnaire was designed to collect the relevant product information to facilitate OMD's assessment of the AI scribe's various technical and functional capabilities, including its integration with existing systems, overall functionality, and the vendor's reputation and reliability.
- **Physicians Participants Survey (Appendix 2):** The Physician's Survey was crafted to gather valuable feedback directly from the physicians using the AI scribe. Its primary purpose is to understand their experiences with the AI scribe, focusing on its impact on their daily clinical practice, documentation efficiency, and patient interactions. This survey elucidates how the AI scribe influences clinical workflows and physician-patient dynamics.
- **Patients Participants Survey (Appendix 3):** This survey captured the patients' perspectives, concentrating on their experience and comfort level with the AI scribe during medical visits. It seeks to understand the AI scribe's impact on patient interactions, satisfaction with care, and overall acceptance of this technology. The survey aims to ensure that the AI scribe's implementation aligns with patient comfort and consent, enhancing the quality of care without compromising patient experience.
- **Time Motion Study:** Including a Time Motion Study in our evaluation provided a quantitative analysis of the AI scribe's impact on clinical efficiency. This study aims to measure the time physicians spend on patient visits and documentation, both before and after implementing the AI solution. By comparing these timeframes, the study seeks to identify any significant changes in efficiency and workload management.

Sampling and Participants

- **Physicians as Pilot Participants**: physicians engaged with the 'WELL AI Voice' AI scribe during routine consultations. Their interactions provided critical insights into the scribe's usability, its impact on their workload, and its effect on patient interactions. These physicians, selected by the WELL Health team, participated in training sessions and subsequently provided valuable feedback through surveys.
    - The recruitment for the evaluation of 'WELL AI Voice' was a collaborative effort between the WELL Health and OMD teams. The WELL Health team was crucial in identifying and recruiting physicians and selecting suitable clinics and physicians to create a diverse and representative sample. They also managed the deployment of the AI scribe solution and provided training and ongoing support. Concurrently, the OMD team oversaw the recruitment and onboarding processes, ensuring adherence to the inclusion criteria and collaborating in training the participants. OMD was responsible for conducting the pilot evaluation, managing data collection, and monitoring the services provided. The OMD team also analyzed the data and compiled the findings into this report.
- **Patient as Pilot Participants**: Patients who visited physicians using the AI scribe during their consultation completed a brief survey post-visit. Their feedback regarding comfort with the AI scribe,

the consent process, and overall satisfaction with 'WELL AI Voice' played a significant role in our evaluation.

All participants were thoroughly briefed on their roles and responsibilities at the start of the evaluation process, ensuring clarity and engagement throughout the study.

Inclusion Criteria

The selection of participants for the evaluation of the AI scribe solution ('WELL AI Voice') was based on:

- **Clinician Types:** We focused exclusively on family physicians practicing in Ontario.
- **Number of Participants**: Our target was to recruit 30 family physicians.
- **Participant Diversity**: We sought a diverse group of physicians to ensure a thorough assessment. This diversity encompassed:
  - **Payment Models**: Physicians practice under different models, including Fee-for-Service (FFS), or Family Health Group (FHG), and Family Health Organization (FHO).
  - **Gender Representation**: Both male and female physicians were included.
  - **Experience Levels**: Physicians at various career stages, from early to mid and late practice, were part of the participant pool.
  - **Speech Variations**: A range of accents was included to test the AI scribe's adaptability to different speech patterns and dialects.
- **Number of Encounters**: Each participating physician was required to provide data from at least 25 patient encounters. This threshold ensured a sufficient amount of data for meaningful analysis.
- **Use Cases**: The evaluation was centred on in-person patient visits with family physicians. We prioritized complex clinical cases - involving more than one health issue - over simpler ones to assess better the AI scribe's capability in handling diverse and challenging clinical scenarios.

Evaluation Procedures

The evaluation process was executed in distinct phases:

- **Pre-Pilot Phase:** The initial phase involved differentiating the contributions of both Tali AI (as the software developer) and WELL Health (as the vendor). This consideration was important when reviewing documents, such as the privacy policy and the FAQs, to guide participant contributions to the pilot and to enhance their understanding. Extensive negotiations took place during this phase, covering the main agreement between OMD and WELL Health to operate the pilot, as well as the terms of service agreement between WELL Health and the participants. Once these documents were finalized, foundational data regarding the technical features and functionality of the AI scribe were collected from WELL Health. Data about the time consumed for patient visits and typing a SOAP note into the EMR was also captured during this stage (Pre-Implementation Time Motion Study phase).
- **In-Pilot Phase:** In this phase, physicians used the AI scribe alongside preliminary training. The time for patient visits and writing a SOAP note on the EMR using the AI scribe were documented (Post-

Implementation Time Motion Study phase). Additionally, patients were asked to complete a brief questionnaire after their visits (patient's survey).

- **Post-Pilot Phase:** After the pilot was completed, users were asked to fill out a detailed survey to provide feedback about their experiences using the AI scribe.

Data Analysis

The analysis employs quantitative and qualitative methods, synthesizing performance metrics, user feedback, and trend analyses to understand the AI scribe's effectiveness.

- **Quantitative Analysis**: Quantitative data, primarily from performance metrics and user surveys, was analyzed to identify potential patterns, associations, and trends. This involved comparing the pre- and post-implementation time tracking data to identify improvements in physician efficiency and reduced workload, aligning with our objectives to alleviate physician burnout and enhance clinical operations.
- **Qualitative Analysis**: Qualitative data derived from user feedback in the surveys was carefully reviewed and thematically analyzed to identify overarching themes and insights. The focus was on understanding user experiences, including satisfaction levels, ease of integration into clinical workflows, and the AI scribe's impact on physician-patient interactions. Responses were examined for feedback on the AI scribe's usability, clinical relevance, and adaptability to diverse speech patterns and clinical environments. This qualitative analysis was crucial in assessing how well the AI scribe met the diverse needs of the participants, as highlighted in our inclusion criteria.

RESULTS

Descriptive Analysis: Physician Recruitment and Data Collection

The recruitment phase was a foundational step in our evaluation, targeting a specific segment of physicians to assess the impact of the 'WELL AI Voice' scribe. We reached out to 207 WELL Health EMR family physician users across Ontario, leveraging the customer relationship management (CRM) systems of WELL Health and OMD for effective communication and data extraction.

- From the total physicians contacted, there was a response rate of 15.5%, with 32 physicians agreeing to be part of the evaluation. 79.7% (165 physicians) declined to participate, while a smaller subset of 4.8% (10 physicians) withdrew during the process.
- There was balanced gender representation in the study, with 17 female physicians (53.1%) and 15 male physicians (46.9%).
- A broad spectrum of experience levels was represented in the study, ranging from three to 45 years of practice.
- Over two-thirds of physicians (68.8% or 22 physicians) were part of a Family Health Organization (FHO). The remainder (31.2% or 10 physicians) were part of a Family Health Group (FHG), which included those operating under a Fee-for-Service (FFS) model.

Lessons Learned and Mitigating Steps

Throughout the recruitment and onboarding process for the evaluation study, we encountered several challenges that provided valuable lessons for future endeavours. The barriers faced and the measures taken to address them were:

- **Recruitment Delays** due to various unforeseen circumstances were noted, leading to an extended recruitment period. Collaborative efforts with OMD Advisors enhanced follow-up efficiency with potential participants. A tracking spreadsheet was developed to monitor progress. However, we recognized the need for a more robust approach, highlighting the importance of an established onboarding process and tools for comprehensive pilot management.
- **Technical Barriers** such as firewall restrictions at clinics presented initial deployment obstacles. The WELL Health team's proactive approach provided clinic administrators with complete whitelisting site lists required for the AI solution, streamlining future deployments and reducing resolution times.
- **Legal and Regulatory Considerations**. In addressing the legal and regulatory considerations, the primary focus was on establishing accountability for information collected (as between physician and vendor) and with respect to establishing limits on the collection, use and disclosure of personal information (PI) and personal health information (PHI). Other important considerations were to ensure the responsible use of PI and PHI by maintaining the privacy and security of such data in compliance with the *Personal Health Information Protection Act*, 2004 (PHIPA). Issues considered included the following:
    - **a)** whether PI or PHI could be processed or stored outside Canada;
    - **b)** requiring subcontractors to apply the same responsibility and safeguards to personal health information;
    - **c)** restrictions on secondary use of transcripts and summary notes;
    - **d)** destroying PI and PHI upon termination or expiration of the pilot;
    - **e)** encryption of data at rest and in transit;
    - **f)** restrictions on the disclosure of PI and PHI to the minimum necessary, disclosing only what is required and when it is necessary only;
    - **g)** de-identifying transcripts and summarized notes before processing and use; and
    - **h)** require notification of any unauthorized access, use, or disclosure of any data.
- **Time Constraints** emerged as a significant concern for participants who perceived the training and evaluation process as time intensive. To mitigate this, we introduced pre-recorded, self-serve training options, allowing physicians to engage at their convenience. Additionally, supplementary materials like "how-to" guides were distributed, aiding in retaining and referencing information from training sessions.
- **Determination of Roles and Responsibilities is an area in need of clarification.** Beyond the legal and regulatory considerations mentioned earlier, OMD also had to take into account contractual considerations. There appears some uncertainty among physicians and vendors regarding their roles and responsibilities in offering an AI scribe tool. Multiple vendors or parties may be involved in offering a solution and there is a need for clarity with respect to the status of these vendors vis-à-vis the physician. Clear delineation of roles and responsibilities, including privacy compliance and consent

procedures, is essential. Vendor terms, conditions, and privacy policies must be easily understood to establish a transparent line of accountability. These considerations were integral to the contract negotiations, leading to the identification of accountability between the AI software developer, the vendor, and the physician. It was necessary to evaluate and enhance patient consent processes, initially identified as time-consuming. In response, we provided concise and clear language materials to expedite the process for patients. Another outcome involved requiring physicians to sign a participation agreement with OMD, allowing us to utilize participants' confidential information for pilot purposes. This streamlined approach will likely remain an important priority in efficient rollout out and adherence to regulatory compliance.

- **Physician Availability** proved challenging as many physicians were too busy or absent during the evaluation period. To accommodate their schedules, we extended deadlines and adopted a rolling schedule. This approach and widening the recruitment and onboarding window allowed for a better fit with physician schedules and increased participation.
- **Supported Languages** offered by the AI scribe were an issue for clinics serving primarily non-English speaking patients (e.g., Chinese, Spanish and Persian). This feedback will be vital for the AI solution vendors to consider in future updates to ensure inclusivity and functionality in multilingual settings.

These lessons have been instrumental in refining our approach to deploying innovative health-care solutions. They have prompted us to establish more flexible and inclusive processes that can adapt to the diverse and dynamic environments in which physicians operate. Moving forward, these insights will guide the structuring of evaluation frameworks to be more physician-centric, ensuring higher engagement and smoother integration into their workflows.

Quantitative Analysis

Physician Survey Results:

In analyzing the physician survey data, questions were grouped by themes to provide a more nuanced understanding of the AI scribe's functionality and its impact on clinical practice:

Theme #1. Clinical Workload Management and Burnout

The theme of "Clinical Workload Management and Burnout" relates to the following subthemes, each represented by a survey question: perceived impact of the AI scribe on physicians' workload distribution; cognitive load during patient visits; time savings in documentation; and the overall effect on timely clinical note-taking. Overall, most respondents (50% and 35% responded "Strongly Agree" and "Agree," respectively) reported benefits in terms of reductions in cognitive load and administrative burden, as well as time savings (Figure 1).

- **Reducing Cognitive Load**
  - Most respondents (58% and 27% responded "Strongly Agree" and "Agree," respectively) agreed that the AI scribe reduced their cognitive load by minimizing the need to recall and document

information from memory. This is key in reducing burnout as physicians can focus more on the consultation and less on administrative tasks such as documentation.

- **Time Savings in Documentation**
  - Most respondents (54% and 31% responded "Strongly Agree" and "Agree," respectively) agreed that using the AI scribe saved time when documenting patient encounters compared to manual methods. This perceived time-saving is crucial in a fast-paced clinical environment where every minute counts, leading to fewer overtime hours and fewer rushed patient visits.
- **Timely Clinical Note-Taking**
  - Most respondents (42% and 38% responded “Strongly Agree” and “Agree,” respectively) agreed that the AI scribe transcribed patient-physician conversations in a timely manner. Quick and accurate transcription is essential for maintaining up-to-date patient records and ensuring no critical information is lost or delayed.
- **Capacity creation**
  - Most respondents (46% and 42% responded "Strongly Agree" and "Agree," respectively) agreed that the AI scribe helped redistribute their workload. According to these physicians, the tool alleviated administrative burdens, potentially freeing time for direct patient care.

Overall Analysis

Overall, the AI scribe appeared well-received by physicians in managing their workload and reducing the time and cognitive demands associated with documentation. The high percentage of agreement across questions within this theme suggests that the AI scribe effectively supported the respondents in these areas.

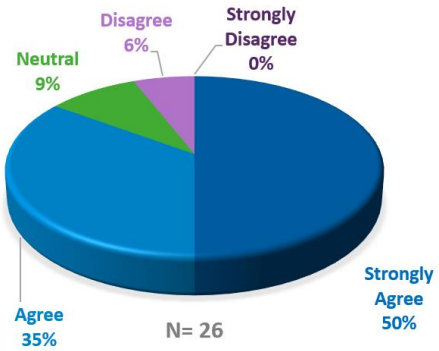

Figure 1 Clinical Workload Management and Burnout

Theme #2. AI Scribe Functionality

The theme of "AI Scribe Functionality" relates to the following subthemes, each represented by a survey question: overall satisfaction with the AI scribe’s functions and features; the AI scribe's adaptability to different accents; the AI scribe’s ability to capture the full scope of conversations comprehensively; and the AI scribe’s comprehension of clinical terminology and context. Overall, most respondents (80% responded

"Strongly Agree" or "Agree") seemed satisfied with the functions and features of the AI scribe (Figure 2). However, some respondents (17%) remained neutral, which could indicate indecision, ambivalence, or a less-than-satisfactory experience.

- **Performance Satisfaction**
  - Most respondents (81% responded "Strongly Agree" or "Agree") reported satisfaction with the AI scribe's performance in terms of functions and features. 20% of respondents remained neutral.
- **Adaptability to Different Accents**
  - Most respondents (85% responded "Strongly Agree" or "Agree") agreed that the AI scribe adapted well to different accents. This is significant for a diverse linguistic environment and suggests that the AI scribe's voice recognition algorithms are robust, at least among the study respondents.
- **Capturing the Full Scope of a Conversation**
  - 69% of physicians agreed (23% neutral; 8% disagreed) that the AI scribe comprehensively captured the full scope of conversations. This represented the lowest percentage of agreement, suggesting challenges in capturing nuances or complexities within clinical dialogue among study respondents.
- **Comprehension of Clinical Terminology and Context**
  - Most respondents (85% responded "Strongly Agree" or "Agree") agreed that the AI scribe effectively comprehended clinical terminology and context. In the study sample, the AI scribe was proficient primarily in recognizing and processing specialized medical language.

Overall Analysis

Most physicians viewed the AI scribe as an effective and beneficial tool, with solid satisfaction noted in its core functional areas. While the data (although limited by the small sample size) indicates a positive trend in the acceptance and utility of AI scribes, the neutral and negative feedback highlight areas where the technology could evolve to meet the needs of particular users better. These insights are crucial for developing and optimizing AI scribe technology to ensure it fully supports physicians in the ever-changing health-care landscape.

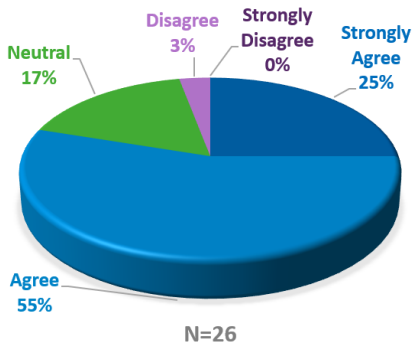

Figure 2 AI Scribe Functionality

Theme #3. Efficiency & Time Management

The theme of "Efficiency and Time Management" relates to the following subthemes, each represented by a survey question: real-time transcription performance, reduction of cognitive load, time savings in documentation, and timeliness in transcribing conversations into clinical notes. Overall, most physicians (49% and 35% responded "Strongly Agree" and "Agree", respectively) agreed that the AI scribe is an efficient tool that saves time and reduces cognitive load (Figure 3).

- **Real-time Transcription Performance**
  - Most respondents (80% responded "Strongly Agree" or "Agree") agreed that the AI scribe performed well in real-time transcription during conversations. In this sample of physicians, the AI scribe appeared to capture conversations as they occurred effectively. This is critical for maintaining the flow of clinical encounters in this study sample.
- **Streamlining Clinical Documentation**
  - Most respondents (58% and 27% responded "Strongly Agree" and "Agree", respectively) agreed that the AI scribe reduced their cognitive load by minimizing the need to recall and document information from memory. The AI scribe supported physicians in this sample by reducing the mental effort required during patient documentation.
- **Time Savings in Documentation**
  - Most respondents (54% and 31% responded "Strongly Agree" and "Agree", respectively) agreed that the AI scribe saved them time when documenting patient encounters compared to manual methods. This is a crucial aspect of the AI scribe's value proposition, indicating that it can streamline the documentation process at least among the physician respondents.
- **Timeliness in Transcribing Conversations into Clinical Notes**
  - Most respondents (42% and 38% responded "Strongly Agree" and "Agree", respectively) agreed that the AI scribe transcribed patient-physician conversations in a timely manner. According to our sample, the AI scribe captured information effectively and swiftly, which can be essential for maintaining up-to-date patient records.

Overall Analysis

Overall, the respondents agreed that the AI scribe positively impacted workflow efficiency and time management. Most physicians who participated in this study appeared to view the AI scribe as a tool that enhances real-time transcription, reduces cognitive load, and saves time in clinical documentation, all of which significantly contribute to improved practice efficiency. While the feedback is predominantly positive, the presence of neutral and negative responses suggests that there are still areas where the AI scribe can be optimized to better serve all users' needs. This feedback is valuable for guiding future enhancements to the AI scribe technology, aiming to maximize efficiency and user satisfaction across all aspects of time management in clinical practice.

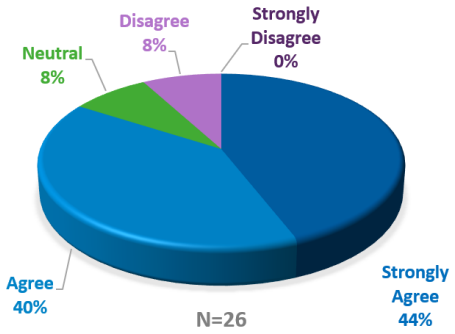

Figure 3 Efficiency & Time Management

Including "Time Savings in Documentation" and "Timeliness in Transcribing Conversations into Clinical Notes" in both Theme #1 and Theme #3 of the analytics report is deliberate and essential for a comprehensive analysis. Firstly, these factors hold overlapping importance in both themes, emphasizing their central role in reducing physicians' workload and enhancing overall efficiency in clinical practice. Secondly, they offer different contextual implications; in Theme #1, they are pivotal for managing workload and reducing burnout, whereas in Theme #3, they focus on enhancing overall efficiency and time management. This repetition also reinforces the findings, highlighting the critical nature of these factors in various aspects of the physicians' work. Moreover, it ensures a comprehensive analysis by acknowledging the multifaceted impact of these factors in clinical settings. Finally, this strategic repetition clarifies for stakeholders the integral role these factors play in improving multiple dimensions of clinical practice, thereby demonstrating the interconnectedness of various elements within the health-care system.

Theme #4. Patient Perception and Satisfaction

The theme of "Patient Perception and Satisfaction" relates to how the AI scribe is perceived by patients, including its impact on their experience during medical visits. The following subthemes, each represented by a survey question, were included: comfort with AI scribe use, patient satisfaction with their care experience, accurate transcription of conversations, and concentration on patient interactions. Overall, most physicians (88% responded "Strongly Agree" or "Agree") agreed that their patients viewed the AI scribe positively (Figure 4).

- **Comfort with AI Scribe Use**
  - Almost all respondents (65% and 31% responded "Strongly Agree" and "Agree," respectively) reported that their patients felt comfortable using AI scribes during consultations. To successfully integrate AI scribes, it is essential that patients do not find the technology intrusive and feel at ease with its presence during consultations.
- **Patient Satisfaction with Care Experience**
  - Most respondents (69% responded "Strongly Agree" or "Agree") agreed that patients were satisfied with their care experience when AI scribes were used. As suggested by these findings, AI scribe has the potential to enhance the quality of patient care.
- **Concentration on Patient Interactions**

- Almost all respondents (54% and 42% responded "Strongly Agree" and "Agree," respectively) agreed that the AI scribe helped them to concentrate more on patient interactions. In our sample of physicians, the AI scribe appeared to have reduced certain administrative burdens related to patient care, such as documentation. Thus, physicians may focus more on the patient, enhance communication, and provide better patient care.

Overall Analysis

According to our study respondents, integrating AI scribes into clinical practice appeared to have positively influenced patient experience. Patient comfort and satisfaction with AI scribes, coupled with a physician's ability to focus on patient interactions, demonstrate the value of AI scribes beyond efficiency gains. AI scribes may also play a significant role in enhancing the quality of interactions between physicians and patients.

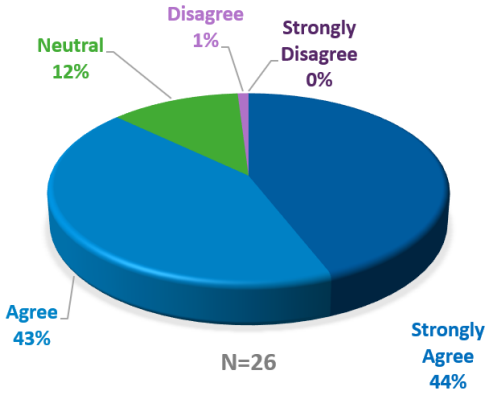

Figure 4 Patient Perception and Satisfaction

Theme #5. User Experience and Interface

The theme of "User Experience and Interface" relates to the usability of the AI scribe system, as well as the quality of the training. Overall, most physicians (50% and 39% responded "Strongly Agree" and "Agree", respectively) reported a positive experience with the AI scribe and training (Figure 5).

- **AI Scribe Vendor's Training**
  - A majority of respondents (85% responded "Strongly Agree" or "Agree") agreed that the training provided by the AI scribe's vendor met their expectations. Some respondents were neutral (8%) or disagreed (8%).
- **AI Scribe's User Interface**
  - A majority of respondents (61% and 31% responded "Strongly Agree" and "Agree," respectively) agreed that the AI scribe's user interface was easy to navigate and user-friendly. User experience (UX) design is an essential factor in adoption and sustained use. However, as with the training, a small portion of users (8% combined for neutral and disagree) suggest that there might be some aspects of the interface that could be enhanced for a better user experience.

Overall Analysis

The general feedback on the AI scribe system is highly positive, with physicians appreciating its initial training and user interface. This feedback is crucial for developers to enhance the system further. However, the small sample size and potential randomness in the responses suggest a more extensive study and direct interviews with physicians for more in-depth insights are needed.

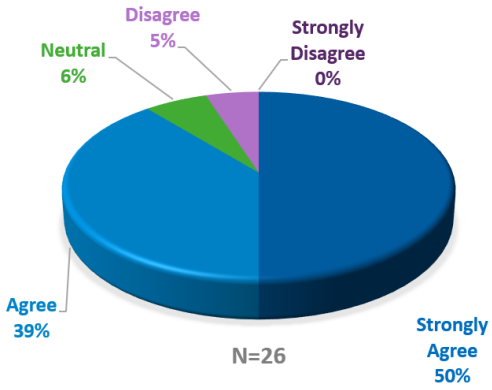

Figure 5 User Experience and Interface

In conclusion, the physician survey results present a broadly positive view of the AI scribe system across various themes, including clinical workload management, AI scribe functionality, efficiency, patient satisfaction, and user experience. Notably, physicians reported benefits such as reduced cognitive load, time savings in documentation, improved patient interactions, and satisfaction with the system's training and user interface. These findings are encouraging and suggest that the AI scribe system has significant potential to enhance clinical practice.

Lessons Learned and Limitations:

- **Sample Size and Diversity:** The primary limitation of this study was the challenge in recruitment, leading to a small and possibly non-representative sample size. This limits the generalizability of the findings.
- **Potential for Bias and Randomness:** The responses, while positive, may contain biases or randomness not accounted for due to the limited sample. A larger and more diverse sample would be essential for more robust conclusions.
- **Depth of Understanding:** While the survey provided initial insights, it lacked depth in fully comprehending the reasons behind physician responses. Future studies should incorporate qualitative methods such as interviews and involve all clinicians, including nurses, to gain a more comprehensive understanding and diverse perspectives.

Recommendations for the Next Phase:

- **Enhanced Recruitment Efforts:** To address the sample size issue, future phases should focus on broader and more effective recruitment strategies to ensure a more diverse and representative sample of physicians.
- **Incorporating Qualitative Research:** Alongside quantitative surveys, incorporating interviews or focus groups will provide more detailed feedback on user experiences and perceptions.
- **Continual Refinement:** The feedback, especially neutral and negative responses, should guide the ongoing refinement of the AI scribe system to address specific user needs and improve overall functionality.
- **Outlining Appropriate Accountability:** Establishing clear roles and responsibilities for the AI software developer, the vendor, and the physician is essential to create a transparent line of accountability within the terms of the agreement and privacy policies.
- **Streamlining of Agreements:** As the vendor's terms of use and privacy policy were unclear regarding roles and responsibilities and did not embody the expected tone expected from the physician, implementing a streamlined approach across all agreements and documents would ensure consistency and alignment with the pilot's goals.

Patient Survey Results:

The patient survey results indicate a strong positive response to using an AI scribe during medical visits. Analysis of the aggregated responses for 'Agree' and 'Strongly Agree' reveals the following:

- **Focus on Conversation (95%):** Nearly all patients agreed that using the AI scribe allowed their physician to focus more on the conversation than documentation. Importantly, this suggests that the AI scribe did not distract from the patient-physician interaction but enhanced it by allowing the physician to be more present and attentive.
- **Comfort with AI Scribe (92%):** Most patients report comfort in using the AI Scribe during visits.
- **Visit Satisfaction (85%):** Most patients reported that their visit was more satisfactory when using the AI scribe. This satisfaction could be associated with a perceived improvement in the quality of interaction with the physician, possibly due to the physician having more time to spend with the patient than on documentation, aligned with physicians' reports in their survey.
- **Enhanced Physician-Patient Relationship (78%):** Most patients reported that using the AI scribe enhanced their relationship with the physician.

Each of these findings underscores the positive reception of AI scribe technology from the patient's perspective. The strong agreement in conversation focus, comfort, satisfaction, and relationship enhancement suggests that the AI scribe is a welcome addition to the patient experience in a clinical setting. Moreover, these results may alleviate concerns about the potential negative impact of technology on the patient-physician dynamic. Instead, they point to an augmentation of the quality of care and patient satisfaction.

The high percentages of agreement in all categories reflect well on integrating the AI scribe into clinical practice from the patient's viewpoint. However, it is essential to consider that this feedback is based on

patients' subjective experiences and perceptions. Potential areas for further research include detailed patient feedback on their comfort and satisfaction levels to understand better the underlying factors contributing to these high levels of approval.

A comparison of physician and patient survey data shows a largely positive reception of AI scribes. In our samples, both physicians and patients appeared to realize the technology's ability to enhance the quality of interactions in clinical settings. There's a notable agreement that AI scribes help maintain focus during visits and foster stronger physician-patient relationships. As reported by patients, high comfort levels with AI scribes are mirrored by physicians' observations, pointing towards the seamless technology integration in health-care environments. The alignment in reports of increased satisfaction with the care experience further affirms the AI scribe's contribution to improved clinical outcomes. Although the current feedback is affirmative, incorporating objective outcome measures could further substantiate the effectiveness of AI scribes.

Time Motion Study Results:

The time-motion study aimed to quantify the impact of AI scribes on time efficiency within clinical settings. The study assessed the total time spent by physicians on consultations and the creation of clinical notes (i.e., SOAP notes) before and after adopting the AI scribe system.

Two measures were derived to analyze the data: change in total minutes and change in average minutes per consult. The change in total minutes was calculated as the total minutes spent in consultations post-AI scribe less the total minutes spent in consultations pre-AI scribe. Negative values represented time savings by using the AI scribe (i.e., less time spent in consultations with using the AI scribe). The change in average minutes per consult represented a standardized (i.e., per consultation) version of the change in total minutes measure. Negative values represented time savings per consultation using the AI scribe (i.e., less time spent per consultation with the AI scribe). Both measures were averaged across the physician participants to obtain the change in total minutes per physician and the change in average minutes per consultation and physician. Findings from the latter measure are reported in this paper.

Average Minutes per Consultation

Focusing on the change in average minutes per consult measure, the average physician who participated in this study saved 3.4 minutes per consultation (Figure 6). The change in average minutes per consult measure ranged from a maximum value of 2 minutes (i.e., time lost) to a minimum value of -11 minutes (i.e., time saved). Most physicians (84%) experienced decreased average minutes spent per consultation using AI scribe.

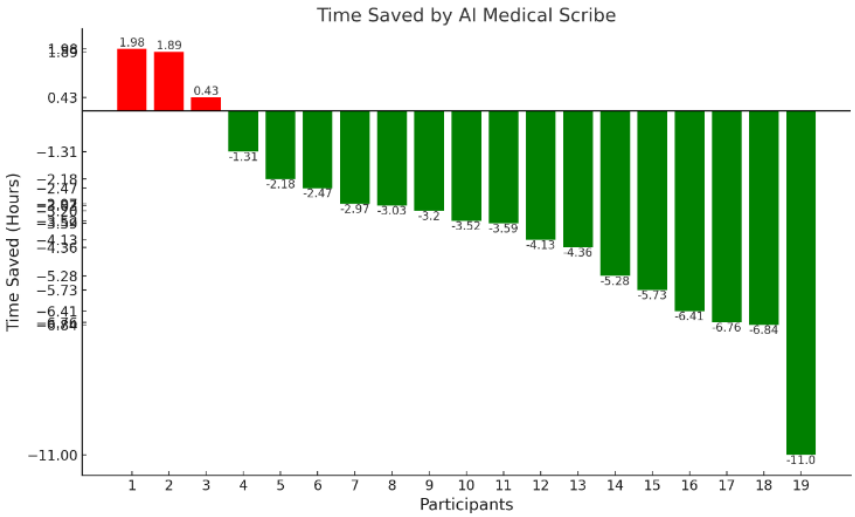

Figure 6 Change in Average Minutes per Consultation

Qualitative Analysis

This section discusses the experiences of physicians with the AI scribe based on an in-depth analysis of the open-ended responses from the physician survey.

Physician experience using AI scribe:

Positive Aspects Highlighted by Physicians:

- **Transcription Accuracy:**
  - 75% of physicians appreciated the AI scribe's ability to accurately capture key elements of the conversation, with some indicating that it required minimal editing.
  - Specifics: A notable proportion mentioned efficiency in capturing psychosocial elements and integrating well with existing EMR systems.
- **Cognitive Load Reduction:**
  - There was a consensus on reducing the cognitive load (90%), as physicians felt reassured during complex encounters and experienced less stress due to decreased manual documentation.
  - Specifics: Several responses indicated a 10% reduction in documentation time per patient, translating to significant cumulative weekly time savings.
- **Patient Interaction:**
  - 85% of physicians reported more direct patient interaction, citing increased eye contact and attention, which they believe could enhance the patient-physician relationship.

- Specifics: Multiple physicians mentioned that the AI scribe enabled them to be more present and attentive, potentially improving the quality of care.

Challenges and Limitations Noted by Physicians:

- **Technical Limitations:**
  - Almost 20% of physicians pointed out the AI scribe's inability to capture physical examination details and medication names accurately, which could be critical in medical documentation.
  - Specifics: Some indicated the need for loud verbalization of findings for accurate capture, which may not always be practical or comfortable in a clinical setting.
- **Formatting Issues:**
  - The repetition and misplacement of information within the SOAP note format were highlighted by almost one-fourth of participants as areas needing improvement.
  - Specifics: Instances of information being erroneously repeated across subjective, objective, and assessment sections were commonly noted.
- **Distinction in Documentation:**
  - 15% of physicians agreed that differentiating between multiple patients' issues or capturing the nuances of family history versus the patient's subjective history was difficult.
  - Specifics: In a few cases, capturing the interactions with more than one person, the AI scribe was unreliable and had difficulty differentiating between patients' subjective history and when patients mentioned family member history.

Impact on Clinical Practice:

- **After-Hours Workload:**
  - 50% of physicians agreed that a reduction in after-hours documentation was a recurring positive theme, with physicians reporting more thorough notes and a reduced likelihood of omitting details.
  - Specifics: Some physicians mentioned saving around 2 hours per week, while others expressed relief at not having to catch up on notes after hours.
- **Clinical Efficiency:**
  - Despite the benefits, there were mixed responses regarding changes in the number of patients seen weekly. Some physicians did not report an increase, often due to already being at full capacity, while 30% of physicians noted improved workflow efficiency.
  - Specifics: Where increases were noted, they ranged from 5 to 15 additional patients per week.

AI scribe impact on documentation time and cognitive load

- **Reduction in Cognitive Load and Documentation Time:**

- A majority (90%) of respondents acknowledged reduced cognitive load and time spent on documentation.
- Specifics: Physicians reported feeling a sense of relief and relaxation, with one stating that the AI scribe's support allowed for more relaxed working conditions.

- **Efficiency Gains:**
  - Many physicians (80%) noted efficiency gains, ranging from several minutes saved per visit to several hours per week.
  - Specifics: The AI scribe's ability to capture clinical notes contributed to some physicians experiencing the best practice days in a long time due to up-to-date and complete charting.
- **Need for Note Review and Completion:**
  - 20% of respondents did not feel the AI scribe significantly impacted their documentation time or cognitive load.
  - Specifics: Physicians still needed to review and manually add information after visits, particularly for unverbalized physical examination details.
- **Non-reduction of Certain Tasks:**
  - Less than 10% of physicians pointed out that the AI scribe did not save time on tasks such as creating requisitions and sending referrals, which comprise a substantial part of charting and paperwork.
  - Specifics: The AI scribe's current capabilities may not extend to streamlining all aspects of clinical documentation, necessitating continued manual input for some administrative tasks.

Impact of AI Scribe on Patient Relationships

- **Enhanced Physician-Patient Interaction:**
  - A substantial majority (85%) of physicians noted a significant positive impact on their patient relationships due to using the AI scribe.
  - Specifics: A physician highlighted the unobtrusive nature of the tool, which was well-received by patients and facilitated more open communication.
- **Increased Direct Time with Patients:**
  - 85% of physicians reported that the AI scribe allowed them to spend more direct time with patients and be more present during encounters.
  - Specifics: Knowing that notes were being generated in the background, physicians could focus more on the patient, enhancing the quality of interactions and making the visits more collaborative.
- **Improved Focus and Engagement:**
  - Physicians (85%) felt their patients appreciated using the AI scribe as it allowed for a deeper focus on health issues during visits.
  - Specifics: This technology enabled more eye contact, less screen distraction, and more interaction through verbal and nonverbal communication. Physicians experienced a shift towards more collaborative conversations about health when the pressure to type and document was alleviated.

AI scribe impact on documentation after-hours time and tasks

- **Reduced Documentation Time and Increased Efficiency:**
  - 60% of physicians agreed that the AI scribe reduced documentation time, enhancing work efficiency.
  - Specifics: The physicians experienced efficiency improvements, such as completing SOAP notes within 5 to 7 minutes and a decreased need to update notes post-encounter. This reduction in time needed for documentation tasks also reduced the necessity to "catch up" on notes.
- **Ability to Complete Work Within Clinic Hours:**
  - Many physicians (70%) noted that the increased efficiency allowed them to finish their work within clinic hours and reduced the need to take work home after hours.
  - Specifics: The time saved on documentation shifted focus towards other administrative tasks, such as inbox management and forms, thereby improving overall workflow management and reducing after-hours work.

CONCLUSION

The OMD evaluation study of WELL AI Voice is an initial step in understanding the impact of AI scribe in modern clinical practice. While the pilot study indicates a promising potential for AI scribes to enhance physician efficiency and patient care, it is essential to consider these findings within the context of the study's limitations.

Key Findings:

- **Clinical Relevance and Accuracy**: The AI scribe demonstrated a solid ability to transcribe clinical conversations, accurately capturing complex medical terminology.
- **Impact on Physician Burnout**: Reduced time and cognitive load associated with documentation were observed, potentially alleviating a critical factor in physician burnout. This shift can allow physicians to concentrate more on patient care, potentially enhancing job satisfaction and overall well-being.
- **Patient Perspective and Experience**: Patients reported increased comfort and satisfaction during WELL AI Voice consultations. The technology complemented, rather than detracted from, the physician-patient relationship, enhancing interaction quality and care.
- **Scalability and Support**: WELL AI Voice can handle increasing volumes of patient data and conversations, backed by continuous development and a dedicated support team.

Limitations and Future Directions:

- The small, non-random sample size limits the generalization of findings and statistical tests.
- Potential biases, including selection bias and response bias, could have influenced the results.
- The incentivization method for participation may have introduced bias.
- Future phases of the study aim to recruit a more extensive and diverse sample of physicians, applying more rigorous research designs and sampling methods to overcome these limitations.
- Qualitative research methods should be employed to deepen the understanding of the AI scribe's impact, exploring nuances in time savings and clinical efficiency.

In conclusion, while the results from this pilot study are encouraging, suggesting a positive trend in the adoption and effectiveness of AI scribes, they should be interpreted cautiously. The potential benefits in clinical practice, patient experience, and physician well-being are evident. However, further research with a larger, more diverse sample and enhanced methodological rigour is necessary to confirm and expand upon these findings. The future of AI scribes in health care looks promising. Still, a balanced approach considering both the technological advancements and the inherent limitations of the current study is essential for their successful integration and optimization in clinical settings.

APPENDICES

Appendix 1: AI Scribe Vendor Questionnaire

1. Could you please describe the level of integration between the 'WELL AI Voice' and WELL's OSCAR Pro EMR system? How seamlessly does the AI scribe integrate and function within this system?
2. And how about other EMR systems? Is your product versatile enough to work with various EMRs across Canada?
3. Does WELL AI Voice collect any Personal Health Information (PHI)?
4. If it does, how does your system go about de-identifying PHI?
5. Please explain how WELL AI Voice complies with crucial data security and privacy laws, such as PHIPA and PIPEDA.
6. Do you ensure all data remains within Canadian borders?
7. Do you employ encryption to secure data during transit and at rest?
8. Can WELL AI Voice handle an increase in patient data and conversations as the needs of the clinic evolve? Could you explain this further?
9. Could you detail how your team's efficiency in solving problems, providing training, and ongoing support? Do you have a specific team assigned for support?
10. Could you share your product's Technology Readiness Level (TRL) with us?
11. Could you give us a high-level explanation of the AI or ML techniques used in WELL AI Voice?
12. What is the product development cycle like for WELL AI Voice? How often do you roll out updates or new features?
13. Can you share any evaluation results you might have done on the WELL AI voice users?
14. How does WELL AI Voice handle different accents or speech patterns? Does it learn and adapt over time?
15. How does WELL AI Voice handle potential errors or inaccuracies in transcriptions? Is there a mechanism for correction or learning from mistakes?
16. What future improvements or developments can users expect to see in the WELL AI Voice?
17. Does the company have a disaster recovery plan in place in case of data breaches or loss?
18. How easy is it to train users on your system? Do you provide any instructional materials or resources?
19. How customizable is WELL AI Voice to a specific clinic’s needs and workflows?
20. What measures are taken to ensure the accessibility of the tool to users with varying levels of tech proficiency?

Appendix 2: Participants Survey

Most of the questions used a 5-point Likert scale for responses, offering multiple choice options ranging from "Strongly Agree" to "Strongly Disagree."

1. The AI scribe **accurately**[1] transcribes patient-physician conversations, correctly capturing individual elements within the transcript, such as words or sentences. *For instance, if a statement like "The patient's blood pressure is 120 over 80" is spoken, the AI scribe captures it verbatim.*
2. The AI scribe **comprehensively**[2] captures the full scope of a conversation, including all relevant information. *For example, in a dialogue discussing multiple symptoms, past diagnoses, current medications, and new treatment plans, the AI scribe includes all these details, capturing the full scope of the conversation.*
3. The AI Scribe transcribes the patient-physician conversation into clinical notes in a **timely manner**[3] (less than 10 seconds).
4. The AI scribe effectively comprehends **clinical terminology** and context, contributing to the overall accuracy and relevance of the transcriptions.
5. The AI scribe's **user interface** is easy to navigate and user-friendly.
6. I am **satisfied** with the performance (functions and features) of the AI scribe:
   - performance: how well the AI scribe executes tasks such as accurate transcriptions and swift generation of clinical notes, and
   - functionality: the specific features and operations the AI scribe is capable of, such as real-time transcription, understanding clinical terminology, and seamless integration with existing systems.
7. How would you describe your experience using AI scribe, and what aspects did you like or dislike? *(open-ended question)*
8. The AI vendor's **training** meets my expectations.
9. The AI scribe performs well in **real-time transcription**[4] during conversations, accurately transcribing even in the face of multiple speakers, interruptions, or overlapping speech.
10. The AI scribe **adapts** well to different accents, dialects, and speech patterns.
11. The AI scribe **saves time** when documenting patient encounters compared to manual methods.
12. The AI scribe reduces my **cognitive load** by minimizing the need to recall and document information from memory.

[1] This refers to how correctly the AI scribe can transcribe and interpret the conversations. It concerns the correctness of individual elements within the transcript, such as words, sentences, or medical terminologies. For example, if a doctor says, "The patient's blood pressure is 120 over 80," an accurate transcription would capture this statement exactly as spoken.

[2] This refers to the ability of the AI scribe to capture all relevant information within a conversation, not just transcribing it correctly. For example, a comprehensive transcription would include all of these details in a conversation between a doctor and a patient discussing multiple symptoms, a past diagnosis, medication being taken, and a new treatment plan. It goes beyond getting the individual words or sentences right and involves capturing the full scope of the conversation.

[3] This refers to the system's ability to process the recorded or input conversation and subsequently generate the corresponding clinical notes. It is essentially asking about the efficiency of the system in converting the conversation into a written record after the conversation has concluded. It may involve elements like natural language processing speed, computing power, and the system's ability to manage and prioritize tasks.

[4] This point is more about the AI scribe's ability to keep up with a live conversation in real time. For example, can it transcribe accurately as people are speaking? Can it handle multiple speakers, interruptions, or overlaps in speech? This functionality could be crucial in a fast-paced clinical environment where live transcription might be beneficial, like in the middle of consultations, surgeries, or clinical team discussions.

13. How has the AI scribe influenced your documentation time during patient encounters compared to manual methods, and have you observed any effects on reducing cognitive load? *(open-ended question)*
14. The AI scribe helps **redistribute**[5] the workload, taking over administrative tasks and enabling me to focus more on direct patient care and other critical tasks.
15. Using the AI scribe has allowed me to concentrate more on patient interactions, influencing the overall **patient-physician relationship**.
16. *Please provide insights into how AI scribes have impacted your patient relationships.*
17. My patients are **comfortable** with the use of AI scribes during their visits.
18. My patients report higher **satisfaction** with their care experience when AI scribes are used.
19. How has the AI scribe influenced your documentation tasks, including its impact on after-hours documentation and administrative duties? Please highlight specific experiences or instances that demonstrate its effect on your daily workflow and the potential reduction in manual notetaking and data entry. *(open-ended question)*
20. Considering your clinical practice both before and after implementing the AI scribe, could you estimate how the average number of patients you see per week has changed? If possible, please refer to your schedules for a more accurate comparison. *(open-ended question)*

[5] If an AI scribe is effective, it should be able to take over some administrative tasks. By accurately transcribing conversations and generating clinical notes, the AI scribe can reduce physicians' time on paperwork. This would allow health-care professionals to focus more on direct patient care, making critical decisions, and other tasks that can't be automated.

Appendix 3: Participants (Patients) Survey

Most of the questions used a 5-point Likert scale for responses, offering multiple choice options ranging from "Strongly Agree" to "Strongly Disagree."

1. The use of the AI scribe allowed my doctor to focus more on our conversation rather than on documentation.
2. “I was comfortable when the doctor used the AI scribe during my visit.” This is about whether you felt okay when the doctor used a computer program to write down what was said during the visit.
3. "My visit was more satisfactory when the doctor used the AI scribe." This is about whether the doctor’s use of the AI scribe improved your experience during the visit.
4. The use of the AI scribe during my visit enhanced my relationship with the doctor.

Appendix 4: WELL AI Voice Product Information

Integration with EMR Systems: WELL AI Voice offers seamless integration with the OSCAR Pro EMR system, facilitating seamless translation of conversations into structured SOAP notes and easy insertion into electronic charts. The platform is also adaptable for use with various EMRs across Canada, allowing physicians to integrate notes into different systems.

Handling of Personal Health Information (PHI): The system captures necessary PHI during conversations but employs robust de-identification processes for transcripts before third-party use. It complies with PHIPA and PIPEDA through data encryption and adherence to privacy laws.

Data Security and Localization: For purposes of the Pilot, WELL AI Voice is committed to ensuring data security with encryption both in transit and at rest, with patient-related data processed and stored within Canada. A portion of operational data is managed in the United States.

System Scalability and Support: The platform can handle increasing patient data volumes and offers efficient problem resolution, training, and ongoing support through a dedicated team.

Technology and Development: WELL AI Voice is at Technology Readiness Level (TRL) 9 and employs advanced speech recognition, natural language processing, and generative AI techniques. It follows a continuous development cycle, with regular updates and new features.

User Feedback and System Adaptability: Anecdotal evidence suggests substantial time savings and improved user clinical experiences. The system continuously improves speech models to handle various accents and learns from feedback to refine transcription accuracy.

Disaster Recovery and User Training: The company has a comprehensive disaster recovery plan for data breaches or loss. Training for new users is straightforward, with additional support through upcoming video resources.

Customization and Accessibility: WELL AI Voice aligns with individual clinic workflows and documentation standards, offering a user-friendly interface and responsive support for users of varying technical proficiencies.