Physicians' Lived Experiences with AI Scribes

Nina Huang*
ninaa.huang@mail.utoronto.ca
University of Toronto
Canada

Katherine Jelich*
katherine.jelich@mail.utoronto.ca
University of Toronto
Canada

Brenna Li
brli@cs.toronto.edu
University of Toronto
Canada

Khai N. Truong
khai@cs.toronto.edu
University of Toronto
Canada

Alex Mariakakis
mariakakis@cs.toronto.edu
University of Toronto
Canada

Abstract

Clinical documentation has become a major source of burnout for physicians, as increasing aspects of the clinical encounter must be recorded in electronic health record (EHR) systems. AI scribes have emerged as a potential solution, automatically converting conversations between patients and physicians into clinical notes, yet little is known about how physicians perceive and integrate these tools into everyday workflows. We conducted clinical observations and semi-structured interviews with 17 primary care physicians in Canada who had experience using AI scribes. We found that physicians primarily used AI scribes as transcribers and legal documenters to reduce their cognitive load and ensure record completeness. However, the scribes fell short of fully satisfying these roles because of repetitive phrasing, irrelevant content, and recommendations that overstepped clinicians' boundaries. Our work aims to explore how the roles and capabilities of AI scribes can be refined and integrated to better support physicians — not only in documentation, but also in broader aspects of clinical workflows and care delivery.

CCS Concepts

• **Human-centered computing → Empirical studies in HCI**; • **Applied computing → Health informatics**.

Keywords

AI scribes, LLM, primary care, documentation burden

ACM Reference Format:
Nina Huang, Katherine Jelich, Brenna Li, Khai N. Truong, and Alex Mariakakis. 2025. Physicians' Lived Experiences with AI Scribes. In *Proceedings of (CHI '25 Workshop on Envisioning the Future of Interactive Health).* ACM, New York, NY, USA, 4 pages.

1 Introduction

The burden of electronic health record (EHR) documentation is a leading contributor to physician burnout [12, 17, 20]. A single patient encounter requires physicians to manage extensive documentation, including structured data like prescriptions, lab orders, billing codes, and unstructured free-text clinical notes [1, 16]. These documentation tasks can consume between 25% to 50% of a physician's working time, amounting to as much as four hours each day [7, 9]. This substantial documentation burden not only leads to burnout but also impairs the quality of care, as physicians spend more time interacting with their EHRs than engaging face-to-face with their patients [21].

Ambient AI scribes, also known as digital scribes, have recently gained traction as a viable solution to alleviate the burden of medical documentation. Designed to function as virtual medical assistants, these systems passively listen to conversations between physicians and patients in order to generate structured clinical notes in real time [3, 5, 11, 13]. Recent advancements in speech-to-text algorithms and large language models (LLMs) have significantly advanced the capabilities of AI scribes [23], making them more accurate and coherent while preserving contextual understanding [22].

Despite the growing adoption of AI scribes, a deeper understanding of how physicians are integrating them into their daily practices is still needed. Preliminary studies have shown that AI scribes reduce perceived burden and documentation time [14, 19]. However, key questions about how well these systems meet physician expectations, the challenges that persist, and the future roles this technology should play in clinical workflows remain largely unexplored.

This work presents a preliminary analysis of our ongoing research investigating these questions. We have been using a combination of clinical shadowing and interviews to understand the roles that AI scribes have been serving and where they have fallen short. In doing so, we seek to provide a comprehensive overview of AI scribes' capabilities and limitations, offering actionable insights for developers, healthcare providers, and policymakers involved in their deployment.

2 Methods

2.1 Participants

We recruited licensed primary care physicians in the Greater Toronto Area through online forums, hospital organizations, and regional health organization mailing lists from September to November 2024. Physicians were only eligible if they had current or prior experience using an AI scribe. At the time of this

*Both authors contributed equally to this research.

CHI '25 Workshop on Envisioning the Future of Interactive Health, Yokohama, Japan

manuscript, 17 physicians (6 males, 11 females) have completed the study protocol, and their average number of years in independent practice was 11.3 ± 10.4 years.

2.2 Study Design

Our study offered two enrollment options: in-person clinical shadowing and virtual interviews. For physicians who were willing to be observed (N = 5), two researchers visited their practices and shadowed their patient interactions for up to four hours. This was followed by a 30-minute interview to triangulate observations with physicians' past experiences and to reflect on what transpired during the shadowing. Physicians who were unable to accommodate clinical shadowing (N = 12) were asked to provide a walkthrough of their typical scribe usage, offering insights into what might be observed in their clinics. The walkthrough was followed by a semi-structured interview similar to one conducted with the other physicians. These sessions were held online and typically lasted 1.5 hours.

It is important to note that the physicians in our study used a range of AI scribe products with different features. Some clinics endorsed and subsidized specific AI scribes, while in other cases, physicians chose products on their own. A product comparison is beyond the scope of our work, but the diversity of options provided additional insight into the potential roles that AI scribes can support.

3 Results

3.1 Primary Benefits of AI Scribes

Physicians often viewed AI scribes as serving two primary roles: transcription and legal documentation. As a transcriber, an AI scribe converts conversations into a summarized physician note in real time. This contrasts with past experiences where physicians had to simultaneously listen to patients while taking notes to document the visit. By serving in this role, AI scribes helped physicians offload the cognitive burden that they typically experienced when multitasking. P14 claimed, *"It's the ability not to have to split my focus during the visit that I have found more valuable than anything else."*

Although complete medical records are crucial for protecting healthcare providers and patients in the event of legal disputes, physicians struggle to document all the details of a visit due to time constraints and limited attention capacity. As a legal documenter, an AI scribe captures consultation details that physicians may not have otherwise typed themselves. P12 noted, *"[The AI scribe] helps to create a much better patient record and narrative. When you're going back in time, you have a much better record of what actually transpired."*

3.2 Primary Limitations of AI Scribes

Despite the aforementioned benefits, physicians recounted times when their AI scribes fell short of meeting their expectations. For example, many felt that AI scribes were unnecessary for transcribing routine visits with a predictable procedure. As explained by P12, *"There are certain encounters where I have a very specific template that I have made that I just have to click. I don't use an AI scribe [in those scenarios] because it takes me no time to finish the note."* In these cases, physicians felt that it was more expedient to directly enter structured information into their EHR templates than to review and edit their AI scribe's notes after the consultation.

Those who viewed an AI scribe as a legal documenter noted that AI scribes often struggle to balance thoroughness and conciseness. While they appreciated having a complete record of their consultations, this objective conflicted with their preference for brevity and comprehensibility. P13 expressed, *"I don't care as much if [all the information is] put on the note. The note is just not as nice when I go back and read it, or doesn't make sense from a clinical perspective."* Since physicians use their notes to simultaneously maintain a clear record and to inform future visits, achieving a balance between these purposes was considered critical.

3.3 Anticipated Benefits & Boundaries for Future AI Scribes

While transcription and legal documentation were the two most prominent benefits experienced by the physicians, they also identified secondary functions that AI scribes were either already fulfilling or could better fulfill in the future. Physicians also commented on boundaries that future AI scribes should not cross.

3.3.1 Administrative Assistance. Beyond reducing documentation burden, most physicians saw opportunities for AI scribes to assist with administrative tasks in their workflows. As P15 described, *"Being able to dictate a note and have the scribe also create lab requisitions and referral letters would be a huge help."* Some physicians envisioned triggering these actions using explicit voice interactions (e.g., "Order a CBC"), while others foresaw AI scribes taking a more autonomous role. For example, P8 explained, *"If I say 'I'm going to refer you to a cardiologist,' the AI should generate the consultation note and find a local specialist."*

3.3.2 Decision Support. The idea of using AI scribes for decision support was met with a mixed reception. Some physicians saw value in AI scribes surfacing clinically relevant information. For example, P3 noted that AI scribes could reduce the need for manual EHR searches by proactively retrieving useful information from past visits and lab tests: *"When patients say, 'Do you remember when I came in two years ago for knee pain?' No, I don't, but the chart does. Right now, I have to manually search for past notes, but an AI could surface them automatically."*

However, concerns emerged regarding how AI scribes could subtly influence clinical decision-making. P17 described losing trust in the system when it began suggesting next steps: *"I was hoping it would be helpful for history taking, but I was disappointed when it would make inferences because that wasn't the role I was using it for."* Other physicians saw the potential for confirmation bias. Reflecting on a time when their AI scribe justified a clinical recommendation using references to literature, P10 commented, *"I worry that it's giving me false reassurance, especially with a diagnosis. It could be providing information supporting what I've already decided."*

Going a step further, physicians overwhelmingly rejected the notion of using AI scribes to automate clinical decisions. They consistently emphasized that AI scribes should assist rather than interpret or modify clinical reasoning. In P4's words, *"I don't want it creating diagnoses based on a compilation of symptoms. I wouldn't*

recognize that because it wasn't in my line of thinking." Some of these concerns were based on past experiences when AI scribes misinterpreted the conversation. For instance, P17 recounted a situation where their AI scribe fabricated the phrase "referred the patient to a specialist" in the note, misrepresenting the outcome of the consultation.

4 Discussion

4.1 Validating Efficiency Benefits

Physicians emphasized that AI scribes must genuinely reduce documentation effort rather than displace cognitive load elsewhere. Some reported that the need to review and correct AI scribe outputs limited the amount of time they saved by using one in the first place. Our work highlights that AI scribes serve two documentation roles with competing priorities. Conciseness is essential for physicians who wish to have their notes transcribed, while completeness is essential for physicians who wish to have legal documentation for their consultations. This tension presents an opportunity for designers to develop AI scribes capable of balancing these needs.

4.2 Adding Roles with EHR Integration

Most physicians used AI scribes that were not fully integrated with their EHR systems, so they frequently expressed frustration with having to manually copy-and-paste text from their AI scribes to their notes. Integrating AI scribes with EHR systems could not only help streamline chart reconciliation and maintain more accurate long-term patient records but also facilitate communication across multiple physicians within a healthcare team [4]. These opportunities have been largely underexplored because of the significant technical and regulatory barriers that have prevented AI scribe integration, particularly with respect to interoperability and privacy [2, 6, 15, 22]. Nevertheless, designers should proactively consider how AI scribes can support such roles.

4.3 Defining the Boundaries of AI Assistance

Physicians were cautious about the potential influence AI scribes could have on clinical reasoning. While retrieving past test results and patient history was seen as beneficial, some physicians expressed concerns that AI-generated suggestions could have undesirable consequences. Previous studies have validated such worries in other clinical applications of AI, reporting cases when physicians overruled their initially correct judgments based on erroneous AI suggestions [10, 18]. This raises another tension in the design of AI scribes: one between automation and autonomy. This tension has been well-reported in human-computer interaction literature, leading to the idea of mixed-initiative systems that are neither fully automated nor completely manual [8]. In fact, Coiera et al. [3] reflected on this tradeoff in their own commentary on AI scribes. By recognizing the different roles that AI scribes serve, we believe that designers can begin to probe the spectrum between automation and autonomy with greater success.

5 Conclusion

As AI scribes become more sophisticated and integrated into clinical systems, it is increasingly important that we reflect on their roles in supporting physicians. Although our preliminary findings have confirmed the belief that AI scribes alleviate documentation burden, we found that this affordance can actually be decomposed into two roles: transcription and legal documentation. We also found that physicians envision AI scribes serving additional roles ranging from administrative tasks to various forms of decision support. Defining these roles and establishing clear boundaries for AI scribes will be essential to maintaining physician trust and ensuring that these tools enhance rather than disrupt clinical workflows.

References

[1] Brian G Arndt, John W Beasley, Michelle D Watkinson, Jonathan L Temte, Wen-Jan Tuan, Christine A Sinsky, and Valerie J Gilchrist. 2017. Tethered to the EHR: primary care physician workload assessment using EHR event log data and time-motion observations. *The Annals of Family Medicine* 15, 5 (2017), 419–426.

[2] Henry Bundy, Jay Gerhart, Sally Baek, Crystal Danielle Connor, McKenzie Isreal, Ajay Dharod, Casey Stephens, Tsai-Ling Liu, Timothy Hetherington, and Jeffery Cleveland. 2024. Can the administrative loads of physicians be alleviated by AI-facilitated clinical documentation? *Journal of general internal medicine* 39, 15 (2024), 2995–3000.

[3] Enrico Coiera, Baki Kocaballi, John Halamka, and Liliana Laranjo. 2018. The digital scribe. *NPJ digital medicine* 1, 1 (2018), 1–5.

[4] Enrico Coiera and Sidong Liu. 2022. Evidence synthesis, digital scribes, and translational challenges for artificial intelligence in healthcare. *Cell Reports Medicine* 3, 12 (2022).

[5] Gregory Finley, Erik Edwards, Amanda Robinson, Michael Brenndoerfer, Najmeh Sadoughi, James Fone, Nico Axtmann, Mark Miller, and David Suendermann-Oeft. 2018. An automated medical scribe for documenting clinical encounters. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations*. 11–15.

[6] Tejal K Gandhi, David Classen, Christine A Sinsky, David C Rhew, Nikki Vande Garde, Andrew Roberts, and Frank Federico. 2023. How can artificial intelligence decrease cognitive and work burden for front line practitioners? *JAMIA open* 6, 3 (2023), ooad079.

[7] Rebekah L Gardner, Emily Cooper, Jacqueline Haskell, Daniel A Harris, Sara Poplau, Philip J Kroth, and Mark Linzer. 2019. Physician stress and burnout: the impact of health information technology. *Journal of the American Medical Informatics Association* 26, 2 (2019), 106–114.

[8] Eric Horvitz. 1999. Principles of mixed-initiative user interfaces. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 159–166.

[9] Erik Joukes, Ameen Abu-Hanna, Ronald Cornet, and Nicolette F de Keizer. 2018. Time spent on dedicated patient care and documentation tasks before and after the introduction of a structured and standardized electronic health record. *Applied clinical informatics* 9, 1 (2018), 46.

[10] Rohan Khera, Melissa A Simon, and Joseph S Ross. 2023. Automation bias and assistive AI: risk of harm from AI-driven clinical decision support. *JAMA* 330, 23 (2023), 2255–2257.

[11] Kundan Krishna, Sopan Khosla, Jeffrey P Bigham, and Zachary C Lipton. 2020. Generating SOAP Notes from Doctor-Patient Conversations. *arXiv preprint arXiv:2005.01795* (2020).

[12] Philip J Kroth, Nancy Morioka-Douglas, Sharry Veres, Stewart Babbott, Sara Poplau, Fares Qeadan, Carolyn Parshall, Kathryne Corrigan, and Mark Linzer. 2019. Association of electronic health record design and use factors with clinician stress and burnout. *JAMA network open* 2, 8 (2019), e199609–e199609.

[13] Brenna Li, Noah Crampton, Thomas Yeates, Yu Xia, Xirong Tian, and Khai Truong. 2021. Automating clinical documentation with digital scribes: Understanding the impact on physicians. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–12.

[14] Stephen P Ma, April S Liang, Shreya J Shah, Margaret Smith, Yejin Jeong, Anna Devon-Sand, Trevor Crowell, Clarissa Delahaie, Caroline Hsia, Steven Lin, et al. 2025. Ambient artificial intelligence scribes: utilization and impact on documentation time. *Journal of the American Medical Informatics Association* 32, 2 (2025), 381–385.

[15] Blake Murdoch. 2021. Privacy and artificial intelligence: challenges for protecting health information in a new era. *BMC Medical Ethics* 22 (2021), 1–5.

[16] J Marc Overhage and David McCallie Jr. 2020. Physician time spent using the electronic health record during outpatient encounters: a descriptive study. *Annals of Internal Medicine* 172, 3 (2020), 169–174.

[17] Louis Raymond, Guy Paré, Ana Ortiz de Guinea, Placide Poba-Nzaou, Marie-Claude Trudel, Josianne Marsan, and Thomas Micheneau. 2015. Improving performance in medical practices through the extended use of electronic medical record systems: a survey of Canadian family physicians. *BMC medical informatics and decision making* 15, 1 (2015), 27.

[18] Emely Rosbach, Jonathan Ganz, Jonas Ammeling, Andreas Riener, and Marc Aubreville. 2024. Automation Bias in AI-Assisted Medical Decision-Making under Time Pressure in Computational Pathology. *arXiv preprint arXiv:2411.00998* (2024).

[19] Shreya J Shah, Anna Devon-Sand, Stephen P Ma, Yejin Jeong, Trevor Crowell, Margaret Smith, April S Liang, Clarissa Delahaie, Caroline Hsia, Tait Shanafelt, et al. 2025. Ambient artificial intelligence scribes: physician burnout and perspectives on usability and documentation burden. *Journal of the American Medical Informatics Association* 32, 2 (2025), 375–380.

[20] Tait D Shanafelt, Lotte N Dyrbye, and Colin P West. 2017. Addressing physician burnout: the way forward. *Jama* 317, 9 (2017), 901–902.

[21] Christine Sinsky, Lacey Colligan, Ling Li, Mirela Prgomet, Sam Reynolds, Lindsey Goeders, Johanna Westbrook, Michael Tutty, and George Blike. 2016. Allocation of physician time in ambulatory practice: a time and motion study in 4 specialties. *Annals of internal medicine* 165, 11 (2016), 753–760.

[22] Aaron A Tierney, Gregg Gayre, Brian Hoberman, Britt Mattern, Manuel Ballesca, Patricia Kipnis, Vincent Liu, and Kristine Lee. 2024. Ambient artificial intelligence scribes to alleviate the burden of clinical documentation. *NEJM Catalyst Innovations in Care Delivery* 5, 3 (2024), CAT–23.

[23] Robert M Wachter and Erik Brynjolfsson. 2024. Will generative artificial intelligence deliver on its promise in health care? *Jama* 331, 1 (2024), 65–69.