

---

**Editorial**

# AI Scribes in Health Care: Balancing Transformative Potential With Responsible Integration

---

Tiffany I Leung<sup>1,2</sup>, MPH, MD; Andrew J Coristine<sup>1,3</sup>, MSc, PhD; Arriel Benis<sup>4</sup>, PhD

<sup>1</sup>JMIR Publications, Toronto, ON, Canada

<sup>2</sup>Department of Internal Medicine (adjunct), Southern Illinois University School of Medicine, Springfield, IL, United States

<sup>3</sup>Division of Cardiology (affiliate), Department of Medicine, McGill University, Montreal, QC, Canada

<sup>4</sup>Department of Digital Medical Technologies, Holon Institute of Technology, Holon, Israel

**Corresponding Author:**

Tiffany I Leung, MPH, MD

JMIR Publications

130 Queens Quay East, Unit 1100

Toronto, ON M5A 0P6

Canada

Phone: 1 4165832040

Email: [tiffany.leung@jmir.org](mailto:tiffany.leung@jmir.org)

---

## Abstract

The administrative burden of clinical documentation contributes to health care practitioner burnout and diverts valuable time away from direct patient care. Ambient artificial intelligence (AI) scribes—also called “digital scribes” or “AI scribes”—are emerging as a promising solution, given their potential to automate clinical note generation and reduce clinician workload, and those specifically built on a large language model (LLM) are emerging as technologies for facilitating real-time clinical documentation tasks. This potentially transformative development has a foundation on longer-standing, AI-based transcription software, which uses automated speech recognition and/or natural language processing. Recent studies have highlighted the potential impact of ambient AI scribes on clinician well-being, workflow efficiency, documentation quality, user experience, and patient interaction. So far, limited evidence indicates that ambient AI scribes are associated with reduced clinician burnout, lower cognitive task load, and significant time savings in documentation, particularly in after-hours electronic health record (EHR) work. One consistently reported benefit is the improvement in the patient-physician interaction, as physicians feel more present during a clinical encounter. However, these benefits are counterbalanced by persisting concerns regarding the accuracy, consistency, language use, and style of AI-generated notes. Studies noting errors, omissions, or hallucinations caution that diligent clinician oversight is necessary. The user experience is also heterogeneous, with benefits varying by specialty and individual workflow. Further, there are concerns about ethical and legal issues, algorithmic bias, the potential for long-term “cognitive debt” from overreliance on AI, and even the potential loss of physician autonomy. Additional pragmatic concerns include security, privacy, integration, interoperability, user acceptance and training, and the cost-effectiveness of adoption at scale. Finally, limited studies describe adoption or evaluation of these technologies by nonphysician clinicians and health professionals. Although ambient AI scribes and AI-driven documentation technologies are promising as potentially practice-changing tools, there are many questions remaining. Key issues persist, including responsible deployment with the goal of ensuring that ambient AI scribes produce clinical documentation that supports more efficient, equitable, and patient-centered care. To advance our collective understanding and address key issues, *JMIR Medical Informatics* is launching a call for papers for a new section on “Ambient AI Scribes and AI-Driven Documentation Technologies.” As editors, we look forward to the opportunity to advance the science and understanding of these fields through publishing high-quality and rigorous scholarly work in this new section of *JMIR Medical Informatics*.

*JMIR Med Inform* 2025;13:e80898; doi: [10.2196/80898](https://doi.org/10.2196/80898)

**Keywords:** ambient AI scribe; digital scribe; AI scribe; virtual scribe; documentation assistant; ambient listening technology; AI assistant; artificial intelligence; clinical documentation; electronic health records; administrative burden

## Introduction

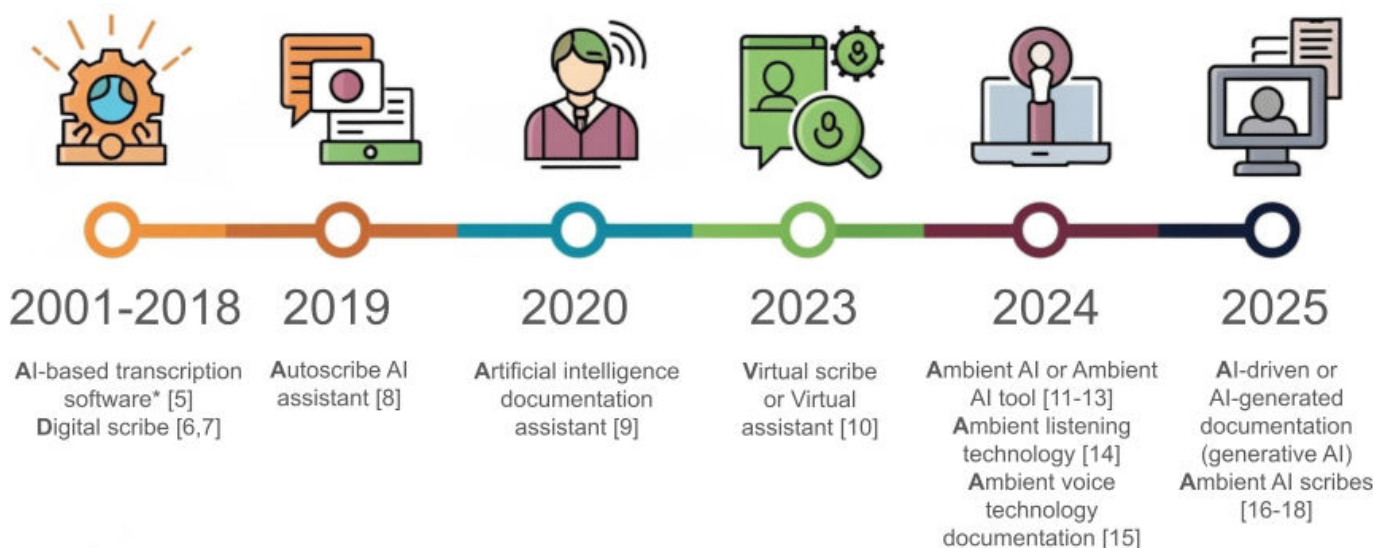
Administrative burdens associated with widespread electronic health record (EHR) adoption are well documented, as are the associated clinician burnout and negative consequences for direct patient care and the patient experience. In response, ambient artificial intelligence (AI) scribes have emerged as promising transformative technologies. These technologies aim to listen to patient-practitioner conversations during clinic visits or other synchronous encounters; then, they generate clinical notes for health care practitioner review, revision, and approval. AI scribes are still in the early stages of adoption and evaluation; however, they are already seen operationally as powerful tools that could combat administrative burdens and clinician burnout. Accelerating efforts to install AI scribes in clinical practices are taking place, with a backdrop of long-standing efforts to automate all aspects of documentation workload, which typically includes clinical note generation and other forms of administrative burden, such as preparation of computerized order entry, prior authorization forms, and medical assessments according to structured requirements, as well as coding and billing [1]. In this editorial, we describe our interpretations of the current landscape of ambient AI scribe technology and opportunities

for further research and publication, as a part of the *JMIR Medical Informatics* call for papers on “Ambient AI Scribes and AI-Driven Documentation Technologies” [2].

## Hype and Hope of Ambient AI Scribes

The conceptualization of ambient clinical documentation has evolved in parallel with the technology over the past several years (Figure 1), with a clear bibliometric trend of increasing research on the subject. There is no taxonomy of AI scribe technologies, although one appears to be emerging. One general operating definition of “digital scribe” is the use of automatic speech recognition technology or natural language processing to support clinical documentation [3]. Although the digital scribe concept does not explicitly exclude the use of AI technology, it is only in more recent years that AI has been explicitly labeled as a component of clinical documentation tools. One published literature review subdivided AI-driven documentation systems into generative AI and ambient AI, even if ambient AI may also make use of architecture comparable to that of generative AI [4].

**Figure 1.** Terminology used for artificial intelligence (AI)-based documentation technologies, based on published literature over the past decade [5-18]. \*Transcription software in this period was based on different types of AI models, including one or more of the following: automated speech recognition, natural language processing, probabilistic graphical models (conditional random field), or logistic regression models [5].



Coiera et al [19] envisioned a progression through the following three stages: (1) human-led systems augmented by tools like dictation and templates (2); mixed-initiative systems where AI assists in converting conversations into summaries; and (3) computer-led systems that autonomously handle documentation, seeking human input only for exceptions. It would seem that in the last 3 years, health care organizations and physicians have been firmly shifting into the second stage, driven by the increasingly advanced technologies available and the salient burdens on clinical and patient care time, particularly in the United States, which has generated the largest proportion of published original

research studies on ambient AI scribes. Professional societies and physician groups around the world are also engaged in the global dialogue on the role and promise of AI scribes in medical practice [20-22]. Ultimately, intelligent clinical environments, which capture and integrate data into the EHR, offer the promise of off-loading human-led documentation tasks onto a machine and doing so with an ability to integrate multimodal data from various sources, thereby freeing up the human in the loop and allowing them to focus cognitive efforts on clinical and medical decision-making. So far, initial evaluations have come from small-scale, short-term

pilot studies [12,13] that often have volunteer participants who may be biased toward technology.

This excitement has no doubt been spurred by the introduction and widespread availability of large language models (LLMs), which, in terms of development and adoption, have rapidly outpaced other AI technologies underlying documentation support. Numerous industry products [1,23] and a rapidly increasing number of publications have emerged regarding ambient AI scribes using LLMs, even though other AI technologies for documentation tasks have been studied and used for the last 2 decades. A simple PubMed literature search on June 1, 2025, resulted in the retrieval of 940 potentially relevant articles published in the last 10 years (Multimedia Appendix 1); of the most relevant articles (ie, based on their titles), 7 have been published in 2025 (as of the search date) [12,13,17,18,23-25], 8 were published in 2024 [14,15,26-31], and 3 were published in 2023 [10,32,33]. Since this search, further studies have been published to guide ambient digital scribe evaluation [5,34]. During the course of preparing this editorial, authors identified more peer-reviewed literature on ambient AI scribes every few days. Undoubtedly, additional research is forthcoming as hype, hope, and operational needs coincide to drive further adoption.

Despite the rapidly growing published literature on ambient AI scribes and AI-driven documentation, we found that many still focus on similar objectives and evaluation metrics. Consequently, we felt that a call for papers on the topic in *JMIR Medical Informatics* would be valuable for collecting and publishing scientific studies and evidence-based perspectives on broader aspects of ambient AI scribe technologies. As a starting point, we synthesized recent literature about their impact on clinical workflows, well-being, note quality, user experience, patient interaction, and medicolegal aspects, identifying opportunities for further investigation in this field (Multimedia Appendices 2-4).

## Ambient AI Scribe Opportunities

The enthusiasm surrounding AI scribes is tempered by caution. An initial focus of evaluating the outputs of AI scribes is the quality of the note [4], which is assessed by using instruments such as the Physician Documentation Quality Instrument (PDQI-9) [35] or Sheffield Assessment Instrument for Letters (SAIL) [36]; usability is assessed with the NASA Task Load Index (NASA-TLX) [37], and burnout is assessed with various inventories. Results suggest that there is risk for hallucinations or fictitious information [4]. Other evaluations of stand-alone tools for audio recordings suggest potentially high rates of errors, including incorrect information, omissions, and hallucinations [11,29,38]. The high rates of omissions and hallucinations found in some studies underscore the need to evaluate for potential diagnostic errors or other long-term safety risks stemming from AI-generated content [4,10,29-32,39-42]. Others express caution regarding the potential for security risks or risks to medical decision-making in cases where LLMs can access and modify sensitive patient data [31]. Overviews on the topic emphasize the need

for careful consideration of the ethical and practical integration of LLMs into clinical practice [1,4], cautioning about potential risks, such as automation bias, privacy concerns, and medicolegal implications [19].

Another major unexamined area is the impact on tangible clinical outcomes and patient safety. Although efficiency, time savings, and productivity are important, there are other measures that are also important. Assessments of note quality, accuracy, and impact on patient safety are essential; Gellert [32] raised a crucial point—there is currently no systematic data collection for evaluating the extent to which clinical errors or negative patient outcomes can be attributed to the use of medical scribes. More comprehensive evaluations of the safe and effective implementation of ambient AI scribe technologies (eg, along the dimensions of the seminal sociotechnological model of health information technology) are still lacking [43]. Additional diversity of clinical specialty applications, clinical disciplines, and practice settings also would be insightful. We identified studies published in a dermatology and urology journal but were unable to retrieve the full-text articles. One pediatrics application indicated positive outcomes during a digital AI scribe pilot in an outpatient pediatric setting [44]. Further, two conference proceeding papers described LLM applications in nursing documentation [27,45], as did a position paper from the Nursing and Artificial Intelligence Leadership Collaborative regarding multimodal LLM support for nursing documentation [46], although these stopped short of discussing AI scribe applications.

The most frequently mentioned gap is the limited study of patients' or caregivers' perspectives regarding AI scribes. Most studies rely on clinicians' perceptions of the patient experience, with very few directly capturing patient viewpoints. Additional research could directly measure patients' experiences, preferences, and patient-reported outcomes [4,47], especially given recent studies solely examining physicians' experiences of the patient-physician visit. Pelletier et al [44] incorporated the assessment of caregiver satisfaction in pediatrics, with the sole statistically significant finding being that caregivers' "provider-specific likelihood to recommend" was higher after the pilot digital scribe implementation. There are also potential patient benefits that remain unexplored. AI scribes may prove useful in providing rapid, patient-friendly visit summaries; outlining the diagnosis and management plan; scheduling appointments; and providing reasons to seek follow-up care [31]. They might also help to bridge communication gaps by highlighting misunderstandings or discrepancies between patient-reported details and those documented in EHRs [45].

The downstream effects of AI-generated notes on clinical communication and reasoning are also unknown. Note bloat—the well-established phenomenon of creating lengthy clinical documentation, which is most often attributed to the copy-paste phenomenon of digitizing documentation [42,48,49]—may also result from AI scribe use; whether large quantities of text redundancy change as a result of AI scribe use is unknown. With regard to clinical reasoning and cognition, as per preliminary evidence from nonmedical

studies, such as a study by Kosmyrna et al [50], one potential consequence of using generative AI for a writing task is “cognitive debt,” where repeated reliance on AI for cognitive tasks may lead to the atrophy of critical thinking and memory skills. The study found that LLM use impaired memory recall and reduced the brain’s neural engagement with the studied task. Although potentially dull, the documentation process may have beneficial effects on retention and memory. Learner experiences and consequences on expertise development, such as the formative diagnostic process for trainees, are underexplored [51].

Finally, system-level and economic outcomes require more rigorous investigation. Studies call for comprehensive cost-benefit analyses to justify the significant expense of AI scribe technologies [17]. On a broader scale, Gellert [32] cautioned about a largely unstudied systemic risk—the widespread adoption of scribes may decouple physicians from their EHRs and thereby impede the user-driven feedback necessary for the long-term evolution of next-generation clinical AI. In clinical informatics networks, user engagement in the co-design and development of AI scribes is seen as an essential component of the appropriate advancement and adoption of the technology, yet only one study has taken

steps to pursue this ideal in evaluating primary care physician needs [9].

## Conclusions

Ambient AI scribes, particularly with the widespread availability of LLMs, offer potential solutions to a previously difficult-to-bridge technological gap in clinical documentation. Although many health systems and physicians are welcoming this potentially practice-changing technology, there are many questions and areas that remain to be fully understood. Substantial clinical documentation also occurs adjacent to or outside of a clinical visit, and the applications of ambient AI scribes and AI-driven documentation technologies in these areas are yet to be explored. Such areas involve various clinician types and health professionals, clinical settings, or community settings. As editors, we look forward to the opportunity to advance the science and understanding of these fields through publishing high-quality and rigorous scholarly work in the *JMIR Medical Informatics* call for papers on “Ambient AI Scribes and AI-Driven Documentation Technologies” [2].

---

### Acknowledgments

Gemini 2.5 Pro and NotebookLM (Google) were used to help with article summarization and synthesis tasks during manuscript preparation. The prompts and responses used for the summarization tasks are available in [Multimedia Appendices 2 and 3](#). Limited reuse of summarized text has been incorporated into portions of the manuscript, with manual editing and revisions by authors for clarity and style. Authors manually verified all claims and citations and are accountable for the content of this manuscript.

---

### Data Availability

Data sharing is not applicable as no datasets were generated or analyzed during the preparation of this manuscript.

---

### Authors' Contributions

Conceptualization: TIL, AJC

Supervision: TIL, AB

Writing – original draft: TIL, AJC

Writing – review & editing: TIL, AJC, AB

---

### Conflicts of Interest

TIL is the scientific editorial director at JMIR Publications and a director on the Board of Directors, American Medical Informatics Association. AJC is a scientific editor at JMIR Publications and the editor in chief of *JMIR Cardio*. AB is the editor in chief of *JMIR Medical Informatics*.

---

### Multimedia Appendix 1

PubMed search and Gemini 2.5 Pro prompts and responses.

[\[DOCX File \(Microsoft Word File\), 25 KB-Multimedia Appendix 1\]](#)

---

### Multimedia Appendix 2

NotebookLM prompts and responses.

[\[DOCX File \(Microsoft Word File\), 29 KB-Multimedia Appendix 2\]](#)

---

### Multimedia Appendix 3

NotebookLM mind map: challenges and limitations.

[\[PNG File \(Portable Network Graphics File\), 1070 KB-Multimedia Appendix 3\]](#)

---

### Multimedia Appendix 4

NotebookLM mind map: future directions and recommendations.

[PNG File (Portable Network Graphics File), 1153 KB-Multimedia Appendix 4]

## References

1. Kunze KN, Bepple J, Bedi A, Ramkumar PN, Pean CA. Commercial products using generative artificial intelligence include ambient scribes, automated documentation and scheduling, revenue cycle management, patient engagement and education, and prior authorization platforms. *Arthroscopy*. May 24, 2025;S0749-8063(25)00397-4. [doi: [10.1016/j.arthro.2025.05.021](https://doi.org/10.1016/j.arthro.2025.05.021)] [Medline: [40419172](https://pubmed.ncbi.nlm.nih.gov/40419172/)]
2. Call for papers - Theme Issue: Ambient AI Scribes and AI-Driven Documentation Technologies. *JMIR Medical Informatics*. URL: <https://medinform.jmir.org/announcements/601> [Accessed 2025-07-11]
3. van Buchem MM, Boosman H, Bauer MP, Kant IMJ, Cammel SA, Steyerberg EW. The digital scribe in clinical practice: a scoping review and research agenda. *NPJ Digit Med*. Mar 26, 2021;4(1):57. [doi: [10.1038/s41746-021-00432-5](https://doi.org/10.1038/s41746-021-00432-5)] [Medline: [33772070](https://pubmed.ncbi.nlm.nih.gov/33772070/)]
4. Bracken A, Reilly C, Feeley A, Sheehan E, Merghani K, Feeley I. Artificial intelligence (AI) - powered documentation systems in healthcare: a systematic review. *J Med Syst*. Feb 18, 2025;49(1):28. [doi: [10.1007/s10916-025-02157-4](https://doi.org/10.1007/s10916-025-02157-4)] [Medline: [39966286](https://pubmed.ncbi.nlm.nih.gov/39966286/)]
5. Ng JJW, Wang E, Zhou X, et al. Evaluating the performance of artificial intelligence-based speech recognition for clinical documentation: a systematic review. *BMC Med Inform Decis Mak*. Jul 1, 2025;25(1):236. [doi: [10.1186/s12911-025-03061-0](https://doi.org/10.1186/s12911-025-03061-0)] [Medline: [40598136](https://pubmed.ncbi.nlm.nih.gov/40598136/)]
6. Quiroz JC, Laranjo L, Kocaballi AB, Berkovsky S, Rezazadegan D, Coiera E. Challenges of developing a digital scribe to reduce clinical documentation burden. *NPJ Digit Med*. Nov 22, 2019;2(1):114. [doi: [10.1038/s41746-019-0190-1](https://doi.org/10.1038/s41746-019-0190-1)] [Medline: [31799422](https://pubmed.ncbi.nlm.nih.gov/31799422/)]
7. Coiera E, Liu S. Evidence synthesis, digital scribes, and translational challenges for artificial intelligence in healthcare. *Cell Rep Med*. Dec 20, 2022;3(12):100860. [doi: [10.1016/j.xcrm.2022.100860](https://doi.org/10.1016/j.xcrm.2022.100860)] [Medline: [36513071](https://pubmed.ncbi.nlm.nih.gov/36513071/)]
8. Lin SY, Shanafelt TD, Asch SM. Reimagining clinical documentation with artificial intelligence. *Mayo Clin Proc*. May 2018;93(5):563-565. [doi: [10.1016/j.mayocp.2018.02.016](https://doi.org/10.1016/j.mayocp.2018.02.016)] [Medline: [29631808](https://pubmed.ncbi.nlm.nih.gov/29631808/)]
9. Kocaballi AB, Ijaz K, Laranjo L, et al. Envisioning an artificial intelligence documentation assistant for future primary care consultations: A co-design study with general practitioners. *J Am Med Inform Assoc*. Nov 1, 2020;27(11):1695-1704. [doi: [10.1093/jamia/ocaa131](https://doi.org/10.1093/jamia/ocaa131)] [Medline: [32845984](https://pubmed.ncbi.nlm.nih.gov/32845984/)]
10. Yim WW, Fu Y, Ben Abacha A, Snider N, Lin T, Yetisgen M. Aci-bench: a novel ambient clinical intelligence dataset for benchmarking automatic visit note generation. *Sci Data*. Sep 6, 2023;10(1):586. [doi: [10.1038/s41597-023-02487-3](https://doi.org/10.1038/s41597-023-02487-3)] [Medline: [37673893](https://pubmed.ncbi.nlm.nih.gov/37673893/)]
11. Balloch J, Sridharan S, Oldham G, et al. Use of an ambient artificial intelligence tool to improve quality of clinical documentation. *Future Healthc J*. Jun 26, 2024;11(3):100157. [doi: [10.1016/j.fhj.2024.100157](https://doi.org/10.1016/j.fhj.2024.100157)] [Medline: [39371531](https://pubmed.ncbi.nlm.nih.gov/39371531/)]
12. Misurac J, Knake LA, Blum JM. The effect of ambient artificial intelligence notes on provider burnout. *Appl Clin Inform*. Mar 2025;16(2):252-258. [doi: [10.1055/a-2461-4576](https://doi.org/10.1055/a-2461-4576)] [Medline: [39500346](https://pubmed.ncbi.nlm.nih.gov/39500346/)]
13. Stults CD, Deng S, Martinez MC, et al. Evaluation of an ambient artificial intelligence documentation platform for clinicians. *JAMA Netw Open*. May 1, 2025;8(5):e258614. [doi: [10.1001/jamanetworkopen.2025.8614](https://doi.org/10.1001/jamanetworkopen.2025.8614)] [Medline: [40314951](https://pubmed.ncbi.nlm.nih.gov/40314951/)]
14. Owens LM, Wilda JJ, Hahn PY, Koehler T, Fletcher JJ. The association between use of ambient voice technology documentation during primary care patient encounters, documentation burden, and provider burnout. *Fam Pract*. Apr 15, 2024;41(2):86-91. [doi: [10.1093/fampra/cmada092](https://doi.org/10.1093/fampra/cmada092)] [Medline: [37672297](https://pubmed.ncbi.nlm.nih.gov/37672297/)]
15. Galloway JL, Munroe D, Vohra-Khullar PD, et al. Impact of an artificial intelligence-based solution on clinicians' clinical documentation experience: initial findings using ambient listening technology. *J Gen Intern Med*. Oct 2024;39(13):2625-2627. [doi: [10.1007/s11606-024-08924-2](https://doi.org/10.1007/s11606-024-08924-2)] [Medline: [38980463](https://pubmed.ncbi.nlm.nih.gov/38980463/)]
16. Ma SP, Liang AS, Shah SJ, et al. Ambient artificial intelligence scribes: utilization and impact on documentation time. *J Am Med Inform Assoc*. Feb 1, 2025;32(2):381-385. [doi: [10.1093/jamia/ocae304](https://doi.org/10.1093/jamia/ocae304)] [Medline: [39688515](https://pubmed.ncbi.nlm.nih.gov/39688515/)]
17. Shah SJ, Devon-Sand A, Ma SP, et al. Ambient artificial intelligence scribes: physician burnout and perspectives on usability and documentation burden. *J Am Med Inform Assoc*. Feb 1, 2025;32(2):375-380. [doi: [10.1093/jamia/ocae295](https://doi.org/10.1093/jamia/ocae295)] [Medline: [39657021](https://pubmed.ncbi.nlm.nih.gov/39657021/)]
18. Shah SJ, Crowell T, Jeong Y, et al. Physician perspectives on ambient AI scribes. *JAMA Netw Open*. Mar 3, 2025;8(3):e251904. [doi: [10.1001/jamanetworkopen.2025.1904](https://doi.org/10.1001/jamanetworkopen.2025.1904)] [Medline: [40126477](https://pubmed.ncbi.nlm.nih.gov/40126477/)]
19. Coiera E, Kocaballi B, Halamka J, Laranjo L. The digital scribe. *NPJ Digit Med*. Oct 16, 2018;1:58. [doi: [10.1038/s41746-018-0066-9](https://doi.org/10.1038/s41746-018-0066-9)] [Medline: [31304337](https://pubmed.ncbi.nlm.nih.gov/31304337/)]
20. Guidance on the use of AI-enabled ambient scribing products in health and care settings. NHS England. URL: <https://www.england.nhs.uk/long-read/guidance-on-the-use-of-ai-enabled-ambient-scribing-products-in-health-and-care-settings/> [Accessed 2025-07-12]

21. Shemtob L, Majeed A, Beaney T. Regulation of AI scribes in clinical practice. *BMJ*. Jun 20, 2025;389:r1248. [doi: [10.1136/bmj.r1248](https://doi.org/10.1136/bmj.r1248)] [Medline: [40541424](https://pubmed.ncbi.nlm.nih.gov/40541424/)]
22. AI scribes. OntarioMD. URL: <https://www.ontariomd.ca/pages/ai-scribe-overview.aspx> [Accessed 2025-07-12]
23. Blaseg E, Huffstetler A. Artificial intelligence scribes shape health care delivery. *Am Fam Physician*. Apr 2025;111(4):304-305. [Medline: [40238969](https://pubmed.ncbi.nlm.nih.gov/40238969/)]
24. Schaye V, DiTullio D, Guzman BV, et al. Large language model-based assessment of clinical reasoning documentation in the electronic health record across two institutions: development and validation study. *J Med Internet Res*. Mar 21, 2025;27:e67967. [doi: [10.2196/67967](https://doi.org/10.2196/67967)] [Medline: [40117575](https://pubmed.ncbi.nlm.nih.gov/40117575/)]
25. Duggan MJ, Gervase J, Schoenbaum A, et al. Clinician experiences with ambient scribe technology to assist with documentation burden and efficiency. *JAMA Netw Open*. Feb 3, 2025;8(2):e2460637. [doi: [10.1001/jamanetworkopen.2024.60637](https://doi.org/10.1001/jamanetworkopen.2024.60637)] [Medline: [39969880](https://pubmed.ncbi.nlm.nih.gov/39969880/)]
26. Bundy H, Gerhart J, Baek S, et al. Can the administrative loads of physicians be alleviated by AI-facilitated clinical documentation? *J Gen Intern Med*. Nov 2024;39(15):2995-3000. [doi: [10.1007/s11606-024-08870-z](https://doi.org/10.1007/s11606-024-08870-z)] [Medline: [38937369](https://pubmed.ncbi.nlm.nih.gov/38937369/)]
27. Chen CJ, Liao CT, Tung YC, Liu CF. Enhancing healthcare efficiency: integrating ChatGPT in nursing documentation. *Stud Health Technol Inform*. Aug 22, 2024;316:851-852. [doi: [10.3233/SHTI240545](https://doi.org/10.3233/SHTI240545)] [Medline: [39176926](https://pubmed.ncbi.nlm.nih.gov/39176926/)]
28. Huang TY, Hsieh PH, Chang YC. Performance comparison of junior residents and ChatGPT in the Objective Structured Clinical Examination (OSCE) for medical history taking and documentation of medical records: development and usability study. *JMIR Med Educ*. Nov 21, 2024;10(1):e59902. [doi: [10.2196/59902](https://doi.org/10.2196/59902)] [Medline: [39622713](https://pubmed.ncbi.nlm.nih.gov/39622713/)]
29. Kernberg A, Gold JA, Mohan V. Using ChatGPT-4 to create structured medical notes from audio recordings of physician-patient encounters: comparative study. *J Med Internet Res*. Apr 22, 2024;26:e54419. [doi: [10.2196/54419](https://doi.org/10.2196/54419)] [Medline: [38648636](https://pubmed.ncbi.nlm.nih.gov/38648636/)]
30. Seo J, Choi D, Kim T, et al. Evaluation framework of large language models in medical documentation: development and usability study. *J Med Internet Res*. Nov 20, 2024;26:e58329. [doi: [10.2196/58329](https://doi.org/10.2196/58329)] [Medline: [39566044](https://pubmed.ncbi.nlm.nih.gov/39566044/)]
31. Tripathi S, Sukumaran R, Cook TS. Efficient healthcare with large language models: optimizing clinical workflow and enhancing patient care. *J Am Med Inform Assoc*. May 20, 2024;31(6):1436-1440. [doi: [10.1093/jamia/ocad258](https://doi.org/10.1093/jamia/ocad258)] [Medline: [38273739](https://pubmed.ncbi.nlm.nih.gov/38273739/)]
32. Gellert GA. Medical scribes: symptom or cause of impeded evolution of a transformative artificial intelligence in the electronic health record? *Perspect Health Inf Manag*. Jan 10, 2023;20(1):1d. [Medline: [37215336](https://pubmed.ncbi.nlm.nih.gov/37215336/)]
33. Socrates V, Gilson A, Lopez K, Chi L, Taylor RA, Chartash D. Predicting relations between SOAP note sections: the value of incorporating a clinical information model. *J Biomed Inform*. May 2023;141:104360. [doi: [10.1016/j.jbi.2023.104360](https://doi.org/10.1016/j.jbi.2023.104360)] [Medline: [37061014](https://pubmed.ncbi.nlm.nih.gov/37061014/)]
34. Wang H, Yang R, Alwakeel M, et al. An evaluation framework for ambient digital scribing tools in clinical applications. *NPJ Digit Med*. Jun 13, 2025;8(1):358. [doi: [10.1038/s41746-025-01622-1](https://doi.org/10.1038/s41746-025-01622-1)] [Medline: [40514413](https://pubmed.ncbi.nlm.nih.gov/40514413/)]
35. Stetson PD, Bakken S, Wrenn JO, Siegler EL. Assessing electronic note quality using the Physician Documentation Quality Instrument (PDQI-9). *Appl Clin Inform*. 2012;3(2):164-174. [doi: [10.4338/aci-2011-11-ra-0070](https://doi.org/10.4338/aci-2011-11-ra-0070)] [Medline: [22577483](https://pubmed.ncbi.nlm.nih.gov/22577483/)]
36. Crossley GM, Howe A, Newble D, Jolly B, Davies HA. Sheffield Assessment Instrument for Letters (SAIL): performance assessment using outpatient letters. *Med Educ*. Dec 2001;35(12):1115-1124. [doi: [10.1046/j.1365-2923.2001.01065.x](https://doi.org/10.1046/j.1365-2923.2001.01065.x)] [Medline: [11895235](https://pubmed.ncbi.nlm.nih.gov/11895235/)]
37. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. *Advances in Psychology*. 1988;52:139-183. [doi: [10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)]
38. Biro J, Handley JL, Cobb NK, et al. Accuracy and safety of AI-enabled scribe technology: instrument validation study. *J Med Internet Res*. Jan 27, 2025;27:e64993. [doi: [10.2196/64993](https://doi.org/10.2196/64993)] [Medline: [39869899](https://pubmed.ncbi.nlm.nih.gov/39869899/)]
39. Chelli M, Descamps J, Lavoué V, et al. Hallucination rates and reference accuracy of ChatGPT and Bard for systematic reviews: comparative analysis. *J Med Internet Res*. May 22, 2024;26:e53164. [doi: [10.2196/53164](https://doi.org/10.2196/53164)] [Medline: [38776130](https://pubmed.ncbi.nlm.nih.gov/38776130/)]
40. Metz C. Chatbots may 'hallucinate' more often than many realize. *The New York Times*. Nov 6, 2023. URL: <https://www.nytimes.com/2023/11/06/technology/chatbots-hallucination-rates.html> [Accessed 2025-07-07]
41. Hatem R, Simmons B, Thornton JE. A call to address AI "hallucinations" and how healthcare professionals can mitigate their risks. *Cureus*. Sep 5, 2023;15(9):e44720. [doi: [10.7759/cureus.44720](https://doi.org/10.7759/cureus.44720)] [Medline: [37809168](https://pubmed.ncbi.nlm.nih.gov/37809168/)]
42. Liu J, Capurro D, Nguyen A, Verspoor K. "Note bloat" impacts deep learning-based NLP models for clinical prediction tasks. *J Biomed Inform*. Sep 2022;133:104149. [doi: [10.1016/j.jbi.2022.104149](https://doi.org/10.1016/j.jbi.2022.104149)] [Medline: [35878821](https://pubmed.ncbi.nlm.nih.gov/35878821/)]
43. Sittig DF, Singh H. A new sociotechnical model for studying health information technology in complex adaptive healthcare systems. *Qual Saf Health Care*. Oct 2010;19 Suppl 3(Suppl 3):i68-i74. [doi: [10.1136/qshe.2010.042085](https://doi.org/10.1136/qshe.2010.042085)] [Medline: [20959322](https://pubmed.ncbi.nlm.nih.gov/20959322/)]

44. Pelletier JH, Watson K, Michel J, McGregor R, Rush SZ. Effect of a generative artificial intelligence digital scribe on pediatric provider documentation time, cognitive burden, and burnout. *JAMIA Open*. Jul 3, 2025;8(4):ooaf068. [doi: [10.1093/jamiaopen/ooaf068](https://doi.org/10.1093/jamiaopen/ooaf068)] [Medline: [40620477](https://pubmed.ncbi.nlm.nih.gov/40620477/)]
45. Thawinwisan N, Liu C, Kishimoto K, Yamamoto G, Mori Y, Kuroda T. Comparing patient perception and physician's records: generative AI performance evaluation. *Stud Health Technol Inform*. Aug 22, 2024;316:671-675. [doi: [10.3233/SHTI240503](https://doi.org/10.3233/SHTI240503)] [Medline: [39176831](https://pubmed.ncbi.nlm.nih.gov/39176831/)]
46. Michalowski M, Topaz M, Peltonen LM. An AI-enabled nursing future with no documentation burden: a vision for a new reality. *J Adv Nurs*. Mar 24, 2025. [doi: [10.1111/jan.16911](https://doi.org/10.1111/jan.16911)] [Medline: [40129115](https://pubmed.ncbi.nlm.nih.gov/40129115/)]
47. Sasseville M, Yousefi F, Ouellet S, et al. The impact of AI scribes on streamlining clinical documentation: a systematic review. *Healthcare (Basel)*. Jun 16, 2025;13(12):1447. [doi: [10.3390/healthcare13121447](https://doi.org/10.3390/healthcare13121447)] [Medline: [40565474](https://pubmed.ncbi.nlm.nih.gov/40565474/)]
48. Rule A, Bedrick S, Chiang MF, Hribar MR. Length and redundancy of outpatient progress notes across a decade at an academic medical center. *JAMA Netw Open*. Jul 1, 2021;4(7):e2115334. [doi: [10.1001/jamanetworkopen.2021.15334](https://doi.org/10.1001/jamanetworkopen.2021.15334)] [Medline: [34279650](https://pubmed.ncbi.nlm.nih.gov/34279650/)]
49. Thornton JD, Schold JD, Venkateshaiah L, Lander B. Prevalence of copied information by attendings and residents in critical care progress notes. *Crit Care Med*. Feb 2013;41(2):382-388. [doi: [10.1097/CCM.0b013e3182711a1c](https://doi.org/10.1097/CCM.0b013e3182711a1c)] [Medline: [23263617](https://pubmed.ncbi.nlm.nih.gov/23263617/)]
50. Kosmyna N, Hauptmann E, Yuan YT, et al. Your brain on ChatGPT: accumulation of cognitive debt when using an AI assistant for essay writing task. *arXiv*. Preprint posted online on Jun 10, 2025. [doi: [10.48550/arXiv.2506.08872](https://doi.org/10.48550/arXiv.2506.08872)]
51. Wright DS, Kanaparthi N, Melnick ER, et al. The effect of ambient artificial intelligence scribes on trainee documentation burden. *Appl Clin Inform*. Jul 2, 2025. [doi: [10.1055/a-2647-1142](https://doi.org/10.1055/a-2647-1142)] [Medline: [40602775](https://pubmed.ncbi.nlm.nih.gov/40602775/)]

## Abbreviations

- AI:** artificial intelligence  
**EHR:** electronic health record  
**LLM:** large language model  
**NASA-TLX:** NASA Task Load Index  
**PDQI-9:** Physician Documentation Quality Instrument  
**SAIL:** Sheffield Assessment Instrument for Letters

*Edited by Tiffany Leung; This is a non-peer-reviewed article; submitted 18.07.2025; accepted 18.07.2025; published 01.08.2025*

*Please cite as:*

Leung TI, Coristine AJ, Benis A  
*AI Scribes in Health Care: Balancing Transformative Potential With Responsible Integration*  
*JMIR Med Inform* 2025;13:e80898  
URL: <https://medinform.jmir.org/2025/1/e80898>  
doi: [10.2196/80898](https://doi.org/10.2196/80898)

© Tiffany I Leung, Andrew J Coristine, Arriel Benis. Originally published in *JMIR Medical Informatics* (<https://medinform.jmir.org>), 01.08.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Medical Informatics*, is properly cited. The complete bibliographic information, a link to the original publication on <https://medinform.jmir.org/>, as well as this copyright and license information must be included.