Can AI Hear Me? Nine Facts About the Artificial Intelligence Scribe

Laura Solano DMSc, MPAS, PA-C, **Robert Smith** DHSc, MS, PA-C, DFAAPA

Corresponding Author

Laura Solano
A.T. Still University
orcid.org/0000-0003-4910-2756

ABSTRACT

This article summarizes important features of artificial intelligence scribe (AIS) and its application as a digital scribe through the utilization of voice recognition as it is integrated into the clinical practice of medicine. This discussion will include 1) digital scribe technology, 2) practical utility, 3) liability, 4) privacy, 5) integration to electronic health records (EHR), 6) technology burden, 7) case use, 8) AIS documentation styles, and 9) cost. The use of AIS technology in healthcare is a promising tool and will result in variability applications requiring healthcare providers to be aware of the complexities of the tool before utilization, during implementation, the effect of it on workflow, and its acceptance by the patients we serve. AIS technology promises to decrease the cognitive workload of the healthcare provider, through application by relieving the ever-increasing burden of documenting the medical record, organizing, and capturing relevant medical data, by improving integration of this data into cumbersome EHRs, thus providing better work life balance. The authors piloted a version of an AIS in a primarily Spanish patient population. This article provides a unique perspective towards the usability of an AIS in a Hispanic patient population.

Keywords: *digital scribe, artificial intelligence, virtual scribe, speech recognition, voice recognition, e-scribe, EHR integration*

1. Introduction

As the documentation burden has increased so has the demand and pressure for healthcare providers to provide evidence of the care being provided. The need to document the level of care, the quality measures being applied, patient education, orders for ancillary services, referrals, and the integration of multiple sources of medical information all being used in decision making model for payment has made it imperative that healthcare providers seek alternative ways to capture and document the care they provide. These dynamics place healthcare providers into a documentation paradox; to document a medical encounter comprehensively and ethically, bill correctly, avoid potential liabilities, all in an efficient and timely manner that allows the provider to interact and bond with the patient.

There have been several solutions initiated to address this issue, for example the use of transcription, while it made the charts more comprehensive and easier to read, it did little to reduce the burdens described. More recently was the integration and use of human scribes as a methodology initiated to address this demand, and while the overall application of this methodology has found that use of human scribes can be effective, that many providers like the premise and the intended relief offered due to more face time with the patient, less time spent documenting; however, the application and wide utilization are limited because of cost and potential return on investment (ROI).[1,2]

This article will focus on the use of the AIS. This virtual scribe is a scribe that utilizes speech/voice recognition technology to "listen" to in-clinic conversations utilizing artificial intelligence to organize the conversation into data format that is useful medical data: (history of present illness, review of systems, physical exams, treatment plan, and diagnosis). Gellert wrote that evolving artificial intelligence will help make the practice of medicine more

satisfying and can potentially eliminate the EHR as a contributor to burnout.[3] A national poll revealed that AI-assisted documentation may be beneficial in that it frees clinicians from less cognitively useful tasks.[4] Ghatnekar, Faletsky, and Nambudiri cited barriers to this technology including upfront costs, time-intensive training, linguistic variations, medical-legal compliance, and interoperability with existing computer systems.[5]

The authors of this article piloted an AIS in two clinics in Texas with approximately 120 + patients a week and with Telemedicine visits. The authors work in primary care and deliver patient care to family practice patients and have an opportunity to utilize an AIS during busy days with acute and chronic cases. Since the population was largely Hispanic (and the authors are two native Spanish speakers) the entire article stems from the authors' experience with a pilot of an AIS in a Hispanic population. Therefore, the article provides a unique perspective regarding the clinical utility of the AIS when working with Hispanic patients.

1.1 Technology

Speech recognition (SR) and Voice Recognition (VR) as an input mechanism that translates speech into text.[6] AISs are intelligent documentation assistants that utilize SR and VR, natural language processing, machine learning (ML), and large language models (LLM) to document clinical encounters. Tran et al share that Ambient clinical documentation technology uses automatic speech recognition (ASR) a combination of SR/VR and natural language processing (NLP) to transform patient–clinician conversations into clinical documentation.[7] NLP is a technology that allows computer software to understand human language. Powered by ambient technology the digital scribe can provide a documentation recommendation after 'listening" to a clinical encounter with a patient. A digital scribe has been defined by Falcetta et al as an automated clinical documentation system that capture the physician–patient conversation and then generate the documentation for the appointment, enabling the clinician to engage with the patient entirely.[8] The speech and voice recognition program can differentiate and understand who is speaking (provider patient/caregiver) and has the capacity to understand multiple languages including Spanish, translating the entire clinic visit into an English format for EHR records. Additionally, the use of the AIS is efficient in translating Spanish. Artificial intelligence driven digital scribes have the ability to listen, "understand," and translate a visit conducted in Spanish into English. Furthermore, in a situation where a human translator is involved, the clinician has AIS functions as a tool to check the veracity and or accuracy of the human translator. AIS may improve the quality of translation in addition to scribing and medical note creation. This is an important aspect of the technology that is useful for the Hispanic population.

1.2 Practical Utility

The AIS concept is presenting itself at a time when documentation requirements are increasing, EHR learning curves are becoming more complex, and change is part of the clinician job description. According to the American Academy of Family Practice, there are different methodologies to assist clinicians in completing documentation: dictation systems, transcription services, ambient systems, virtual human scribes, live human scribes, and artificial intelligence assistance. [9] The task of AIS is to essentially reduce the clinicians' burden of documentation while meeting the complexities and needs of healthcare providers in the medical industry. This basic process could be easily accomplished by non-medical systems such as a dictation service or a transcription service. However, the potential benefit of artificial intelligence scribing is that it can organize the accrued data into medical notes. This digital scribe aims to meet the challenge of emerging divergence of many patient complaints and assist the healthcare provider in completing the documentation requirements while reducing the burden.

The AIS intends to both reduce the documentation burden of the clinician and facilitate efficiency in completing the notes. Another benefit of the AIS is its availability for clinician use, it does not require paid time off and is available whenever the clinician is providing care (evenings, weekends, and holidays) and may provide as significant ROI. Last, the digital scribe may function as a memory aid, Wang et al highlight that "using the digital scribe to document information during patient encounters, providers may more easily recall encounter events when editing multiple notes."[8,10]

1.3 Clinician liability

EHRs are riddled with opportunities for errors. A clerical error may be a precursor to a wrong diagnosis and consequent plan including erroneous medications. Template utilization allows for preselected answers that clinicians may need to take the time to correct. A similar situation may occur for the AIS if there is an error in the format process resulting in something erroneously being heard, recognized, or translated. Similarly, errors may occur when the AIS does not distinguish who is the speaker in the recording (clinician, family member, or patient). The errors that may occur due to the AIS, speech recognition and or voice recognition can be a precursor to clinician liability.

Further, Azamfirei, et al highlighted that LLMs are prone to hallucinations, where the AIS generates data based on data sets and probabilities and cannot verify when it is wrong.[11] Therefore, in programs like Chat GPT, the AI may" hallucinate" or compose erroneous data. The authors wrote that the LLMs may not have "an output to verify correctness."[11] Also, Azamfirei, et al emphasize that it is important to recognize the limitations of all medical tools including the limitations of the AIS.[11]

In 2023, Tang et. al conducted a systematic study of large language models (LLM) the potential and possible limitations of zero-shot prompt-based LLMs on medical evidence summarization using GPT-3.5 and ChatGPT models.[12] The authors explored the summarization of medical evidence findings in the context of evidence synthesis and meta-analysis evaluated systems such as Chat GPT for errors in language models by completing a large review of medical summaries and sub-divided the errors into the following categories:

- **Factual consistency** - measures whether the statements in the summary (made by the AI) were supported by the systematic review.
- **Medical Harmfulness or Risk** - refers to the potential of a summary that leads to physical or psychological harm or unwanted changes in therapy or compliance due to the misinterpretation of information.
- **Comprehensiveness** - evaluates whether a summary contains sufficient information to cover the objectives of the systematic review.
- **Coherence** - refers to the ability of a summary to build a coherent body of information about a topic through sentence-to sentence connections.[12]

Other authors have evaluated error margins of AI-driven systems. In 2015, Hodgson and Coiera completed a study comparing speech recognition (SR) to dictated transcription (DT) evaluating document turnaround time (TAT) and error rates.[6] SR consistently improved TAT compared to DT at a greater level (16.41% to 82.34%), across all studies the improvement was 0.90% per year.[4] SR accuracy was reported in ten studies (88.90% to 96.00%) and appears to improve by 0.03% per year as the technology matured, meaning the longer it was used the greater the efficacy.[1] However, the mean number of errors per report increased using SR (0.05 to 6.66) compared to DT (0.02 to 0.40). [6]

Non-AI driven transcription services may have errors as well. Zhou et. al completed a cross-sectional study evaluating how accurately clinical documents are dictated by speech recognition software (Dragon Medical 360) edited and

reviewed by both medical transcriptionists and physicians.[13] Their findings demonstrated an overall error rate greater than 7% (7 errors per 100 words) and supported the importance of review and editing.[13] In addition, the healthcare provider should be cautious concerning automatization bias, where the information produced by a machine is not verified and it is trusted as truth.[13,14] Given the above, the clinician should always verify the accuracy of the recommendation by the digital scribe. An example is that the healthcare provider may say they are going to the Y (referring to a gym), and the AIS may type that the patient is going to Hawaii. Healthcare providers must be prudent to verify the accuracy of output of the digital scribe to be most successful. Although this type of mistake may not occur frequently the astute clinician must be aware of the possibility of this type of error. As the AIS continues to evolve the premise is that errors of transcription are possible depending on the speaker's accent or dialect, however it is anticipated that the overall the error rates will remain low and will continue decline over time with improvement technologies and application. It is possible that AI and the AIS will continue to be adopted over time.[15]

1.4 Privacy

The Health Insurance Portability and Accountability Act of 1996 (HIPAA) a federal law required the creation of national standards to protect sensitive patient health information from being disclosed without the patient's consent or knowledge. Before this period, medical records were the purview of the clinician and patient and relied on paper files to be maintained and secured by the provider. The digitization of information; word processing, faxing, EHRs, and the evolution of the World Wide Web made it necessary to create a standardized method for protecting personal information. The advent of third-party AI adds complexity to the nature and essence of the documentation of the medical visit raising concerns that include access, use, and control of patient data in private hands. These public–private partnerships for implementing AI have resulted in poor privacy protection.[16]Another concern relates to the external risk of privacy breaches through AI-driven methods, even with the ability to de-identify or anonymize patient health data it may be compromised or even nullified with new algorithms that have successfully reidentified such data which increases the risk to patient data under private custodianship.[17]

Other related concerns of privacy are the issues of a digital presence in the healthcare setting because of previous medical atrocities such as the Guatemalan Syphilis Study in 1946 by the U.S. Government that resulted in a higher level of distrust in the Latino community of the medical community than Caucasians counterparts.[18,19] Hispanics may perceive a breach of confidentiality, that somebody is listening and will use the information against them for other reasons than providing care. This is particular the concern is for patients that may be undocumented and may cause Hispanics to avoid seeking medical care. This will require organizations such as National Hispanic Medical Association (NHMA), American Medical Association (AMA), National Medical Association (NMA) and other professional groups to educate providers and the lay community about the evolution of the AIS, and its' utilization and integration into the clinical setting.

1.5 Integration to EHRs

Developing the AI systems underpinning digital scribes will require access to clinical datasets for machine learning. SR works well for note dictation, however, it is less effective when used with modern EHRs, introducing errors with the potential for patient harm. The question regarding integration directly correlates with the usability

of the tool. If the integration feature is not seamless then the clinician will spend more time migrating the AI data, increasing interactions with technology, and decreasing time spent with patients. Therefore, the integration workflow is an important consideration when considering different versions of digital scribes.

Appropriate integration of SR/VR will be a must and the need for standards will need to be established. Also, clinicians may consider "optimizing ambient AI scribe output for accuracy, relevance, and alignment in the physician–patient relationship."[20]

1.6 Technology requirements

The efficient use of the AIS requires a laptop, or applicable mobile device/tablet. This device must be durable (transported from room to room), highly mobile (small and lightweight), cost-effective, Wi-Fi and software compatible, HIPPA compliant and have long-life batteries. A high bandwidth quality internet connection is essential to efficient workflows without affecting productivity. Lieberman et al suggest that there is a need to "improve digital tools to facilitate family history collection."[21.]

There are implications with technology, cost, and computer resources in poor or rural communities. Specifically, there may be a large digital divide for access to high quality internet infrastructures (computers/laptops/electronic pads) for the Latino community. Clinicians that provide care to these populations have increased limited financial resources and challenges to provide a high level of computer generated EHR or AI supported software. Cost is a significant barrier as costs to EHRs may be limiting, for example the cost of an Epic System or similar may prevent the Latino community from access to helpful technology as the digital scribe.

1.7 Case use

There are variable applications of case use for the AIS including utilizing it as a dictation device. This speech recognition system transcribes information and functions as an artificial assistant. Utilized as an artificial assistant, the AIS may be able to complete many of the tasks the provider typically completes directly. However, the technology burden may push the clinician to choose to utilize the digital scribe in only selected instances. Further, the clinician may utilize some instances as a dictation device (speak to the AI after the patient encounter) or have the AIS listen to the entire clinical encounter to create a medical note. For example, the clinician may choose to utilize all the features of the digital scribe with a new patient or a wellness exam and may choose not to utilize the software for straightforward cases acute cases (acute nasopharyngitis). Therefore, the clinician has flexibility with regards to how and when to utilize the AIS.

1.8 Documentation styles with AI Scribing

Some versions of the allow for different formatting of notes-narrative, bulleted, or a combination. Prior notes may have been short to increase efficiency or relied heavily on templates. The new AIS notes may be too detailed, capture irrelevant facts, or document too many complaints or errors, which may affect the care being delivered. Over-documenting may in fact- increase the risk and responsibility of the clinician. Essentially, this may increase the clinicians' stress level and inadvertently increase the number of demands necessary in documenting and caring for the patient. Therefore, the clinician may choose to utilize all or only a portion of the AIS potential recommendation for documentation. Or the clinician may choose to utilize different styles of notes depending on if the AI brand/type and type of patient visit (acute/chronic).

1.9 Cost

The upfront costs associated with implementing an AIS platform in healthcare can vary widely based on the complexity and scope of the AI project. Stokes discussed several upfront costs however also discussed potential cost savings and long-term return on investment (ROI) and revenue enhancements:[15]

- **Technology Acquisition:** This includes the purchase or development of AI software and hardware infrastructure, including specialized AI algorithms, computing resources, and data storage systems.
- **Data Preparation:** Preparing and cleaning healthcare data for AI analysis can be labor-intensive. Data preparation costs may include data extraction, cleansing, and transformation.
- **Staff Training:** Healthcare professionals and IT staff may require training to effectively use AI systems. Training costs include workshops, courses, and time away from regular duties.
- **Integration**: Integrating AI systems with existing healthcare IT infrastructure, such as Electronic Health Records (EHR) systems, may require significant resources and expertise.
- **Regulatory Compliance:** Ensuring that the AI implementation complies with healthcare regulations, such as HIPAA, may entail additional costs for legal and regulatory consulting.
- **Pilot Programs:** Running pilot programs to test the AI system in a real healthcare setting may involve costs related to trial implementation and monitoring.[15]

Despite the initial cost investment, the AIS may produce savings by increasing efficiency, improving the billing process, and decreasing the costs of human scribes which can be a significant expense at $50K per application annually.[5] Ghatnekar, Faletsky, and Nambudiri discuss the high cost of in person live scribes.[5] Administrators must carefully consider the cost of the AIS vs the costs of a human scribes in the overall evaluation on the ROI. The reduction in cost of implementing an AIS savings versus cost of human scribe can include time intensive training, high turnover rates,-and other HR issues such as missing work, benefits, overtime, and paid time off.[5]-

There are some negative consequences of the AIS that include loss of function that current AIS platforms cannot perform and that human scribes can fulfill writing prescriptions, making referrals, and completing complex coding requirements that easily flow from gathered data in its current capacity. Additionally, human scribes may be able to assess the patient/provider interaction and determine authenticity and body language that affects the consultation that cannot be seen or captured by a digital platform. Over time these specific limitations may be negated with improvements in technology.

2. Discussion

Incorporating the AIS and digital scribing with SR/VR into healthcare can be transformative, however, it requires careful planning and consideration. A robust AIS decision framework serves as a guide to help healthcare stakeholders navigate the complex terrain of the AIS implementation. Considerations include addressing data quality, ethics, clinical validity, integration, cost-benefit, and training. Healthcare providers must make informed decisions about AIS's suitability for their healthcare infrastructure. The digital transformation of healthcare is inevitable. Since the AIS allows for more robust notes, utilization of this technology may (in time) become the standard. Therefore, the clinician of today may consider early adoption as a proactive approach to the possible pending avalanche of AIS ubiquity in healthcare.-Integrating the AIS into practice promises to decrease the clinician's burden of documenting the EHR record and will allow the clinician to focus more time on building patient-clinician rapport during patient encounters.

According to the author's experience with the AIS, the AIS is useful in providing care for Hispanic patients. A positive factor includes allowing the AIS to formulate a comprehensive medical report. The primary care clinician of today is tasked with increasing administrative burdens which is increasing the imbalance of work life balance. Although, the

AIS does not eradicate this burden it may reduce workflow pressures. Hispanic patients have higher burden of comorbidities such as diabetes mellitus, chronic kidney disease, hypertension, and obesity the care of this population requires significant time. Although there is a benefit to the AIS in the Latino population, cost is a barrier and may prohibit access to an AIS in underserved settings. Despite these barriers, the need for the AIS is imperative for future growth and application due to its inherent properties of increasing face to face time with the clinician that will allow more time to provide medical education to the Latino population.

Clinician liability may be decreased by clarifying to the AI who is in the room. In order to assist the AIS in distinguishing the speaker during clinical encounters through speech/voice recognition the clinician may consider introducing the participants of the encounter to the AIS in the following manner, "I am Dr. Ramirez, I will be caring for you today, and you are Mrs. Garcia the mom, and you Maria are the patient." Then, the clinician may conduct the interview.

The AIS may not necessarily improve productivity; however, studies demonstrate that there is a significant increase in quality of notes, and clinical care.[20] AIS may improve clinician satisfaction, reduce in administrative burden, and improve patient satisfaction.[20] Additional outcomes may improve value-based care and quality metrics. Future development and ongoing development of the AIS should include standardized methodology for application of the AIS modality. This may include creating a standardized level of accuracy, including a certain panel of languages, and uniform language model training. Further development of the AIS could include a robust artificial intelligence platform to create a more comprehensive standard AIS and standard medical artificial intelligence.

3. Conclusion

In conclusion, the AIS is an effective tool although it is in early development it is promising for widespread applications and clinical utilization. In our pilot project, we discovered that the AIS has comprehensive speech and voice recognition abilities that can be useful for transcription, translation, and documentation purposes when treating the Latino population with limited English-speaking abilities. In order to make AIS accessible for clinicians treating Latino and other underserved populations the AIS will require cost effective applications that provide quality and accessible care. Further studies AIS and the use of AI in clinical applications are necessary to meet the needs of the growing Latino population.

~

References

1. Ziemann M, Erikson C, Krips M. The use of medical scribes in primary care settings: a literature synthesis. Med Care. 2021;59(Suppl 5):S449-S456. doi:10.1097/MLR.0000000000001605
2. Piersa AP, Laiteerapong N, Ham SA, et al. Impact of a medical scribe on clinical efficiency and quality in an academic general internal medicine practice. BMC Health Serv Res. 2021;21(1):686. Published 2021 Jul 11. doi:10.1186/s12913-021-06710-y
3. Gellert GA. Medical sScribes: symptom or cause of impeded evolution of a transformative artificial intelligence in the electronic health record? PubMed. 2023;20(1):1d1d.
4. Hong G, Wilcox L, Sattler A, et al. Clinicians' experiences with EHR documentation and attitudes toward AI-assisted documentation (2020). med.stanford.edu/healthcare ai and health. Google.
5. Ghatnekar S, Faletsky A, Nambudiri VE. Digital scribe utility and barriers to Implementation in Clinical Practice: A Scoping Review. Health and Technology.2021;11. doi:https://doi.org/10.1007/s12553-021-00568-0

6. Hodgson T, Coiera E. Risks and benefits of speech recognition for clinical documentation: A systematic review. Journal of the American Medical Informatics Association. 2015;23(e1):e169-e179. doi:https://doi.org/10.1093/jamia/ocv152
7. Tran BD, Latif K, Reynolds TL, et al. "Mm-hm," "Uh-uh": Are non-lexical conversational sounds deal breakers for the Ambient Clinical Documentation technology? Journal of the American Medical Informatics Association. 2023;30(4):703- 711. doi:10.1093/jamia/ocad001
8. Falcetta FS, de Almeida FK, Lemos JCS, Goldim JR, da Costa CA. Automatic Documentation of Professional Health Interactions: A Systematic Review. Artificial Intelligence in Medicine. 2023;102487(137):102487.doi:https://doi.org/10.1016/j.artmed.2023.102487
9. TAFP. Using an AI assistant to reduce documentation burden in family medicine. TAFP.org. Published November 15, 2021.Accessed January 31, 2024. https://www.tafp.org/media/healthy-practice/report-suki-assistant-documentation burden.pdf
10. Wang J, Lavender M, Hoque E, Brophy P, Kautz H. A patient-centered digital scribe for automatic medical documentation. JAMIA Open. Published online February 17, 2021. doi:https://doi.org/10.1093/jamiaopen/ooab003
11. Azamfirei, R., Kudchadkar, S.R. & Fackler, J. Large language models and the perils of their hallucinations. Crit Care 27, 120 (2023). https://doi.org/10.1186/s13054-023- 04393-x
12. Tang L, Sun Z, Betina Idnay, et al. Evaluating Large Language Models on Medical Evidence Summarization. NPJ digital medicine. 2023;6(1). doi:https://doi.org/10.1038/s41746-023- 00896-7
13. Zhou L, Blackley SV, Kowalski L, et al. Analysis of errors in dictated clinical documents assisted by speech recognition software and professional transcriptionists. JAMA Network Open. 2018;1(3):e180530. doi:https://doi.org/10.1001/jamanetworkopen.2018.0530
14. Coiera E, Liu S. Evidence synthesis, digital scribes, and translational challenges for artificial intelligence in healthcare. Cell Reports Medicine. 2022;3(12):100860. doi:https://doi.org/10.1016/j.xcrm.2022.100860
15. Coiera E. The price of artificial intelligence. Yearbook of Medical Informatics. 2019;28(01):014-015. doi:https://doi.org/10.1055/s-0039-1677892
16. Murdoch B. Privacy and artificial intelligence: Challenges for protecting health information in a new era. BMC Medical Ethics. 2021;22(1). doi:https://doi.org/10.1186/s12910-021- 00687-3
17. Stokes D. Navigating the Future of Healthcare: A robust AI decision framework (Part 2). Ellit Groups. Published October 18, 2023. Accessed February 4, 2024. https://ellitgroups.com/navigating-the-future-of-healthcare-a-robust-ai-decision framework part-2/
18. Armstrong K, Ravenell KL, McMurphy S, Putt M. Racial/ethnic differences in physician distrust in the United States. Am J Public Health. 2007;97(7):1283-1289. doi:10.2105/AJPH.2005.080762
19. Rodriguez MA, García R. First, do no harm: the US sexually transmitted disease experiments in Guatemala. Am J Public Health. 2013;103(12):2122-2126. doi:10.2105/AJPH.2013.301520
20. Tierney AA, Gayre G, Hoberman B, et al. Ambient Artificial Intelligence Scribes to Alleviate the Burden of Clinical Documentation. NEJM catalyst innovations in care delivery. 2024;5(3). doi:https://doi.org/10.1056/cat.23.0404
21. Liebermann E, Taber P, Vega AS, et al. Barriers to family history collection among Spanish speaking primary care patients: a BRIDGE qualitative study 2024 Mar 27;4:100275]. PEC Innov. 2022;1:100087. doi:10.1016/j.pecinn.2022.100087