

Letter to the Editor

AI Scribe Safety: Measuring What Happens After Signing

Vera Sorin¹, MD; Eyal Klang², MD

¹Department of Radiology, Mayo Clinic Hospital, Rochester, MN, United States

²Department of Radiology, Beth Israel Deaconess Medical Center, Boston, MA, United States

Corresponding Author:

Eyal Klang, MD
Department of Radiology
Beth Israel Deaconess Medical Center
330 Brookline Avenue
Boston, MA 02215
United States
Phone: 1-617-754-9500
Email: eyalklang@gmail.com

Related Article:

Comment on: <https://medinform.jmir.org/2026/1/e89337/>

Abstract

Coiera and Fraile-Navarro question whether AI scribes are being evaluated on metrics that truly impact care. While current evaluations focus on the quality of the initial draft, signed clinical notes are dynamic, as their content can be copied, summarized, coded, and re-ingested by downstream AI tools. We argue that safety must be measured downstream, focusing on how small errors in initial documentation can compound across the patient's electronic health record.

JMIR Med Inform 2026;14:e103162; doi: [10.2196/103162](https://doi.org/10.2196/103162)

Keywords: artificial intelligence; AI; digital scribes; medical scribes; electronic documentation; patient safety; error prevention

In a recent *JMIR Medical Informatics* editorial, Coiera and Fraile-Navarro ask whether artificial intelligence (AI) scribes are being measured on what matters [1]. We agree with their central point: scribe output shapes what downstream clinicians and systems see and act upon. We build on that point by focusing on what happens to documentation errors after a note is signed and integrated into the health record.

This is far from theoretical. Reddy et al [2] compared notes from 11 ambient AI scribe tools with human notes across 5 standardized primary care visits; they found that human notes scored higher in thoroughness, organization, and usefulness in clinical context. Similarly, a pragmatic pilot study also found omissions, hallucinations, and other errors in AI-generated clinician notes [3]. These flaws can persist beyond the initial draft and propagate after the note is signed.

A signed clinical note is not the end of a visit. Another clinician may read it, a later note may copy it forward, or a referral may summarize it. Orders, billing codes, quality measures, registries, and downstream AI tools may rely on it. Thus, small errors can compound.

For instance, if dyspnea is omitted, the next clinician may not know it was discussed. If a medication dose is wrong, the next medication list may repeat it. If the reason for a plan is missing, a later clinician may follow it without knowing why. In the electronic health record (EHR), an AI-generated omission may be copied, summarized, coded, or used as input for another AI system.

This risk is familiar: EHRs already spread errors through copy-paste, copy-forward, and reused text [4]. Ambient AI scribes add another entry point: one error can become source material for later care.

While clinician review is thus necessary, it cannot catch every flaw. Omissions are hard to detect because the notes do not show omissions. Future studies should follow errors across the lifecycle of the clinical note: from the visit to the AI draft, signed note, and subsequent notes, referrals, handoffs, orders, patient instructions, registries, and AI summaries. Was the information complete and correct when reused? Was it corrected? Did it affect care? These questions align with established EHR data-quality work [5].

Our goal is not to reject ambient AI scribes, as documentation burden is a critical challenge. However, as Coiera and Fraile-Navarro argue, safety should be measured across the clinical workflow. We maintain that studies should follow

the note after signing. The question is not only whether AI scribes produce acceptable first drafts, but whether their errors stay contained or become part of the patient's clinical story.

Acknowledgments

Generative AI was used for spelling, grammar, and language refinement. The authors reviewed and take full responsibility for the final text.

Funding

The authors declared no financial support was received for this work.

Data Availability

Data sharing is not applicable to this article as no data sets were generated or analyzed.

Authors' Contributions

Conceptualization: VS, EK.

Writing – original draft: VS, EK.

Writing – review & editing: VS, EK.

Conflicts of Interest

None declared.

References

1. Coiera E, Fraile-Navarro D. AI scribes: are we measuring what matters? JMIR Med Inform. Feb 6, 2026;14:e89337. [doi: [10.2196/89337](https://doi.org/10.2196/89337)] [Medline: [41650281](https://pubmed.ncbi.nlm.nih.gov/41650281/)]
2. Reddy A, Gunnink E, Wheat CL, et al. Rapid evaluation of artificial intelligence technology used for ambient dictation in primary care: comparing the quality of documentation of artificial intelligence-generated and human-produced clinical notes. Ann Intern Med. Jun 2026;179(6):765-772. [doi: [10.7326/ANNALS-25-02772](https://doi.org/10.7326/ANNALS-25-02772)] [Medline: [41996184](https://pubmed.ncbi.nlm.nih.gov/41996184/)]
3. Taylor SL, Jost M, MacDonald S, et al. Quality of clinical notes created by ambient listening generative ai: pragmatic prospective pilot study. JMIR Med Inform. Apr 17, 2026;14:e86474. [doi: [10.2196/86474](https://doi.org/10.2196/86474)] [Medline: [41996389](https://pubmed.ncbi.nlm.nih.gov/41996389/)]
4. Tsou AY, Lehmann CU, Michel J, Solomon R, Possanza L, Gandhi T. Safe practices for copy and paste in the ehr. systematic review, recommendations, and novel model for health IT collaboration. Appl Clin Inform. Jan 11, 2017;8(1):12-34. [doi: [10.4338/ACI-2016-09-R-0150](https://doi.org/10.4338/ACI-2016-09-R-0150)] [Medline: [28074211](https://pubmed.ncbi.nlm.nih.gov/28074211/)]
5. Penev YP, Buchanan TR, Ruppert MM, et al. Electronic health record data quality and performance assessments: scoping review. JMIR Med Inform. Nov 6, 2024;12:e58130. [doi: [10.2196/58130](https://doi.org/10.2196/58130)] [Medline: [39504136](https://pubmed.ncbi.nlm.nih.gov/39504136/)]

Abbreviations

AI: artificial intelligence

EHR: electronic health record

Edited by Amanda Iannaccio, Andrew Coristine; This is a non-peer-reviewed article; submitted 31.May.2026; final revised version received 16.Jul.2026; accepted 17.Jul.2026; published 10.Aug.2026

Please cite as:

Sorin V, Klang E

AI Scribe Safety: Measuring What Happens After Signing

JMIR Med Inform 2026;14:e103162

URL: <https://medinform.jmir.org/2026/1/e103162>

doi: [10.2196/103162](https://doi.org/10.2196/103162)

© Vera Sorin, Eyal Klang. Originally published in JMIR Medical Informatics (<https://medinform.jmir.org>), 10.Aug.2026. 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.