Reimagining Clinical Documentation With Artificial Intelligence

Steven Y. Lin, MD; Tait D. Shanafelt, MD; and Steven M. Asch, MD, MPH

From the Division of Primary Care and Population Health, Department of Medicine (S.Y.L., S.M.A.), and WellMD Center (T.D.S.), Stanford University School of Medicine, Stanford, CA; and VA Center for Innovation to Implementation, Palo Alto, CA (S.M.A.).

Clinical documentation and review may be the leading cause of lost physician productivity in the United States. Physicians spend 34% to 55% of their work day creating notes and reviewing medical records in the electronic health record (EHR),[1-3] which is time diverted from direct patient interactions. Although some of this time bolsters ongoing care, much of it instead serves billing documentation, defense against litigation risk, and regulatory compliance, explaining why EHR use is strongly associated with dissatisfaction and burnout.[4] Time spent unnecessarily on documentation and review carries enormous economic implications. Based on the number of practicing US physicians (861,000 in 2015)[5] and a mean annual salary of $294,000,[6] this time may cost as much as $90 billion to $140 billion in physician time per year. Eliminating even a fraction of this time may mean significant benefits to patients, physicians, health systems, and payers.

Frustration with EHR documentation has given rise to the medical scribe industry. Scribes tackle major physician concerns with the EHR, including time-consuming data entry, interference with face-to-face patient care, degradation in the quality of clinical documentation (eg, templated notes, misuse of cut-and-paste), and performance of clerical tasks below their level of training. Patients readily accept them, and a recent randomized trial shows that scribes improve charting efficiency and physician satisfaction.[7] The industry projects a demand for 100,000 scribes, or one scribe for every nine physicians in the United States by 2020.[8]

Although effective, human scribes may offer only a stopgap strategy. Some experts believe that by undermining market demand for change, scribes may hinder the technological evolution of EHRs toward a more physician-centric design.[8] As EHR developers aim for radical redesign, a process that may take decades to achieve, intermediate informatics solutions to ease documentation could unlock physician time and reduce burnout. Because writing and reviewing notes takes up 78% of all physician interactions with EHRs,[1] we believe that these solutions should be designed around the process of note creation.

Artificial intelligence (AI) promises to accomplish this task through automating documentation.[9] Speech recognition and natural language processing technology can support the creation of notes in real time by listening in on patient-physician conversations.[10] Artificial intelligence can collect, sort, and assemble clinical information from multiple sources (eg, previous notes, laboratory results, radiology reports, pharmacy records) faster than humans.[11] During the note creation process, AI-enhanced decision support software can analyze a note's content and provide real-time evidence-based recommendations to physicians (eg, differential diagnosis, suggested evaluation, treatment guidelines) using dynamic clinical data mining.[12] Risk scores such as atherosclerotic cardiovascular disease to guide statin treatment, or CHA_2DS_2-VASc to guide anticoagulation decisions, can be calculated as part of note creation to augment clinical decision making.[13] Last, AI can automate and optimize the coding and billing process based on risk-adjustment factors[14] and draft level-of-service recommendations for office or hospital visits. The Figure presents a prototype for an "autoscribe" AI assistant.

What are the design requirements that developers might embrace in building the first autoscribe AI? Any autoscribe will need to take a nonlinear, fragmented, multilayered patient history and generate an accurate, cohesive, logical narrative by placing the right data in the right place in the note. It must be able to handle multiple voice inputs (eg, the patient, family members, physician, other members of the care team) that may overlap, be able to distinguish who is who, and handle varying accents and

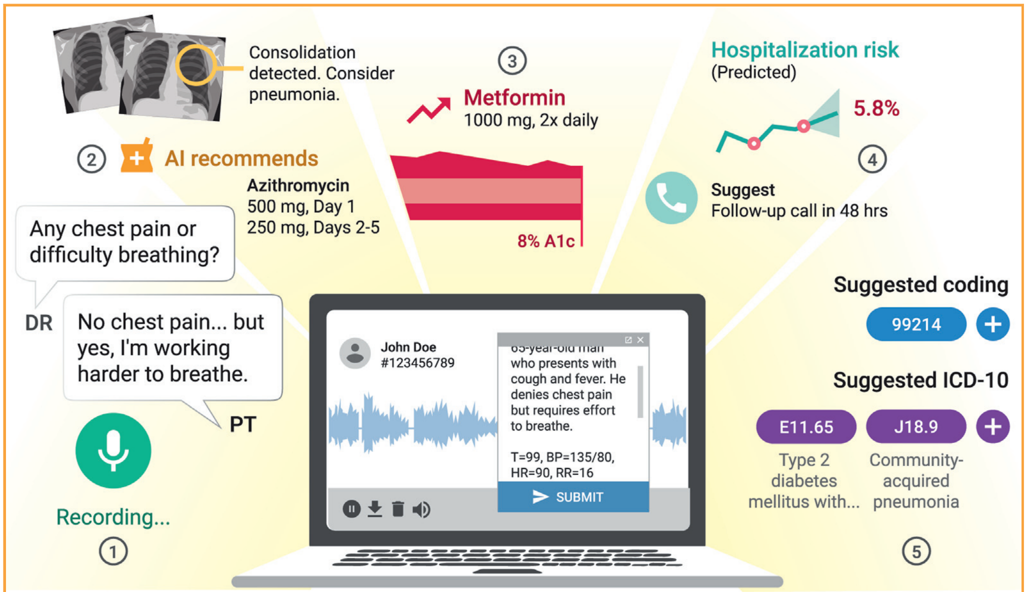

FIGURE. Theoretical prototype for an autoscribe artificial intelligence (AI) assistant in the electronic health record.

dialects. It must be able to automatically assemble clinical information from other sources—including previous notes and pertinent laboratory and radiology results from within a health system and externally—and summarize without causing data overload and rendering notes unreadable. It must seamlessly embed decision support (eg, differential diagnosis, recommended evaluation, evidence-based guidelines) into the primary user interface in a nondistracting way that minimizes pop-up menus, pointing, and clicking. Any autoscribe should be able to create notes that retain the "human-ness" and narrative interpretability of a physician writer or scribe. This is of particular importance in the era of OpenNotes,[15] when provider notes are becoming increasing accessible to patients via electronic portals. Last, all this technology should ideally be invisible to the patient and physician in order not to disrupt their conversations.

Most of the technologies needed to make autoscribe a reality already exist or are being developed. Now is the time for physicians, EHR developers, and AI technologists to engage with each other in the shared goal of transforming EHRs from passive data entry tools into intelligent digital assistants of the future. Physicians cannot afford to stay on the sidelines during the next EHR revolution, not when there is so much dissatisfaction and burnout created by the status quo. Working together with physicians, AI technologists and EHR developers will better understand the technical challenges posed by the nuances and complexities of real-world clinical encounters.

Although the promise of an autoscribe AI is appealing, the process for its development and testing is less clear. Creating even a prototype would require the collection of hundreds of thousands of hours of recorded patient-physician conversations for machine learning. Questions about how to collect, deidentify, store, and protect such data remain unanswered and are of particular concern in a time of high-profile hacks and public data breaches. Digital recordings of clinical encounters are already widespread, but policies regarding their use are lacking.[16] Technical challenges, such as how to incorporate nonverbalized elements (eg, the physical examination) into the note will need to be overcome. Questions around the financial costs of implementing such technology, the risk that AI might "un-train" physicians from important skills, and the medicolegal ramifications of errors made by AI will need to be addressed.[17] Similar to any new AI invention, autoscribe should be vigorously tested to reveal both its capabilities and limitations.[13] Ultimately, no AI is perfect, and what AI needs is human intelligence to command, nurture, and amplify it as a partner in patient care.[9]

This reimagining of clinical documentation, powered by AI, fundamentally alters the relationship between the physician and the EHR and moves the industry closer to the true spirit and vision of "meaningful use." Some fear that AIs will disrupt the physician-patient relationship by replacing human interaction; instead, autoscribe has the potential to unshackle physicians from EHRs and restore this sacred relationship. Artificial intelligence can transform EHRs into intelligent digital assistants that can exceed the demonstrated utility of human scribes, likely at a fraction of the cost. In an era in which AI-powered consumer devices, from watches to refrigerators, are becoming more ubiquitous, it is time for EHRs to join this revolution. Working together, physicians, EHR developers, and AI technologists can free physicians to do what they do best—taking care of patients in the service of healing, not documentation.

ACKNOWLEDGMENTS

We thank Katherine Chou, Elin Pedersen, Rebecca Rolfe, and Kasumi Widner for their input on technical concepts presented herein related to deep learning and artificial intelligence.

Potential Competing Interests: Dr Lin is the principal investigator on a research project sponsored by Google to understand how deep learning techniques and automatic speech recognition can facilitate clinical documentation. Google had no role in the preparation of this manuscript and the decision to approve publication of the finished manuscript.

Correspondence: Address to Steven Y. Lin, MD, 211 Quarry Rd, Suite 405, MC 5985, Palo Alto, CA 94304 (stevenlin@stanford.edu).

REFERENCES

1. Sinsky C, Colligan L, Li L, et al. Allocation of physician time in ambulatory practice: a time and motion study in 4 specialties. *Ann Intern Med*. 2016;165(11):753-760.
2. Tai-Seale M, Olson CW, Li J, et al. Electronic health record logs indicate that physicians split time evenly between seeing patients and desktop medicine. *Health Aff (Millwood)*. 2017;36(4):655-662.
3. Arndt BG, Beasley JW, Watkinson MD, et al. Tethered to the EHR: primary care physician workload assessment using EHR event log data and time-motion observations. *Ann Fam Med*. 2017;15(5):419-426.
4. Shanafelt TD, Dyrbye LN, Sinsky C, et al. Relationship between clerical burden and characteristics of the electronic environment with physician burnout and professional satisfaction. *Mayo Clin Proc*. 2016;91(7):836-848.
5. Association of American Medical Colleges. 2016 Physician Specialty Data Report. https://www.aamc.org/data/workforce/reports/457712/2016-specialty-databook.html. Updated December 31, 2015. Accessed August 29, 2017.
6. Grisham S. Medscape Physician Compensation Report 2017. http://www.medscape.com/slideshow/compensation-2017-overview-6008547. Updated April 5, 2017. Accessed August 29, 2017.
7. Gidwani R, Nguyen C, Kofoed A, et al. Impact of scribes on physician satisfaction, patient satisfaction, and charting efficiency: a randomized controlled trial. *Ann Fam Med*. 2017;15(5):427-433.
8. Gellert GA, Ramirez R, Webster SL. The rise of the medical scribe industry: implications for the advancement of electronic health records. *JAMA*. 2015;313(13):1315-1316.
9. Verghese A, Shah NH, Harrington RA. What this computer needs is a physician: humanism and artificial intelligence. *JAMA*. 2018;319(1):19-20.
10. Klann JG, Szolovits P. An intelligent listening framework for capturing encounter notes from a doctor-patient dialog. *BMC Med Inform Decis Mak*. 2009;9(suppl 1):S3.
11. Deliberato RO, Celi LA, Stone DJ. Clinical note creation, binning, and artificial intelligence. *JMIR Med Inform*. 2017;5(3):e24.
12. Celi LA, Marshall JD, Lai Y, Stone DJ. Disrupting electronic health records systems: the next generation. *JMIR Med Inform*. 2015;3(4):e34.
13. Chen JH, Asch SM. Machine learning and prediction in medicine: beyond the peak of inflated expectations. *N Engl J Med*. 2017;376(26):2507-2509.
14. Rose S. A machine learning framework for plan payment risk adjustment. *Health Serv Res*. 2016;51(6):2358-2374.
15. Delbanco T, Walker J, Darer JD, et al. Open notes: doctors and patients signing on. *Ann Intern Med*. 2010;153(2):121-125.
16. Elwyn G, Barr PJ, Castaldo M. Can patients make recordings of medical encounters? what does the law say? *JAMA*. 2017;318(6):513-514.
17. Cabitza F, Rasoini R, Gensini GF. Unintended consequences of machine learning in medicine. *JAMA*. 2017;318(6):517-518.