



Artificial Intelligence in Medicine & ChatGPT: De-Tether the Physician

Anthony M DiGiorgio¹ · Jesse M Ehrenfeld²

Published online: 4 March 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

A computer, ChatGPT, has now successfully passed the United States Medical Licensure Exam (USMLE) without any training [1]. However, put real world symptoms in its prompt and it gives a very canned response with extensive hedging and qualifiers, as if reading directly from WebMD.

Even the most advanced algorithms and AI-enabled tools still can't diagnose and treat diseases; this is the wrong approach. The probabilistic algorithms are too narrow. They simply can't substitute for judgement, nuance and thought. Crucially, the forthcoming FDA regulatory framework for AI enabled devices is proposing to be much more stringent on AI tools that make diagnoses and recommend treatments, especially if it is an algorithm that continues to adapt and learn over time [2]. At this moment, from a technological and regulatory standpoint, where AI can excel and should be leveraged is in unburdening physicians, de-tethering us from our computers and restoring the patient-physician relationship.

AI: Not Ready to Replace Doctors

The push for algorithm-based medicine has been a large part of "quality drive," on the heels of the *To Err is Human* report. More protocolized medicine should lead to fewer medical errors, the argument goes. Doctors largely pushed back against this idea. We realized that the algorithms substituted for thought, process measures substituting for

clinical judgement. Suddenly, we had to justify our decision making to a computer program (or an administrator armed with a computer program).

Algorithms are great for the textbook patient or solving a very narrow clinical question. That's why ChatGPT could pass the USMLE. It's full of such cases. But patients aren't standardized question stems. They are human beings with thoughts and emotions. They have very complex medical, social and psychiatric backgrounds. They rarely follow the textbook.

It's this complex person that requires a patient-physician relationship. This sacred bond emphasizes the individual. There is much a computer can't assess, such as nuance in patient body language, tone of voice and family interaction. Each patient has differing goals, and a "successful" outcome will not be the same for each patient. Even if a computer could analyze all that, the algorithm would still come to a generic answer. Zebras exist, but a probabilistic reasoning algorithm would never look for them.

These algorithms can also be flawed, based off faulty data or studies that are eventually overturned. The era of big data in medicine is still in its infancy, with most large data sets consisting of basic demographic information and ICD codes. The medical community is just now seeing the importance of collecting data on social determinants of health. Algorithms that are based off incomplete data can be harmful and even perpetuate systemic disparities [3].

The implementation of AI can't simply lead to electronic health record add-ins that generate multiple pop-up warnings as part of a clinical decision-making tool. It can't become part of a government mandated quality program that influences Medicare funding. It can't become a sword of Damocles which administrators hold over clinicians, demanding adherence, lest they be replaced by APPs who will obediently follow the algorithms.

✉ Anthony M DiGiorgio
Anthony.digiorgio@ucsf.edu

Jesse M Ehrenfeld
jehrenfeld@mcw.edu

¹ Department of Neurological Surgery & Institute for Health Policy Studies, University of California, San Francisco, 1001 Potrero Ave, Building 1, Room 101, Box 0899, San Francisco, CA 94143, USA

² Advancing a Healthier Wisconsin Endowment, Medical College of Wisconsin, 8701 W Watertown Plank Rd, Wauwatosa, WI 53226, USA

Utilizing AI to Unburden Physicians

AI and algorithms can assist doctors, but it must happen without impairing the patient-physician relationship. Imagine a helpful AI that integrates into the electronic health record to improve the clinician workflow. We see AI do this in our daily lives, from predictive text in our email to identifying people in our pictures.

The first requirement must be that AI is seamless and unobtrusive. It can't exist as pop-up warnings questioning a physician's decision making. It must provide helpful suggestions, reduce clicks and de-tether the physician from the computer. It must reduce clicks, not add to them.

AI could seamlessly integrate to provide predictive text for physician notes. As the assessment and plan are being populated in the note, the AI could start filling out orders, selecting ICD codes and capturing the appropriate CPT. The electronic health record could then present the orders and codes to the physician who then accepts or changes with a few clicks.

Gathering pertinent clinical history could also be performed by AI. Instead of clicking through endless reams of clinical documentation, a physician could have an AI generated relevant patient history that can be quickly verified with the patient. Instead of physicians being tethered to largely irrelevant inbox messages, AI could filter out the meaningful results.

Many inefficient behind-the-scenes billing processes could be greatly improved with AI as well. Instead of a cumbersome prior-authorization process, AI could be used by payors to improve utilization review. For example, algorithms could track surgeon cases and determine who had consistently solid, cost-effective indications. AI could be utilized to track true quality metrics without annoying coding queries.

ChatGPT is ready to tackle many administrative tasks now. It can generate CPT codes from operative reports with moderate accuracy. A little training and it will likely replace coders. It can also extract ICD-10 codes from clinical notes. This isn't an abstract futuristic concept. It can happen now.

Fostering Success

The AI marketplace must be competitive. Regulations must allow truly transformative, innovative, and useful AI to rise from the bottom-up in a competitive ecosystem. Mandating

specific algorithms from the top down is the wrong approach. This was tried with the current quality metrics and DRG systems, making them inequitable and reliant on specific documentation and coding. AI that is developed from the ground up would be much nimbler and more effective in supporting the care of patients.

A competitive marketplace however requires regulatory flexibility from the FDA. Regulation of AI systems is still in its infancy [2] but AI that improves physician workflow should require less regulatory oversight, than algorithms that make diagnoses, recommend treatments or otherwise impact clinical decision making. While AI algorithms may one day independently learn to read CT scans, identify skin lesions and provide medical diagnoses, the low-hanging fruit is in improving physician efficiency, de-tethering clinicians from the computer. This should be embraced by the healthcare industry now.

Funding AMD – Research funding from DePuy Synthes and The Mercatus Center at George Mason University.
JME – None.

Declarations

Disclosures/ Conflicts None to declare.

References

1. Kung, T.H., et al., Performance of ChatGPT on USMLE: Potential for AI-Assisted Medical Education Using Large Language Models. medRxiv, 2022: p. 2022.12.19.22283643.
2. US Food and Drug Administration. Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD). 2022 1/21/23]; Available from: <https://www.fda.gov/media/122535/download>.
3. Madhusoodanan, J., Is a racially-biased algorithm delaying health care for one million Black people? *Nature*, 2020. 588(7839): p. 546–547.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.