JOURNAL OF INSURANCE MEDICINE

J Insur Med 2023;50:000–000

OPINION

ChatGPT: How Closely Should We Be Watching?

Timothy Meagher, MB FRCP(C)

ChatGPT is about to make major inroads into clinical medicine. This article discusses the pros and cons of its use.

Address of Correspondent:
Munich Re, 1000 Rue de la Gauchetière Ouest, 20e étage, Montréal Québec, H3B 4W5; 514-392-5069; tmeagher@munichre.ca

Correspondent: Timothy Meagher, MB, FRCP(C); Vice-President and Medical Director, Munich Re, Montréal; Associate Professor of Medicine, McGill University, Montréal, Québec.

Key words: ChatGPT, artificial intelligence, technology, foundation models, clinical medicine, risk selection.

ChatGPT (Chat Generated Pre-trained Transformer) is widely considered a transformational advance in artificial intelligence. Comparisons to the arrival of the internet and the iPhone have been made. ChatGPT can perform an impressive list of tasks, from translation to creative writing, to generating code to solving mathematics problems, and all of these via a friendly conversational interface.

It also has medical skills: it can provide answers to many factual knowledge queries and has successfully passed steps 1-3 of the USMLE examination. Improved iterations have already arrived: GPT4 is even more powerful; its plug-ins allow interactions with third-party applications and provide a whole array of new utilities. And, by all accounts, this is just the beginning of the new world of "generative" artificial intelligence. Generative artificial intelligence describes algorithms that can generate new content, such as text or images, in response to a query or command.

LIFE ALTERING CHANGES

Could this technology alter the science and art of risk selection? The answer must be a loud "yes." This prediction is based on the following observations: ChatGPT– or whatever iteration or competitive model replaces it– is about to make major inroads into clinical medicine. Eventually, it will be omnipresent. In addition to reducing the administrative headaches that beset current practice, diseases will be detected earlier, and diagnoses will be more accurate. Treatments will be customized and will be more successful. Ultimately mortality will improve. As a result, insurance medicine will rethink

some of its time-honored risk-selection paradigms, while basking in the glow of mortality improvement.

WHY THIS BULLISH PREDICTION?

One could reasonably argue that, over the past 10 years, artificial intelligence has made modest inroads into clinical practice and has fallen short of its predicted success. True, it has made impressive gains in image analysis, where it occasionally surpasses human performance, and its value in the assessment of molecular data is indispensable. But by no measure has it become an indispensable feature of modern practice.

WHAT HAS CHANGED?

The main answer is the arrival of the so-called "foundation" models of artificial intelligence, of which ChatGPT is an example. Foundation models are deep-learning models that are trained on vast quantities of data and learn its general patterns and features. They then serve as a "foundation" for the development of subsequent models that can be adapted to a wide range of specific tasks, such as image recognition or translation. These models are capable of self-supervised learning, wherein the underlying algorithms can, with minimal training, process vast quantities of data and learn on the fly. This contrasts with the "supervised learning" models that power current medical applications. These latter models require heavily annotated training datasets, a labor-intensive process that relies on expert input. One example is the labeling of thousands of abnormal CT scan images, that create the "ground truth" from which algorithms can subsequently learn.

The second answer is the ability of foundation models to simultaneously analyze data of many different types, such as text, images, genomic sequences, and voice. This multimodal capacity in many ways mirrors the clinical exercise, where the integration of multiple pieces of information underpins diagnostic reasoning and treatment choices. By comparison, supervised learning models analyze data of a single type, eg, images or text, and not at the same time. Not surprisingly, such models, while undoubtedly successful, have had modest clinical utility and have enjoyed minimal impact. The 'unimodal' limitation is a substantial barrier to evolution and clinical applicability.

With this combination of self-supervised learning and multimodal inputs, it is inevitable that foundation models will improve the practice of medicine. More imperatively, it is unlikely that medicine can continue to progress without such models. The sheer volume of data that is presently generated in both the clinical care and biomedical research spheres is such that goal of universal personalized medicine will remain unattainable, without the assistance of artificial intelligence.

CHANGES FOR RISK SELECTION AND LIFE INSURANCE MEDICINE

Here are three ways that risk selection and life insurance medicine may be impacted: The first is improved mortality. The closer we can get to universal precision medicine the more mortality will be improved. Artificial intelligence will be a catalyst in this regard, accelerating earlier disease detection (eg, blood-borne cancer signals, retinal images to predict Alzheimer's disease, voice analysis to detect Parkinson's disease), more accurate diagnoses (eg, molecular analyses, high-precision imaging) and more effective treatments (individualized immunotherapies, gene editing, novel drug discovery, protein structure prediction).

While improved mortality is inevitable, the other 2 outcomes are a little less certain, although the optimist will quibble. Novel predictive associations will be discovered. Indeed, this phenomenon has already begun. In the early days of the COVID-19 pandemic, artificial intelligence models learned to predict, better than physicians, which patients

would fare poorly and require intensive care. In the next years, as more and more clinical and research data from ever-widening sources is fed into foundation models, it is inevitable that other predictive associations will emerge, whose impact will spread beyond a few days, and well into the future. Risk selection, whose core is the prediction of mortality, will benefit.

The final outcome, and arguably the most attractive from an insurer's perspective, is a redefinition of health. The majority of today's insurance applicants are considered healthy and are insured without additional premiums. Yet the label of "healthy" is a generously inclusive one. It more closely describes the absence of disease, than it does a state of perfect health. But then, what is perfect health or indeed health? Most would agree, it is an elusive construct. However, as clinical medicine progresses toward its personalized goal, it is likely that health will be defined in more granular terms, based on the predictive associations mentioned above, in addition to insights from novel data, such as gene profiles and wearable metrics. A deep phenotype of health will emerge, allowing health to be segmented in the same way that we segment illnesses. Again, insurance medicine will profit.

NOW TO THE NEGATIVE LEDGER

The list of challenges that artificial intelligence faces in the medical sphere is long and sobering. Many are unsolved. Generative artificial intelligence and foundation models, (computer scientists must be World Champions of Uninspired Names) add new flaws. Some experts are alarmed about their potential for harm.

ChatGPT has shown that errors are common, and some are howlers. Further, the new models have not been trained on the vast databases of medical data that will be key to their accuracy and success, and these datasets are dispersed and not easily accessed. However, the most preoccupying challenge may well be the 'black box' dilemma, wherein the inner workings of algorithms are not fully understood. For a profession that demands high-quality, evidential proof for most innovations, models that self-supervise their learning, but tell no one how they do it, generate mistrust. A mindset shift will be needed for the new models to be accepted by both regulators and clinicians.

One can reasonably ask: can all these problems be solved? There is no categorical answer. However, given the striking advantages that artificial intelligence is poised to deliver and the extraordinary sums of money that are being invested, it seems likely that most challenges will be solved.

Finally, as artificial intelligence transforms clinical medicine and the biomedical enterprise, it will also empower the consumer. The mobile phone will store a wealth of personal health data; on request, a customized model will provide interpretations and insights. The insurance applicant of the future will possess more accurate information about their health than ever before. While this should encourage better lifestyle and healthcare choices, it will also provide a better appreciation of future risk and thus the benefits of life and disability insurance. Greater information asymmetry at risk selection is a distinct possibility.

So, we should be watching closely. It may be a protracted watch, but it will be fascinating. More importantly, not to do so would be at our peril.

REFERENCES

1. Acosta JN, Falcone GJ, Rajpurkar P, et al. Multimodal biomedical AI. *Nat Med*. 2022;28:1773-1784. https://doi.org/10.1038/s41591-022-01981-2
2. Rajpurkar P, Chen E, Banerjee O, et al. AI in health and medicine. *Nat Med*. 2022;28:31-38. https://doi.org/10.1038/s41591-021-01614-0
3. Haupt CE, Marks M. AI-Generated Medical Advice-GPT and Beyond. *JAMA*. 2023;329:1349-1350. doi: 10.1001/jama.2023.5321. PMID: 36972070.

4. Moor M, Banerjee O, Abad ZSH, et al. Foundation models for generalist medical artificial intelligence. *Nature.* 2023;616:259-265. https://doi.org/10.1038/s41586-023-05881-4
5. Krishnan R, Rajpurkar P, Topol EJ. Self-supervised learning in medicine and healthcare. *Nat Biomed Eng.* 2022;6:1346-1352. https://doi.org/10.1038/s41551-022-00914-1
6. Will ChatGPT transform healthcare? *Nat Med*. 2023;29:505-506. https://doi.org/10.1038/s41591-023-02289-5
7. Li R, Kumar A, Chen JH. How Chatbots and Large Language Model Artificial Intelligence Systems Will Reshape Modern Medicine: Fountain of Creativity or Pandora's Box? *JAMA Intern Med.* Published online April 28, 2023. doi:10.1001/jamainternmed.2023.1835
8. Haug C, Drazen J. Artificial Intelligence and Machine Learning in Clinical Medicine. *N Engl J Med*. 2023; 388:1201-1208. doi:10.1056/NEJMra2302038
9. Lee P, Bubeck S, Petro J. Benefits, Limits, and Risks of GPT-4 as an AI Chatbot for Medicine. *N Engl J Med*. 2023;388:1233-1239. doi:10.1056/NEJMsr2214184
10. Metz C. The Godfather of A.I. Leaves Google and Warns of Danger Ahead. *New York Times*. May 1, 2023.