



# Rise of the Machines: AI and ChatGPT in Medical Writing

Since the advent of spell checkers on word processing platforms, humans have been using computers to help craft our works of writing. I am writing this introduction using a voice dictation program to transfer my words into text. As the text appears, Microsoft Word analyzes my spelling and grammar, while Grammarly runs a plagiarism checker. No matter what I write, it almost always finds a 1% incidence of plagiarism, as no combination of words is ever completely novel.

We will briefly review the history of the development of computer-assisted writing and address issues of citation, medical writing, and ethical considerations. ChatGPT was used in crafting this letter.

## History

In the 1980s, word processing programs such as WordPerfect and Microsoft Word were developed to provide basic document editing and formatting. Word introduced spell and grammar checking in the 1990s. Predictive text came a decade later, although many did not find it helpful.

In the last 10 years, deep learning has advanced sufficiently to assist document preparation. ChatGPT, now in its fourth generation, is a language generation model developed by OpenAI. It is based on the GPT (Generative Pre-training Transformer) architecture, which uses deep learning to generate human-like text. ChatGPT is trained on an enormous dataset of text. Large language models generate responses to user input appropriate to the question and the context. The output reads as natural conversation with proper grammar, spelling, and punctuation.

Large language models such as ChatGPT are based on neural networks, unusually flexible structures loosely based on the interconnectedness of biological neurons. Neural networks are adept at identifying hidden relationships. For years, websites and newspapers have used algorithms to generate reports, such as stock trends or sports predictions. These algorithms are based on machine learning, which is an extension of conventional regression techniques. What makes neural networks different is that they can be very broadly adapted. In the case of ChatGPT, it is hard to find a subject in

which the program can't provide a useful summary.

## ChatGPT as a coauthor

AI-assisted writing is likely to become as accepted tomorrow as Word is today. ChatGPT is already being used to generate articles that are submitted for publication. The obvious benefit is that the generated text will be polished and grammatically correct. However, it may also be inappropriately grandiose, utterly bland, or even meaningless.

Medical and scientific writing is not only about the final product but also the journey of creating something new. Every research report, meta-analysis, and clinical guideline is a narrative of the personal experience and perspective of the authors. AI-generated text cannot capture the emotional investment of the writer in the study design, the conduct of the research, the data analysis, and the final creation of a manuscript. The individuality and creativity associated with human-written text may be replaced by homogenization of content, where every manuscript starts to sound the same.

Text generated by an AI model must be reviewed and validated by human authors, just as authors are expected to validate the words of their coauthors. As described below, AI-generated text introduces an additional risk: the possibility of introducing bias, misinformation, and plagiarism. Authors should diligently screen for these whenever AI is used in manuscript creation.

Language generation models should be cited when used to generate the text, just as we have done here. Researchers need to be transparent about each contribution to the manuscript. Just as researchers would cite the tools used for statistical analysis or the source of laboratory reagents, the source of words must be cited in the methods of the paper.

Large language models raise the possibility of trivializing writing to the point of meaninglessness. Consider a paper written by a chatbot and submitted for publication (currently happening). The journal editor assigns chatbots as the reviewers (currently happening). The author chatbot revises the submission as requested by the reviewer chatbot. The revision is accepted and published, where the text is used to train the next large language models. Is

anything of value being written? What is the role of humans when computers increasingly talk to each other without us?

## Ethical considerations

An AI system has no ethical values or standards of its own. As a result, AI models will reflect the ethical standards and biases of the training text. For example, if a model is trained on text disproportionately written by men, the generated text will offer a male-centric perspective. A model trained on a dataset with misinformation will produce text with the same misinformation. AI chatbots can also be used for malicious purposes, such as generating fake news, impersonating others, or phishing.

Privacy is an important concern. Users can describe their medical conditions and receive highly personalized responses with patient-specific medical information. As of this writing, ChatGPT appears always to include “check this with your doctor” with every medical response. Users who enter sensitive personal information need to understand the potential of large language models to incorporate the new information into their training, possibly even including personally identifying medical information in responses to questions from other users.

Lack of transparency is also a concern. Neural networks (much like our synapses) act as millions of tiny nodes running in parallel. It can be very difficult to understand how a particular output arose from this network. This obfuscates the basis for a conclusion and may make it challenging to correct factual errors in the model. It also needs to be clarified how large language models balance peer-reviewed medical literature with websites that offer unsubstantiated personal opinions.

Since large language models are built on text written by others, plagiarism is fundamental in these models. Large language models are sufficiently facile at rearranging words, text, and ideas, so overt plagiarism is unlikely to appear. However, there is the potential for generated text to reflect the verbatim training text. When an author uses the words of another author without attribution, that is considered theft, with the approbation society associates with theft. Because AI models do not have ethical values, it is

not clear that the same approbation is appropriate when a language model uses the verbatim words of a human author. Unlike a human author, the language model truly doesn't “know” that it did anything wrong. It isn't clear that there is value in teaching neural networks to scrupulously avoid using verbatim text from a human author. By way of example, if you ask Google Assistant, Alexa, or Siri a question, the answer will be verbatim text by a human author found on the Internet. We have come to expect this from computers.

Authorship and ownership are also issues in the use of large language models. Who is the author when much of the text is generated by ChatGPT? How should authorship be divided among the person asking the question, the person who trained the model, and the model itself? These questions are being debated in the literary community.

Job displacement is another concern. Automated text generation can be used to automate tasks that were previously done by human writers, editors, and others. This could lead to job displacement and economic disruption. The use of AI-generated text can also lead to a dependence on the technology, which can be harmful if the models fail.

Large language models like ChatGPT have the potential to revolutionize medical writing and other language processing tasks. There are practical and ethical concerns with their use, including bias, misinformation, privacy, lack of transparency, job displacement, and dependence. Potential strategies to address these concerns are automated bias and misinformation detection, privacy auditing, transparently linking source material to the generated text, and workforce training so that skills replaced by AI are supplanted by new skills that leverage AI. Additionally, the output generated by these models should be reviewed and validated by experts before being used in any clinical or medical context. By considering these ethical concerns and taking appropriate measures, we can ensure that the benefits of these powerful tools are maximized while minimizing any potential harm.

The history of computer-assisted writing shows that advances in digital technology are invariably adopted by authors.

It is clear that large language models will rapidly become an essential writing tool. It is not clear how this will change how we write, how we teach, and how we learn. ■

Alexander Doyal, MD, MPH, FASA  
Monika Nanda, MBBS, MPH  
David Sender, MD  
Chapel Hill, North Carolina

\* \* \*

*The views and opinions expressed in the "Letters to the Editor" are those of the authors and do not necessarily reflect the views of ASA or the ASA Monitor Editorial Board. Letters submitted for consideration should not exceed 300 words in length. The Editor-in-Chief has the authority to accept or reject any letter submitted for publication. Personal correspondence to the Editor by letter or e-mail must be clearly indicated as "Not for Publication"*

*by the sender. Letters must be signed (although name may be withheld on request) and are subject to editing and abridgement. Send letters to [ASAMonitor@asahq.org](mailto:ASAMonitor@asahq.org).*

### **ASA Monitor ChatGPT Policy**

Authors who use AI tools in the writing of an article must be transparent in disclosing how the AI tool was used and which tool was used. Authors are fully responsible for the content of their article, even those parts produced by an AI tool, and are thus liable for any breach of publication ethics.