

**Article:**

Nader Rifai.

*Large Language Models for Scientific Publishing: Please, Do Not Make Them a Foe*  
Clin Chem 2024; 70(3): 468–70. <https://doi.org/10.1093/clinchem/hvad219>**Guest:** Dr. Nader Rifai from Harvard Medical School and Boston Children’s Hospital.

Bob Barrett:

This is a podcast from *Clinical Chemistry*, a production of the Association for Diagnostics & Laboratory Medicine. I’m Bob Barrett. In November 2022, the release of ChatGPT fundamentally changed the writing world. Text generated by artificial intelligence was now available to anyone with a computer and internet access, leaving journal editors and authors scrambling to determine how to incorporate this new tool in a way that abides by established standards of ethics and scientific integrity. Initial enthusiasm quickly gave way to serious questions with profound implications. Is AI-generated text equivalent to content produced by a human? What constitutes authorship? Should author still receive academic credit if they relied heavily on ChatGPT to draft certain sections or even the entirety of a scientific manuscript? What if AI-generated content looks convincing but is factually incorrect? On the other hand, proponents argue that AI models are powerful tools that would be foolish to exclude from the writing and publishing process.

A perspective article appearing in the March 2024 issue of *Clinical Chemistry* addresses both sides of this debate and shares the expected impact of large language models on editors, reviewers, and authors of scientific manuscripts. In this podcast, we are pleased to speak with the author of that perspective article. Dr. Nader Rifai is the Orah S. Platt Chair in Laboratory Medicine at Harvard Medical School and the Director of Clinical Chemistry at Boston Children’s Hospital. Dr. Rifai, ChatGPT initially generated a great deal of enthusiasm as it offered AI-generated writing for the masses, but soon after, signs of concerns started to emerge. Can you help give us a broader view of this landscape?

Nader Rifai:

Of course, Bob. ChatGPT was initially introduced in November 2022, and as you said, Bob, with the promise of making artificial intelligence generated writing accessible for everyone. In fact, the quality of the writing was better than expected and it became clear that it’s going to be difficult to distinguish between human writing and that of a machine. The dilemma first appeared in schools and universities among students who were using ChatGPT for their homework assignments. I recall once reading an article about it in *New York Times*, where a college professor confessed that the only

way he uncovered that the assignment was done by a machine is because the writing was almost perfect. So, after more similar stories appearing in both the lay and scientific press, it became clear that because of the potency of this technology and potential widespread utility, we are going to be forced to deal with issues that we have not experienced in the past.

One thing is certain, large language models, or LLM, learn from their own mistakes and get better with time, and newer and more powerful ones are being introduced. For example, GPT-4, introduced in March 2023, is expected to have significantly greater capability than ChatGPT. So many of the identified deficiency of LLM today will be remedied and resolved with time, thus making machine writing far more challenging to detect. An interesting problem that will be with us for the foreseeable future.

Bob Barrett: Well, it's easy to see the advantage in the appeal of the technology but can you elaborate on the pitfalls, at least as they pertain to scientific publishing?

Nader Rifai: Bob, in the scientific publishing world, there are four entities that are responsible for the creation of a paper. You have the publisher, the editor, the reviewer, and the author. LLM presents a unique set of challenges to each of the four parties. Let's start with publishers. Publishers' main fear from LLM is the proliferation of so-called paper mill companies, which are specialized in creating and selling manuscripts or authorships to boost profits. Understandably so, publishers do not want to find themselves in a position of having to investigate the authors and institutions to see if the work has actually been done. What publishers really want is for regulators to create means to monitor the use of AI in scientific publishing, such as the inclusion of watermark for LLM.

As for the editors, their biggest challenge would be to determine the required information needed to be provided by the author to properly evaluate the use of LLM in submitted work, and to find knowledgeable reviewers who can evaluate papers that are employing this technology. For reviewers, the currently available AI system are not really sophisticated enough to provide a thorough critique of the work. They just provide a summary of the article. So, they are not terribly helpful in that regard at the present time. Confidentiality, however, is a major concern about the use of LLM in manuscript review, and several major publishers have prohibited the uploading of their manuscripts to generative AI platforms.

As for investigators and author, the issue of using LLM in manuscript preparation is a bit more nuanced. I admit that I hold a bit more liberal views about it than some of my purest

colleagues, who disapprove of the idea of using machine written text in a scientific paper. I believe that the writing of the introduction and discussion sections, which normally explain the rationale of the study and the significance of the finding can be done by LLM with adequate keyword input from the author. Of course, there's always the risk of generating fake facts, known as hallucinations. But the author can detect and correct that. So, I think LLM tools are fine in the preparation of the first draft of a manuscript.

The other issue I wish to bring up is the fact that English has been the primary language in scientific journals, and those investigators whose native language is not English are at a clear disadvantage. So, the use of LLM in manuscript preparation may help in leveling the playing field. So, as you see, Bob, the issue is complicated and there are valid reasons to be concerned, but we have to take advantage of this technology, this tool, harness its power to advance our interests. Of course, there are other important questions that also have to be addressed.

Bob Barrett: Okay, understood. Well, then Dr. Rifai, what are these important questions that must be addressed for this technology to be more widely accepted?

Nader Rifai: Bob, there are many important legal, logistical, and philosophical questions that remain to be answered. The technology is here to stay. Many surveys in the literatures have already documented its use in the scientific community for various applications including study design, advanced data analysis, and manuscript preparation. Here in this podcast, we are focusing on its use in scientific publishing. So, the type of questions of interest are, can a copyright be granted for an LLM generated text or graphics, particularly if they were based in someone else's style? What is the acceptable percentage of LLM generated materials in a particular manuscript? How do we deal with the contribution of LLM to study design? Are we comfortable with challenging the established norms for truth? What constitutes an author? Et cetera, et cetera, et cetera.

These are not simple questions obviously, and one would think that an entity such as the National Academies of Sciences, Engineering, and Medicine or NIH will step up and take on this challenge. Otherwise, I'm afraid we are going to end up with contradictory policies from different publishers, journals and scientific societies that will not serve the interests and protect the integrity of the scientific publishing enterprise.

Bob Barrett: So, I'm curious now. Do you currently use artificial intelligence tools in any of your projects?

Nader Rifai: Yes. In fact, I do, Bob. As you know, in the past almost 10 years, I have been working on building an educational program called The Learning Lab for Laboratory Medicine, that is an AI-driven platform and based on the concept of adaptive learning. The closest to personalized education.

Basically, through sophisticated algorithms, the platform interacts with the learner and assesses their level of competency in a particular subject, then provides them with only the needed information to remedy the deficiency. As you know, we currently have over 120 advanced courses and 90 others that are specifically designed for practicing medical laboratory specialist, or MLS. Last year, we started a very ambitious project that entails the translation of the entire program to nine different languages. The way we are doing that is by using AI to do the initial translation to the language of interest. Then, work with the native expert to verify the translation. Of course, it's still a lot of work, but we have saved enormous amount of time by using AI assisted translations.

Bob Barrett: So, how far are you planning to take this technology in The Learning Lab, for example?

Nader Rifai: That's a very good question. Let me give you an example on how I envision using LLM in The Learning Lab. At the present time, an expert in a particular area will prepare a course from scratch. Building a course is not a trivial matter. It takes authors about 300 hours to complete the task. And if you add the editors, reviewers, beta testers, and the program administrators' time, you will find that the time needed to build the course is actually over 400 hours. So, we are planning to experiment with the use of LLM to build a course by working first with an expert to develop an extensive list of keywords and crucial short statements. The LLM then develops the document from which the course materials will be derived. This document will then be reviewed and edited by the experts, and finally, learning objectives, questions in different formats, and learning resources will be developed by LLM using this document.

We will see if LLM will be able to generate for us a high-quality course. I personally believe that it will.

Bob Barrett: Finally, Dr. Rifai, what are the take-home messages from all of this?

Nader Rifai: Well, Bob, we should not forget that LLM is just a tool, like a word processor and a spell checker. However, like all other powerful disruptive tools, it has risk associated with its use that must be recognized. There is no doubt that it's going to transform scientific publishing and help to create articles that are more interactive for a human and more extractable for a

machine. And that is one of the main reasons for my optimism about it.

Bob Barrett:

That was Dr. Nader Rifai from Harvard Medical School and Boston Children's Hospital. He wrote a perspective article on the use of large language models in scientific publishing in the March 2024 issue of *Clinical Chemistry*, and he's been our guest in this podcast on that topic. I'm Bob Barrett. Thanks for listening.