HHS Public Access

Beware of references when using ChatGPT as a source of information to write scientific articles

Luis Sanchez-Ramos, MD, PhD[1], **Lifeng Lin, PhD**[2], **Roberto Romero, MD, DMedSci**[3,4,5]

[1]Division of Maternal-Fetal Medicine, Department of Obstetrics and Gynecology, University of Florida College of Medicine, Jacksonville, Florida, USA

[2]Department of Epidemiology and Biostatistics, The University of Arizona, Tucson, Arizona, USA

[3]Pregnancy Research Branch, Division of Obstetrics and Maternal-Fetal Medicine, Division of Intramural Research, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, U.S. Department of Health and Human Services, Bethesda, Maryland, and Detroit, Michigan, USA

[4]Department of Obstetrics and Gynecology, University of Michigan, Ann Arbor, Michigan, USA

[5]Department of Epidemiology and Biostatistics, Michigan University, East Lansing, Michigan, USA

To the Editors:

The use of artificial intelligence (AI) chatbots in obstetrics and gynecology has been the subject of 2 recent articles in *American Journal of Obstetrics & Gynecology* (*AJOG*).[1,2] Chavez et al[1] addressed the potential value of ChatGPT, an artificial intelligence chatbot, in medical education and scientific writing. AI chatbots promise to facilitate the writing of scientific articles; however, we wish to inform readers of *AJOG* about a potential liability of ChatGPT. We have found that ChatGPT frequently provides erroneous and fictitious references that may include incorrect authors, journals, titles of articles, years of publication, and PubMed identifiers (PMID and DOI). Sometimes, the title of the article is correct, but the chatbots include incorrect authors or other information. Citations are intended to allocate appropriate credit for previous contributions, identify sources of ideas and chart the progress of a line of investigation. Errors in references undermine the credibility of the authors and trustworthiness of a scientific report.[3–5] Therefore, verification of all components of references provided by chatbots is necessary. The causes for the inaccuracies of ChatGPT are related to the vast amount of text data from diverse sources and inconsistency errors or inaccuracies in the primary data, which may influence the AI generated response. It is expected that the current limitations of AI chatbots will improve with time. A specific limitation of ChatGPT is that it relies on a fixed database with a particular knowledge cutoff date (November 2021). Recently, Bard, another AI chatbot, has been released by Google. This AI chatbot can be integrated with a Google Search and, therefore, may improve the accuracy of the information. Other AI assisted programs are available for identification

Corresponding Author: luis.sanchez@jax.ufl.edu (LS-R).

of relevant bibliography such as Elicit, "the AI research assistant," which can search for specific scientific literature (eg randomized clinical trials). In conclusion, authors need to verify the accuracy of chatbots outputs before submitting manuscripts for publication.

References

1. Chavez MR, Butler TS, Rekawek P, Heo H, Kinzler WL. Chat Generative Pre-trained Transformer: why we should embrace this technology. Am J Obstet Gynecol 2023. [Epub ahead of print].

2. Grünebaum A, Chervenak J, Pollet SL, Katz A, Chervenak FA. The exciting potential for ChatGPT in obstetrics and gynecology. Am J Obstet Gynecol 2023. [Epub ahead of print].

3. Thorp HH. ChatGPT is fun, but not an author. Science 2023;379:313. [PubMed: 36701446]

4. Santini A. The importance of referencing. J Crit Care Med (Targu Mures) 2018;4:3–4. [PubMed: 29967893]

5. Flier JS. Credit and priority in scientific discovery: a scientist's perspective. Perspect Biol Med 2019;62:189–215. [PubMed: 31281118]