Peer Review File

Reviewer A

You summarize the complexities of the original paper very well and make it quite digestible. I agree with you that I think these sorts of research articles deserve to be discussed by the academic community at large. I would ask you to consider this description of an editorial commentary from the journal's website: "The problems addressed by the relevant paper/report/event/topic should be considered within the wider context of the field." Your commentary would be strengthened by doing just that. I think the summarization of the original article is excellent and I appreciate the comparison of the paper's results to other similar papers, but I think this commentary is lacking your views on what "next steps" might be necessary for the field or what this paper's potential implications are. Some appropriate ideas might come from considering limitations of the study and future directions of next studies. For example:

Comment 1: The study only investigated ChatGPT. What about Claude or Bard/Gemini or other AI-powered chatbots?
- Reply 1: Thank you for the thoughtful and extensive review. We were able to look further into comparisons regarding these other chatbot tools you mentioned.
- Changes to the text: Please see lines 50-54 for discussion regarding other chatbots.

Comment 2: Prompt-engineering and its implications here. Might the manner in which the researchers wrote their questions impact the model's response? What implications are there here for future studies and the manner in which prompts are constructed?
- Reply 2: Thank you for providing this suggestion. We briefly discuss this in paragraph 5.
- Changes to the text: Please refer to new information in paragraph 5.

Comment 3: This study was performed in early 2023 on an older, less-powerful LLM. Might the results be even more accurate and appropriate with more powerful models that are used nearly a year later?
- Reply 3: Thank you for this thoughtful comment. Davis et al. utilized ChatGPT version 3.5. The newest version of Open AI's chatbot is GPT-4. We agree that some additional comments on the version of the LLM utilized and differences between this vs GPT-4 will be helpful. It is also important to note though that ChatGPT 3.5 is still the current version available for free, while GPT-4 is the paid version of this platform. We appreciate your comment prompting additions to our commentary which have strengthened the discussion.
- Changes to the text: Paragraph 4 is now a new addition discussing the comparison of

ChatGPT 3.5 vs 4.0.

Comment 4: These are just ideas, but I would encourage you to consider them and expand your discussion about how this study may have implications in the wider field.

- Reply 4: Thank you for each of your comments and suggestions. This allowed us to develop our discussion further.
- Changes to the text: Please refer to new paragraphs 4, 5, 8.

Comment 5: This opinion letter does not add any evidence to scientific literature background.
- Reply 5: We added in further paragraphs with studies that compared ChatGPT 3.5 and 4.0. We hope the addition of this new information will provide more insight into the scientific background literature on this pertinent topic.
- Changes to the text: Paragraph 4 was a new addition discussing the comparison of ChatGPT 3.5 vs 4.0. Paragraph 8 was a new addition discussing important factors to consider when creating a new evaluation tool.

Comment 6: Please indicate whether the ChatGPT referred in the manuscript is ChatGPT 3.5 or 4.0.
- Reply 6: Davis et al. utilized ChatGPT version 3.5.
- Changes to text: Please refer to line 25, this is where we mention that ChatGPT 3.5 was utilized.

Comment 7: Please illustrate the implication of the study you commented on and the future direction.
- Reply 7: Thank you for your comment. We now discuss future directions including the need for a validated tool to truly evaluate and compare the responses from these chatbots.
- Changes to text: We enhanced the discussion section in response to this comment and also added information about the future direction with a potential validated scoring tool specifically for chatbot responses by adding Paragraph 8.

The authors have presented a comprehensive and articulate summary of the research paper. I found the discussion particularly intriguing, especially the segment pertaining to the development of a novel tool for evaluating chatbot health information.

Comment 8: To further enhance the editorial, I would recommend a more detailed exploration of this innovative tool. Specifically, it would be beneficial to delineate how a proposed "AI DISCERN" tool, for example, will differ from the existing DISCERN framework. Such elaboration could provide a clearer understanding of its unique features and potential advantages.

Reply 8: Thank you for the thoughtful review. We added in a paragraph discussing important factors to consider when creating a new evaluation tool as well as specific ways it should be different than the existing DISCERN tool.

- Changes to text: We then enhanced the discussion surrounding the future of a validated scoring tool specifically made for chatbots by adding in paragraph 8.

Comment 9: Additionally, I noted a few minor grammatical inaccuracies within the text:

Line 48: DISERN should be DISCERN

Line 58: Consider changing "making a zero" to "making zero"

- Reply 9: Thank you for these keen edit suggestions for these grammatical inaccuracies within the text.

- Changes to text: The prior lines 48 and 58 were edited per directions above; now lines 82 and 111

Comment 10: A good, substantive comment, drawing attention to the advantages of the article the authors refer to, but also emphasizing the limitations of ChatGPT in Disseminating Urologic Information.

- Reply 1: Thank you for your comment.

- Changes in the text: No specific text changes made regarding this comment.

Comment 11: The authors provided an effective overview on the topic, posing valid arguments. No further comments on my sides.

- Reply 1: We appreciate your review and comment.

- Changes in the text: No specific text changes made regarding this comment.

Comment 12: Interesting commentary. Please also add the following reference, as it is one of the most recent commentaries on challenges of ChatGPT in medicine

(https://doi.org/10.1016/j.jacr.2023.10.019)
- Reply 1: Thank you for your comment and idea for additional reference addition to the manuscript.
- Changes in the text: Please refer to Paragraph 5, lines 68-71.