

**Received:** 2023-12-05    **Reviewed:** 2023-12-20    **Revised:** 2023-12-23    **Accepted:** 2023-12-25    **Published:** 2023-12-29**E d i t o r****R e v i e w e r s**

Ismail Dergaa<sup>id</sup>  
High Institute of Sport and Physical  
Education of Kef, Jendouba, Kef,  
Tunisia.  
Primary Health Care Corporation  
(PHCC), Doha, Qatar.  
Email: Phd.dergaa@gmail.com

**Reviewer 1:** Noomen Guelmami<sup>id</sup>  
High Institute of Sport and Physical Education of Kef, University of Jendouba, Kef,  
Tunisia  
Postgraduate School of Public Health, Department of Health Sciences (DISSAL),  
University of Genoa, Genoa, Italy  
Email: noomen.gualmemi@issepkef.u-jendouba.tn

**Reviewer 2:** Helmi Ben Saad<sup>id</sup>  
University of Sousse, Farhat HACHED Hospital, Sousse, Tunisia  
Email: helmi.bensaad@rns.tn

**1. Round 1***1.1 Reviewer 1*

Date: 19 November 2023

Dear Authors,

Your submission presents an intriguing analysis of ChatGPT and Google's Bard in the context of medicine. However, before I can recommend its acceptance, several revisions are necessary to enhance its structure and academic rigor.

1. Please provide a concise and informative title for your paper.

2. References Add citations to support statements on pages 3 and 9.

Adopt the Vancouver style for your references.

3. The current figures are inadequate. Please create larger, clearer figures, ideally dedicating a full page to each.

The conclusion section should be more concise and focused.

*1.2 Reviewer 2*

Date: 20 November 2023

I read with a great interest the paper titled "A Comparative Analysis of ChatGPT and Google's AI's "Bard" in Medicine". Bard is a conversational generative artificial intelligence chatbot developed by Google. It was developed as a direct response to the meteoric rise of OpenAI's ChatGPT, and was released in a limited capacity in March 2023 to lukewarm responses before expanding to other countries in May.

The rationale of the study is interesting. Comparing different language models, such as ChatGPT and Bard AI systems, in the context of medicine could be interesting for several reasons. First, comparing ChatGPT and Google's AI, Bard, in the field of medicine allows for an assessment of their performance in understanding and generating medical content. This includes evaluating their ability to provide accurate information, answer medical queries, and communicate effectively with users. Second, medical language is highly specialized, and evaluating how well these models handle medical terminology, stay updated on the latest medical research, and provide relevant information can be crucial. Comparisons can reveal which model demonstrates a better understanding of medical concepts. Third, both ChatGPT and Google's Bard may be used in the future to assist healthcare professionals in making clinical decisions. Comparisons can shed

light on their capabilities in providing reliable and evidence-based advice, helping healthcare practitioners in their decision-making process. Fourth, assessing the ability of these models to interact with patients in a natural and empathetic manner is important. Effective communication in healthcare settings, especially when dealing with patients, is crucial, and comparing these models can provide insights into their conversational abilities. Fifth, comparisons can extend beyond technical capabilities to ethical considerations. For example, how well do these models handle issues such as patient privacy, data security, and bias in medical information? Analyzing their ethical implications is crucial as AI systems become increasingly integrated into healthcare. Sixth, medicine is a dynamic field with continuous updates and new research findings. Comparing these models in their ability to adapt to and incorporate the latest medical information is important for assessing their relevance and reliability over time. Seventh, comparing the user experience when interacting with ChatGPT and Bard in a medical context can provide insights into which system is more user-friendly, intuitive, and adaptable to the needs of healthcare professionals and patients.

Before recommending the acceptance of this paper, a revision is needed.

- A. The authors are asked to determine the study design. According to me, it is a letter to editor, or an update or a short review. I recommend considering it as an editorial. They can ask help from the EIC
- B. Authors are asked to add a short title for their paper
- C. Authors are asked to delete the abstract –not needed for an editorial. If not, the authors are asked to define (at first use) Bard (if possible), AI and GPT...
- D. For key words, please avoid citing as keywords some terms used in the title or the abstract, opt for MeSH terms, and classify the key words in alphabetical order.
- E. Authors are asked to add inside the papers the links for Bard and ChatGPT 3.5.
- F. Authors need to clarify that they compared bard and ChatGPT 3.5 (not 4)
- G. Text: define all abbreviated (if possible) at first use: Bard, AI, GPT, ....page 3, add a reference after meaningful dialogues; professionals worldwide, medical documentation, medication errors, page 5: add references after patient outcomes, feasible for most, page 9: add references after subtleties in wording, peer-reviewed sources, page 9 define GDPR, HIPAA, and delete etc...page 10: define LLMs (1st use)
- H. Page 9: correct: this can lead Bard to produce....
- I. For the references, opt for the Vancouver style. Delete reference 3 as it is a preprint, ref 13: what is the journal name? Please homogenize your references.
- J. MAJOR points: figures are very bad. Please opt for big figures (one figure per page).
- K. Please shorten the conclusion
- L. I tried to ask bard the same first question (date December 17,2023) and I got different response than the one obtained by the authors (see below the response). Therefore, the authors are asked to discuss the reproducibility of the responses?

### 1.3 Authors

We would like to thank you for the very constructive and insightful feedback for optimizing this manuscript. We have revised our manuscript based on your helpful suggestions; these changes have significantly improved the manuscript and we thank for the time of the reviewers and editor.

Please find the point-by-point responses along with the details of the corresponding line changes to the manuscript. We have provided both a tracked and clean copy of the revised manuscript.

Thank you for the opportunity to revise the manuscript and we hope that it is now suitable for publication.

Detailed response:

Dear Reviewers and Editor-in-Chief,

We would like to thank you for the very constructive and insightful feedback for optimizing this manuscript. We have revised our manuscript based on your helpful suggestions; these changes have significantly improved the manuscript and we thank for the time of the reviewers and editor.

Please find the point-by-point responses along with the details of the corresponding line changes to the manuscript. We have provided both a tracked and clean copy of the revised manuscript.

Thank you for the opportunity to revise the manuscript and we hope that it is now suitable for publication.

Best wishes,

Ethan Waisberg, MB BCh BAO, Joshua Ong, MD, Mouayad Masalkhi, Nasif Zaman MSc, Prithul Sarker MSc, Alireza Tavakkoli PhD, Andrew G. Lee, MD

#### *1.4 Editor And Revisions*

I read with a great interest the paper titled “A Comparative Analysis of ChatGPT and Google’s AI’s “Bard” in Medicine”. Bard is a conversational generative artificial intelligence chatbot developed by Google. It was developed as a direct response to the meteoric rise of OpenAI’s ChatGPT, and was released in a limited capacity in March 2023 to lukewarm responses before expanding to other countries in May.

The rationale of the study is interesting. Comparing different language models, such as ChatGPT and Bard AI systems, in the context of medicine could be interesting for several reasons. First, comparing ChatGPT and Google’s AI, Bard, in the field of medicine allows for an assessment of their performance in understanding and generating medical content. This includes evaluating their ability to provide accurate information, answer medical queries, and communicate effectively with users. Second, medical language is highly specialized, and evaluating how well these models handle medical terminology, stay updated on the latest medical research, and provide relevant information can be crucial. Comparisons can reveal which model demonstrates a better understanding of medical concepts. Third, both ChatGPT and Google’s Bard may be used in the future to assist healthcare professionals in making clinical decisions. Comparisons can shed light on their capabilities in providing reliable and evidence-based advice, helping healthcare practitioners in their decision-making process. Fourth, assessing the ability of these models to interact with patients in a natural and empathetic manner is important. Effective communication in healthcare settings, especially when dealing with patients, is crucial, and comparing these models can provide insights into their conversational abilities. Fifth, comparisons can extend beyond technical capabilities to ethical considerations. For example, how well do these models handle issues such as patient privacy, data security, and bias in medical information? Analyzing their ethical implications is crucial as AI systems become increasingly integrated into healthcare. Sixth, medicine is a dynamic field with continuous updates and new research findings. Comparing these models in their ability to adapt to and incorporate the latest medical information is important for assessing their relevance and reliability over time. Seventh, comparing the user experience when interacting with ChatGPT and Bard in a medical context can provide insights into which system is more user-friendly, intuitive, and adaptable to the needs of healthcare professionals and patients.

Before recommending the acceptance of this paper, a revision is needed.

The authors are asked to determine the study design. According to me, it is a letter to editor, or an update or a short review. I recommend considering it as an editorial. They can ask help from the EIC

**We agree with the reviewer, this would better fit as an editorial.**

Authors are asked to add a short title for their paper

**A short title has now been added “ChatGPT versus Google’s AI’s “Bard” in Medicine”**

Authors are asked to delete the abstract –not needed for an editorial. If not, the authors are asked to define (at first use) Bard (if possible), AI and GPT...

**Abstract has now been deleted.**

For key words, please avoid citing as keywords some terms used in the title or the abstract, opt for MeSH terms, and classify the key words in alphabetical order.

**Keywords are not repeated from the title: “generative artificial intelligence; large language model; GPT-3.5; LLM; transformer network”. However, if there are any suggested MeSH terms we would be happy to include.**

Authors are asked to add inside the papers the links for Bard and ChatGPT 3.5.

**These links have now been added.**

Authors need to clarify that they compared bard and ChatGPT 3.5 (not 4)

**ChatGPT 3.5 was used, this has now been clarified.**

Text: define all abbreviated (if possible) at first use: Bard, AI, GPT, ....page 3, add a reference after meaningful dialogues; professionals worldwide, medical documentation, medication errors, page 5: add references after patient outcomes, feasible for most, page 9: add references after subtleties in wording, peer-reviewed sources, page 9 define GDPR, HIPAA, and delete etc...page 10: define LLMs (1<sup>st</sup> use)

**Line 1- Added abbreviation for AI.**

**There is no abbreviation for Bard, not included.**

Line 152- Added reference.

Page 9: correct: this can lead Bard to produce....

**Corrected.**

For the references, opt for the Vancouver style. Delete reference 3 as it is a preprint, ref 13: what is the journal name? Please homogenize your references.

**Reference 3 has now been deleted. Referencing style has been changed to Vancouver.**

MAJOR points: figures are very bad. Please opt for big figures (one figure per page).

**Agreed, images should appear larger in published version.**

Please shorten the conclusion

**Conclusion has now been shortened.**

I tried to ask bard the same first question (date December 17,2023) and I got different response than the one obtained by the authors (see below the response). Therefore, the authors are asked to discuss the reproducibility of the responses?

**Added lines 177-180 "Another final concern regarding Bard is reproducibility. When asked the exact same question, with the same wording, Bard can produce multiple different responses. This lack of consistency in response generation can lead to further difficulties when evaluating the responses generated by Bard."**

**2. Revised**

Editor's decision after revisions: Accepted.

Editor in Chief's decision: Accepted.