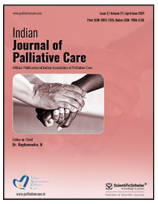
https://jpalliativecare.com/

Indian Journal of Palliative Care



Short Communication

Debunking Palliative Care Myths: Assessing the Performance of Artificial Intelligence Chatbots (ChatGPT vs. Google Gemini)

Prakash Gyandev Gondode[1], Vaishali Mahor[1], Deepa Rani[2], Rupavath Ramkumar[1], Pooja Yadav[1]

[1]Department of Anaesthesiology Pain medicine and Critical care, All India Institute of Medical Sciences, [2]Department of Anesthesiology, Vardhaman Mahavir Medical College, New Delhi, India.

ABSTRACT

Palliative care plays a crucial role in comprehensive healthcare, yet misconceptions among patients and caregivers hinder access to services. Artificial Intelligence (AI) chatbots offer potential solutions for debunking myths and providing accurate information. This study aims to evaluate the effectiveness of AI chatbots, ChatGPT and Google Gemini, in debunking palliative care myths. Thirty statements reflecting common palliative care misconceptions were compiled. ChatGPT and Google Gemini generated responses to each statement, which were evaluated by a palliative care expert for accuracy. Sensitivity, positive predictive value, accuracy, and precision were calculated to assess chatbot performance. ChatGPT accurately classified 28 out of 30 statements, achieving a true-positive rate of 93.3% and a true-negative rate of 3.3%. Google Gemini achieved perfect accuracy, correctly classifying all 30 statements. Statistical tests showed no significant difference between chatbots' classifications. Both ChatGPT and Google Gemini demonstrated high accuracy in debunking palliative care myths. These findings suggest that AI chatbots have the potential to effectively dispel misconceptions and improve patient education and awareness in palliative care.

Keywords: Palliative care, Misconceptions, Artificial intelligence, Chatbots, Patient education, Myth debunking

INTRODUCTION

Palliative care is an indispensable component of comprehensive healthcare, yet pervasive misconceptions persist among patients, caregivers and health-care providers, impeding access to essential services and contributing to undue suffering.[1] These misconceptions, ranging from misunderstandings of palliative care's scope to fears perpetuated by misinformation, underscore the critical need for effective education and awareness initiatives in this domain.[2]

In recent years, artificial intelligence (AI) has emerged as a transformative tool in healthcare, offering promising solutions to various challenges, including communication barriers and knowledge dissemination.[3] Language models, such as ChatGPT and Google Gemini, present an innovative approach to addressing misinformation and debunking myths surrounding palliative care, potentially revolutionising patient education and support mechanisms.[4]

This study aims to evaluate the effectiveness of AI chatbots, specifically ChatGPT and Google Gemini, in debunking prevalent myths associated with palliative care. By leveraging the capabilities of these AI-driven technologies, we seek to provide accurate information, dispel misconceptions, and ultimately improve access to quality palliative care services.

MATERIALS AND METHODS

Statement compilation

Thirty statements reflecting common misconceptions about palliative care were compiled from reputable sources, like the official website of the Indian Association of Palliative Care. (https://www.palliativecare.in/iapcs-infographics/) Infographics on 'General myths about palliative care' Part 1, 2 and 3. The compilation process ensured the inclusion of a diverse range of misconceptions, covering various aspects of palliative care domains such as basic introduction, Population, Setting, Children, End of Life, and Morphine. Ten myths statements from each of the three infographics, a total of 30 statements were compiled and then used for further evaluation.

***Corresponding author:** Prakash Gyandev Gondode, Department of Anaesthesiology Pain medicine and Critical care, All India Institute of Medical Sciences, New Delhi, India. drprakash777@gmail.com

Received: 22 February 2024 **Accepted:** 16 July 2024 **EPub Ahead of Print:** 09 August 2024 **Published:** 23 August 2024 **DOI:** 10.25259/IJPC_44_2024

AI chatbot response generation

Each of the 30 compiled statements was individually inputted into two AI chatbots: ChatGPT and Google Gemini. These AI models were chosen for their demonstrated capabilities in generating human-like text and addressing inquiries across different domains, including healthcare. The responses generated by ChatGPT and Google Gemini indicated whether each statement was classified as true or false according to the AI models' analysis. The chatbots' responses were recorded for further evaluation.

Expert evaluation

An expert in the field of palliative care, possessing extensive knowledge and experience in the domain, was engaged to evaluate the accuracy of the AI-generated responses and the explanations given in the infographics. The expert's assessment was based on current evidence, established best practices, and professional expertise in palliative care. Each of the 30 statements was reviewed individually by the expert, who determined whether the statements were true or false according to established palliative care principles.

Comparative analysis

The evaluations provided by the expert were compared with the responses generated by ChatGPT and Google Gemini for each of the 30 statements. This comparative analysis aimed to assess the performance of AI chatbots in accurately debunking palliative care myths. Discrepancies between the expert's evaluations and the chatbots' responses were noted and analysed to identify patterns and areas for improvement.

Statistical evaluation

To quantify the performance of the AI chatbots, various statistical metrics were calculated. These metrics included sensitivity, positive predictive value (PPV), accuracy, precision, and Chi-square tests. Sensitivity measured the proportion of true positives identified by the chatbots, while PPV measured the proportion of true positives among the statements classified as positive by the chatbots. Accuracy represented the overall correctness of the chatbots' classifications, while precision measured the proportion of true positives among the statements classified as positive by the chatbots. Chi-square tests were conducted to determine the significance of differences between the classifications made by ChatGPT and Google Gemini.

Ethical considerations

Ethical considerations were carefully addressed throughout the study to ensure the integrity and credibility of the research outcomes. While ethical clearance from the Ethical Committee was not pursued due to the retrospective design and absence of direct engagement with human subjects or sensitive personal data, all methodologies and procedures employed in the research adhered to established ethical principles and guidelines.

RESULTS

Performance of AI chatbots

In evaluating the effectiveness of AI chatbots in debunking palliative care myths, both ChatGPT and Google Gemini demonstrated notable performance.

ChatGPT accurately classified 28 out of 30 statements, achieving a true-positive rate of 93.3% and a true-negative rate of 3.3%. Specifically, it correctly identified 27 statements as true positives and 1 statement as a true negative [Figure 1]. On the other hand, Google Gemini achieved perfect accuracy by correctly classifying all 30 statements. It identified all 30 statements as either true positives or true negatives, resulting in a true positive rate and a true negative rate of 100%.

Comparative analysis

The Chi-square test was conducted to determine the significance of the differences between the classifications made by ChatGPT and Google Gemini. The calculated Chi-square value was approximately 0.00018. Comparison with the critical value of the Chi-square distribution at a significance level of $\alpha = 0.05$ indicated that the observed differences were not statistically significant. Therefore, it can be concluded that there is no significant difference between the classifications made by ChatGPT and Google Gemini for the given dataset.

Statistical metrics

Various statistical metrics were calculated to quantitatively evaluate the performance of the AI chatbots in debunking palliative care myths. ChatGPT demonstrated a sensitivity of 93.3%, indicating that it correctly identified 93.3% of the true positives. In addition, it achieved a PPV of 100%, signifying that all statements classified as positive by ChatGPT were indeed true positives. The overall accuracy of ChatGPT was 96.7%, indicating the proportion of correct classifications out of the total statements evaluated. Moreover, ChatGPT exhibited a precision of 100%, implying that all statements classified as positive by ChatGPT were true positives [Table 1]. Similarly, Google Gemini exhibited excellent performance across all statistical metrics. It demonstrated a sensitivity, PPV, accuracy, and precision of 100% each, indicating perfect performance in classifying palliative care myths.

DISCUSSION

Interpretation of results

The findings of this study demonstrate the high accuracy of AI chatbots, including ChatGPT and Google Gemini, in debunking palliative care myths. Both chatbots exhibited remarkable performance, with ChatGPT achieving a true positive rate of 93.3% and Google Gemini achieving perfect accuracy. These results underscore the potential

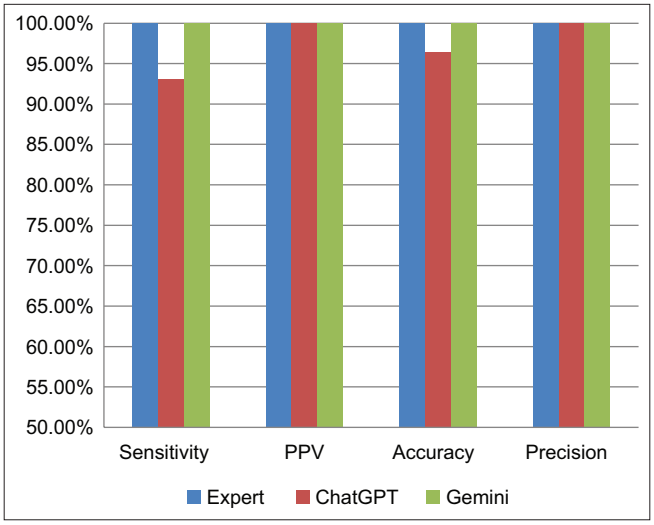


Figure 1: Comparison of metrics: Expert versus ChatGPT versus Gemini.

Table 1: Metrics derived post-evaluation for ChatGPT and Google Gemini.

Metric	ChatGPT results	Google Gemini results
TP	27	29
TN	1	1
FP	0	0
FN	2	0
Sensitivity	0.931	1
PPV	1	1
Accuracy	0.964	1
Precision	1	1

TP: True positive, TN: True negative, FP: False positive, FN: False negative, PPV: Positive predictive value

of AI chatbots as effective tools for patient education and communication in palliative care. By providing accurate information and dispelling misconceptions, AI chatbots can empower patients and caregivers to make informed decisions about their care and improve overall health outcomes.

Contextualisation and comparison

Our findings align with previous research highlighting the utility of AI-driven technologies in healthcare communication.[5,6] Comparisons with existing literature suggest that AI chatbots have the potential to significantly enhance patient education, dispel misconceptions, and improve access to quality palliative care services.[7] The perfect accuracy achieved by Google Gemini further validates the efficacy of AI chatbots in this context. However, it is important to note that while AI chatbots offer promising solutions, they should complement rather than replace human interaction in healthcare settings. Human empathy and expertise remain essential components of effective patient communication and care delivery.

Exploration of variations

While both chatbots demonstrated high accuracy overall, variations in performance were observed. These variations may be attributed to differences in algorithmic design, training data, or implementation. Further, exploration of these factors is warranted to understand better their impact on the reliability and generalizability of the findings. In addition, investigating user feedback and preferences could provide insights into the perceived effectiveness and usability of AI chatbots in addressing palliative care misconceptions. Understanding user perspectives is crucial for refining chatbot functionalities and optimising their performance in real-world healthcare settings.[8]

Future directions

Future research could explore ways to optimise AI chatbots for healthcare communication, such as fine-tuning algorithms and incorporating more diverse training data. User interaction studies conducted in real-time healthcare settings could provide valuable insights into the effectiveness of AI chatbots in addressing patient inquiries and concerns. In addition, integrating AI chatbots into clinical practice settings may offer opportunities to improve patient outcomes and healthcare delivery. Collaborative efforts between AI developers, healthcare providers, and patients can facilitate the co-design of chatbot interfaces and functionalities tailored to the specific needs and preferences of palliative care recipients.

Reflection on study limitations

It is important to acknowledge the limitations of this study, including the small sample size and lack of real-time interaction with users. These limitations may have influenced the findings and should be considered when interpreting the results. Future research should aim to address these limitations and further validate the findings in diverse healthcare settings. In addition, exploring potential biases introduced by the selection of statements or expert evaluation process could enhance the robustness of future studies in this area.

CONCLUSION

AI chatbots, specifically ChatGPT and Google Gemini, demonstrated high accuracy in debunking prevalent palliative care myths. Their ability to provide accurate and accessible information underscores their potential as powerful tools for patient education. By effectively dispelling misconceptions, these chatbots can significantly improve understanding of palliative care, leading to better informed decisions and improved patient outcomes. Further research and development are necessary to fully harness the potential of AI in palliative care patient education.

Ethical approval

The Institutional Review Board approval is not required.

Declaration of patient consent

Patient's consent was not required as there are no patients in this study.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

Use of artificial intelligence (AI)-assisted technology for manuscript preparation

The authors confirm that they have used artificial intelligence (AI)-assisted technology for assisting in the writing or editing of the manuscript or image creations.

REFERENCES

1. Hawley P. Barriers to Access to Palliative Care. Palliat Care 2017;10:1178224216688887.
2. Li WW, Chhabra J, Singh S. Palliative Care Education and Its Effectiveness: A Systematic Review. Public Health 2021;194:96-108.
3. Zahlan A, Ranjan RP, Hayes D. Artificial Intelligence Innovation in Healthcare: Literature Review, Exploratory Analysis, and Future Research. Technol Soc 2023;74:102321.
4. Bragazzi NL, Garbarino S. Assessing the accuracy of generative conversational artificial intelligence in debunking sleep health myths: mixed methods comparative study with expert analysis. JMIR Formative Research 2024; 16:e55762.
5. Shaw G. ChatGPT Will See You Now: The Promise and Peril of Chatbots in Patient Communication. Neurol Today 2023;23:14-5.
6. Gordon EB, Towbin AJ, Wingrove P, Shafique U, Haas B, Kitts AB, *et al*. Enhancing Patient Communication with Chat-GPT in Radiology: Evaluating the Efficacy and Readability of Answers to Common Imaging-related Questions. J Am Coll Radiol 2024;21:353-9.
7. Reddy V, Nafees A, Raman S. Recent Advances in Artificial Intelligence Applications for Supportive and Palliative Care in Cancer Patients. Curr Opin Support Palliat Care 2023;17:125-34.
8. Singh SK, Kumar S, Mehra PS. Chat GPT and Google Bard AI: A Review. In: 2023 International Conference on IoT, Communication and Automation Technology (ICICAT). United States: IEEE; 2023. p. 1-6.

How to cite this article: Gondode PG, Mahor V, Rani D, Ramkumar R, Yadav P. Debunking Palliative Care Myths: Assessing the Performance of Artificial Intelligence Chatbots (ChatGPT vs. Google Gemini). Indian J Palliat Care 2024;30:284-7. doi: 10.25259/IJPC_44_2024