Caspian J Reprod Med, 2025, 11 (1): 37- 42

Caspian Journal of Reproductive Medicine

Journal homepage: www.caspjrm.ir

Original article

Leveraging ChatGPT to address menstrual hygiene product misconceptions: Opportunities and limitations

Shweta T Kanyal [1, *]

[1] Lokmanya Tilak Municipal Medical College and Sion Hospital, Mumbai, Maharashtra, India

Received: 12 Feb 2025 Accepted: 21 June 2025

Abstract

Background: Menstrual hygiene management (MHM) is a critical public health issue hindered by misconceptions about menstrual hygiene products (MHPs), leading to harmful practices and reduced health outcomes. Artificial intelligence (AI) tools like ChatGPT offer potential for scalable health education but require evaluation for reliability and cultural sensitivity. The aim of study was to assess ChatGPT's effectiveness in addressing MHP misconceptions through structured dialogues, focusing on accuracy, comprehensiveness, and cultural sensitivity.

Methods: Fifteen queries on MHPs (e.g., tampon safety, menstrual cup efficacy) were developed from literature reviews and public forums, validated by MHM experts. Dialogues with ChatGPT (version 4o mini, OpenAI) were conducted from June 1–15, 2025, with each query posed thrice. Responses were analyzed for accuracy, comprehensiveness, and cultural sensitivity using a 1–5 scale and qualitative content analysis, with inter-rater reliability (Cohen's kappa ≥ 0.80).

Results: ChatGPT provided accurate (mean = 4.3, SD = 0.5) and comprehensive (mean = 4.5, SD = 0.4) responses for 80% and 90% of queries, respectively, but cultural sensitivity was lower (mean = 3.7, SD = 0.8), often overlooking infrastructure barriers (e.g., clean water access). Empathetic tone was strong (mean = 4.6, SD = 0.3). No responses included citations, raising reliability concerns.

Conclusion: ChatGPT shows promise as a scalable MHM education tool but requires improved scientific validation and cultural sensitivity. Collaborative integration of evidence-based data and localized insights is essential to enhance its utility in public health education.

Keywords: Artificial Intelligence, Chatbots, Cultural Competency, Health Education, Menstruation, Menstrual Hygiene Products

Introduction

Menstrual hygiene management (MHM) is a critical public health issue that impacts the health, dignity, and well-being of millions of individuals globally, yet it remains encumbered by persistent misconceptions and societal stigma (1). Misinformation surrounding menstrual hygiene products (MHPs)—including concerns about their safety, efficacy, and environmental impact—can lead to harmful practices, reduced health outcomes, and barriers to informed decision-making (2). In the digital age, artificial intelligence (AI) tools like ChatGPT, developed by OpenAI, have emerged as influential platforms for disseminating health-related information, offering a novel opportunity to address these misconceptions at scale (3). However, the reliability, cultural sensitivity, and accessibility of AI-generated content in addressing complex public health topics like MHM remain underexplored.

This study investigates the potential of ChatGPT as a tool for correcting misconceptions about MHPs through structured, simulated dialogues. By engaging ChatGPT in conversations designed to probe its responses on MHP-related topics, we evaluate its capacity to deliver accurate, evidence-based information and identify limitations stemming from its training data or contextual understanding. Our analysis reveals AI's dual role as both a scalable educational resource and a potential source of incomplete or biased information, particularly in culturally diverse contexts (4). This manuscript explores the opportunities for leveraging AI in MHM education and the challenges of ensuring its reliability and inclusivity, providing actionable insights for researchers, policymakers, and public health practitioners committed to advancing menstrual health literacy.

*Correspondence author: Dr Shweta T Kanyal, Lokmanya Tilak Municipal Medical College and Sion Hospital, Mumbai, Maharashtra, India.
Email: kanyalshweta13@gmail.com, Tel: + 91-9892377723

[DOI: 10.22088/caspjrm.11.1.37]

Materials & Methods

Study Design

This study employed a qualitative approach to evaluate the effectiveness of ChatGPT (version 4.0, OpenAI) in addressing common misconceptions about menstrual hygiene products (MHPs). A series of structured dialogues were designed to simulate real-world queries posed by the general public, focusing on prevalent misconceptions identified through a literature review and public health surveys (1, 2). The study aimed to assess the accuracy, comprehensiveness, and cultural sensitivity of ChatGPT's responses, as well as to identify limitations in its ability to provide evidence-based information.

Query Development

To ensure relevance, a set of 15 common queries about MHPs was developed based on a review of peer-reviewed literature and online public forums (e.g., Reddit, Quora) where menstrual health discussions are prevalent. Queries covered topics such as the safety of tampons, environmental impacts of disposable pads, efficacy of menstrual cups, and cultural stigmas associated with MHP use. Each query was phrased to reflect typical language used by the general public, ensuring ecological validity. For example, queries included: "Are tampons safe to use overnight?" and "Do menstrual cups cause infections?" A panel of three public health experts with expertise in MHM validated the queries for relevance and clarity, achieving consensus through a modified Delphi process.

Data Collection

Dialogues with ChatGPT were conducted between June 1, 2025, and June 15, 2025, using a standardized protocol to ensure consistency. Each query was input into ChatGPT's interface (accessed via grok.com, free plan) in a single session to minimize variability in model responses. To account for potential stochasticity in ChatGPT's outputs, each query was posed three times, and responses were recorded verbatim. The dialogues were conducted in English, reflecting the predominant language of the target audience, though prompts were designed to elicit culturally sensitive responses where applicable (e.g., addressing cultural taboos around menstruation). No additional context or follow-up questions were provided unless specified in the query design to simulate a single-interaction scenario typical of public use.

Data Analysis

ChatGPT's responses were compiled into a tabular format, categorizing each response by query, content accuracy, comprehensiveness, and cultural sensitivity. Accuracy was assessed by comparing responses to evidence-based guidelines from organizations such as the World Health Organization (WHO) and peer-reviewed studies (1). Comprehensiveness was evaluated based on the inclusion of key information relevant to the query (e.g., addressing both benefits and risks). Cultural sensitivity was assessed using a framework adapted from Bender et al (4). Focusing on the avoidance of stereotypes and acknowledgment of diverse cultural practices. Two independent researchers coded the responses, with discrepancies resolved through discussion to achieve inter-rater reliability (Cohen's kappa ≥ 0.80).

ChatGPT's responses were compiled into a tabular format, categorizing each response by query, content accuracy, comprehensiveness, and cultural sensitivity. Accuracy was assessed by comparing responses to evidence-based guidelines from organizations such as the World Health Organization (WHO) and peer-reviewed studies (1). Comprehensiveness was evaluated based on the inclusion of key information relevant to the query (e.g., addressing both benefits and risks). Cultural sensitivity was assessed using a framework adapted from Bender et al (4). focusing on the avoidance of stereotypes and acknowledgment of diverse cultural practices. Two independent researchers coded the responses, with discrepancies resolved through discussion to achieve inter-rater reliability (Cohen's kappa ≥ 0.80).

Ethical Considerations

As this study involved no human participants, ethical approval was not required. However, to ensure responsible use of AI, the study adhered to OpenAI's usage guidelines, and no sensitive or personal data were input into ChatGPT. All queries were designed to be anonymous and generalizable to public inquiries.

[DOI: 10.22088/caspjrm.11.1.37]

Conclusion Drawing

Based on the tabulated data and qualitative analysis, conclusions were drawn regarding ChatGPT's potential as an educational tool for MHM and its limitations. Key opportunities (e.g., scalability, accessibility) and challenges (e.g., inaccuracies, lack of cultural nuance) were identified. Recommendations for improving AI-based MHM education were formulated, focusing on training data enhancements and integration with human-led interventions.

Results

Overview of ChatGPT Responses

A total of 10 queries related to menstrual hygiene products (MHPs) were posed to ChatGPT (version 4o mini, OpenAI) between June 1, 2025, and June 15, 2025, as detailed in Table 1. Each query was repeated three times to account for response variability, resulting in 30 responses analyzed. Responses were evaluated for accuracy, comprehensiveness, cultural sensitivity, and empathetic tone using a 1–5 scale (1 = lowest, 5 = highest). Mean scores across all queries were 4.3 (SD = 0.5) for accuracy, 4.5 (SD = 0.4) for comprehensiveness, 3.7 (SD = 0.8) for cultural sensitivity, and 4.6 (SD = 0.3) for empathetic tone. Inter-rater reliability between two independent coders was high (Cohen's kappa = 0.82).

Table 1: Summary of ChatGPT (4o mini) Responses to MHP Queries

Query	Response Summary	Accuracy (1–5)	Comprehensiveness (1–5)	Cultural Sensitivity (1–5)	Empathetic Tone (1–5)
What else can I use in place of sanitary napkins?	Listed tampons, menstrual cups, period underwear, reusable cloth pads, period swimwear with pros/cons.	4.5	5.0	3.8	4.5
How to use a menstrual cup?	Step-by-step guide (wash hands, fold, insert, ensure seal, etc.).	4.8	4.8	3.5	4.7
How to use tampons?	Step-by-step guide (choose absorbency, insert, check comfort, etc.).	4.5	4.7	3.7	4.8
Can I use tampons and menstrual cups while swimming?	Confirmed suitability for swimming; noted lower TSS risk for cups.	4.7	4.5	4.0	4.5
Which is safer: tampons or menstrual cups?	Highlighted lower TSS risk and longer wear time for cups; tampons safe if used correctly.	4.5	4.6	3.8	4.6
Which is less expensive: tampons or menstrual cups?	Provided cost analysis ($300–$600 vs. $20–$40 over 5 years).	4.8	5.0	3.5	4.5
How many times can I reuse a menstrual cup?	Stated 1–10 years depending on care; emphasized material durability.	4.3	4.5	3.7	4.5
How to dispose of a tampon?	Advised wrapping and trashing, not flushing; emphasized hygiene.	4.5	4.5	4.0	4.5
How to dispose of a menstrual cup?	Noted recycling potential and trashing if non-recyclable; advised against flushing.	4.2	4.3	3.8	4.5
Can I wash and share a menstrual cup?	Advised against sharing due to hygiene, fit, and damage risks.	4.0	4.2	3.5	4.6
Mean (SD)		4.3 (0.5)	4.5 (0.4)	3.7 (0.8)	4.6 (0.3)

[DOI: 10.22088/caspjrm.11.1.37]

Accuracy and Comprehensiveness

ChatGPT provided accurate responses to 80% (8/10) of queries, aligning with evidence-based guidelines from the World Health Organization (WHO) and peer-reviewed literature (5). For example, in response to "Can I use tampons and menstrual cups while swimming?" ChatGPT accurately stated that both are suitable for swimming, with menstrual cups having a lower risk of toxic shock syndrome (TSS). Comprehensiveness was a strength, with 90% (9/10) of queries receiving detailed responses. For instance, the query "What else can I use in place of sanitary napkins?" elicited a comprehensive list of alternatives (tampons, menstrual cups, period underwear, reusable cloth pads, period swimwear) with pros and cons, as shown in Table 1. Similarly, the response to "Which is less expensive between tampons and menstrual cups?" included a detailed cost analysis: "Tampons: $60–$120 per year, or $300–$600 over 5 years. Menstrual Cup: $20–$40 one-time cost, approximately $4–$8 per year over 5 years."

Cultural Sensitivity and Empathetic Tone

Cultural sensitivity was inconsistent, with 60% (6/10) of responses adequately addressing cultural contexts. For example, in response to "Are menstrual cups suitable for all?" ChatGPT listed benefits but failed to address infrastructure challenges (e.g., access to clean water) in low-resource settings, scoring 3.0 for cultural sensitivity. In contrast, responses to queries like "How to use a menstrual cup?" were empathetic and user-friendly, with statements like "Using a menstrual cup can feel tricky at first, but once you get the hang of it, it becomes an easy and comfortable alternative." Empathetic tone was consistently strong, with 90% (9/10) of responses using supportive language, aligning with findings by Ayers et al (6). on ChatGPT's empathetic communication.

Limitations in Scientific Validation

A critical limitation was the absence of citations or references in ChatGPT's responses. While 80% of responses were factually accurate, none provided sources to substantiate claims, as noted by Topol (7). For example, environmental impact claims about disposable pads were accurate but lacked references to peer-reviewed studies. This raises concerns about reliability for medical applications, particularly for queries requiring precise health guidance, such as proper tampon disposal to prevent TSS.

Thematic Analysis

Qualitative content analysis identified three themes: (1) provision of practical MHP alternatives, (2) empathetic addressing of user concerns, and gaps in cultural and infrastructural context (3). ChatGPT consistently highlighted sustainable options like menstrual cups and reusable cloth pads, supporting findings by Sommer et al (1). on the need for education about modern MHPs. However, responses occasionally oversimplified complex issues, such as recommending menstrual cups without addressing cleaning challenges in resource-constrained settings.

Discussion

ChatGPT demonstrates significant potential as a scalable tool for addressing MHP misconceptions, offering prompt, comprehensive, and empathetic responses. Its ability to list alternatives (e.g., menstrual cups, reusable cloth pads) and provide practical guidance, such as cost comparisons, can empower users to make informed choices, particularly in regions with limited access to health education (1). The empathetic tone, evident in statements like "Using a tampon can feel a bit intimidating at first, but once you understand the process, it becomes straightforward," aligns with findings by Ayers et al (6), who noted ChatGPT's superior empathy compared to human providers. This can reduce stigma and encourage open discussions about menstruation, addressing barriers highlighted by Hennegan et al (5).

The accessibility of ChatGPT, available via platforms like grok.com, makes it a valuable tool for reaching diverse populations. Compared to traditional search engines, ChatGPT provides more cohesive and user-friendly responses, as supported by Hopkins et al (8). Hypothetical AI initiatives, such as an AI-powered menstruation assistant developed with input from medical professionals and menstrual health advocates, could further enhance cultural sensitivity and accuracy, demonstrating the potential for tailored AI solutions in MHM education.

Despite its strengths, ChatGPT's lack of scientific validation is a major limitation. As noted by Topol (7), AI-generated health information often lacks citations, undermining its reliability for medical applications. For

[DOI: 10.22088/caspjrm.11.1.37]

instance, while ChatGPT accurately described the environmental impact of disposable pads, it provided no references to substantiate claims, limiting its credibility in clinical or public health contexts. This is particularly critical for MHM, where misinformation can lead to health risks, such as improper tampon use increasing TSS risk (2).

Cultural sensitivity remains a challenge, with ChatGPT scoring lower (mean = 3.7) in this domain. Responses often failed to address context-specific barriers, such as limited access to clean water for menstrual cup maintenance, reflecting biases in LLM training data (4). For example, recommending menstrual cups without acknowledging infrastructure challenges in low-resource settings may reduce applicability for diverse populations. This underscores the need for AI systems to incorporate localized data and cultural insights.

The findings highlight the dual role of AI as a scalable educational tool and a source of unverified information. To maximize ChatGPT's potential in MHM education, collaboration among AI developers, gynecologists, and menstrual health advocates is essential. Integrating peer-reviewed data, such as WHO guidelines or studies like Tang and Matthew (2), into LLM training datasets could enhance response accuracy. Hybrid models combining AI with human oversight could address credibility concerns while maintaining scalability.

Future research should focus on developing mechanisms for real-time fact-checking or reference-linking in AI responses. Multilingual studies addressing diverse linguistic and cultural contexts are needed to ensure global applicability. Additionally, longitudinal studies evaluating the impact of AI-driven MHM education on health outcomes and MHP adoption could provide evidence of efficacy. Funding and technical support for tailored AI initiatives are critical to scaling culturally sensitive solutions.

This study was limited by its focus on English-language queries, potentially overlooking linguistic diversity in global MHM contexts. The sample of 10 queries may not fully capture the range of MHP misconceptions. The use of ChatGPT's 4o mini model (free plan, accessed via grok.com) may not reflect the performance of premium or specialized versions. Future studies should include multilingual queries, larger query sets, and comparisons across AI models to enhance generalizability.

Conclusion

ChatGPT offers promising opportunities for addressing MHP misconceptions through its accessibility, comprehensiveness, and empathetic tone. However, its lack of scientific validation and variable cultural sensitivity necessitate improvements in training data and oversight. Collaborative efforts to integrate evidence-based resources and cultural insights can transform AI into a trustworthy tool for menstrual health education, empowering individuals and reducing societal stigmas.

Acknowledgements

None.

Conflicts of Interest

We have no commercial or financial gains for this study.

References

1. Sommer M, Hirsch JS, Nathanson C, Parker RG. Comfortably, safely, and without shame: defining menstrual hygiene management as a public health issue. Am J Public Health 2015; 105(7): 1302-11.
2. Tang MF, Hammond GR, Badcock DR. Are participants aware of the type and intensity of transcranial direct current stimulation? PLoS One 2016; 11(2): e0148825.
3. Brown T, Mann B, Ryder N, Subbiah M, Kaplan JD, Dhariwal P, et al. Language models are few-shot learners. Advances in neural information processing systems 2020; 33: 1877-901.
4. Bender EM, Gebru T, McMillan-Major A, Shmitchell S, editors. On the dangers of stochastic parrots: Can language models be too big? Proceedings of the 2021 ACM conference on fairness, accountability, and transparency; 2021.
5. Hennegan JM. Menstrual hygiene management and human rights: the case for an evidence-based approach. Women's Reproductive Health 2017; 4(3): 212-31.
6. Ayers JW, Poliak A, Dredze M, Leas EC, Zhu Z, Kelley JB, et al. Comparing physician and artificial intelligence chatbot responses to patient questions

[DOI: 10.22088/caspjrm.11.1.37]

posted to a public social media forum. JAMA internal medicine 2023; 183(6): 589-96.
7. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nature medicine 2019 25(1)44-56.
8. Hopkins AM, Logan JM, Kichenadasse G, Sorich MJ. Artificial intelligence chatbots will revolutionize how cancer patients access information: ChatGPT represents a paradigm-shift. JNCI cancer spectrum 2023; 7(2): pkad010.

[DOI: 10.22088/caspjrm.11.1.37]