Assessing artificial intelligence responses to common patient questions regarding inflatable penile prostheses using a publicly available natural language processing tool (ChatGPT)

Nader A. Shayegh, BA,[1] Danae Byer, BS,[1] Yasmine Griffiths, MPH,[1]
Pamela W. Coleman, MD,[2] Leslie A. Deane, MD,[2] Jeremy Tonkin, MD[2]

[1]Howard University College of Medicine, Washington, DC, USA
[2]Department of Surgery, Division of Urology, Howard University College of Medicine, Washington, DC, USA

SHAYEGH NA, BYER D, GRIFFITHS Y, COLEMAN PW, DEANE LA, TONKIN J. Assessing artificial intelligence responses to common patient questions regarding inflatable penile prostheses using a publicly available natural language processing tool (ChatGPT). *Can J Urol* 2024;31(3):11880-11885.

Introduction: *The evolving landscape of healthcare information dissemination has been dramatically influenced by the rise of artificial intelligence (AI) driven chatbots, providing patients with accessible and interactive platforms to obtain knowledge about medical procedures and conditions. Among the various surgical interventions in urology, inflatable penile prosthesis (IPP) is a common treatment for men with erectile dysfunction. As patients increasingly seek comprehensive resources to understand what this procedure entails, AI-based chat technologies, such as ChatGPT, have become more prominent. This study aimed to assess the capacity of ChatGPT to provide accurate and easily understandable responses to common questions regarding the IPP procedure.*
Materials and methods: *Ten frequently asked questions (FAQ) about the IPP procedure were presented to the ChatGPT chatbot in separate conversational sessions without follow up questions or repetitions. An evidence-based approach was employed to assess the accuracy of the chatbot's responses. Responses were categorized as "excellent response not requiring clarification," "satisfactory requiring minimal clarification," "satisfactory requiring moderate clarification," or "unsatisfactory requiring substantial clarification."*
Results: *Upon review, 70% of ChatGPT's answers to questions regarding the IPP procedure were rated as "excellent," not necessitating further clarification. Twenty percent were considered "satisfactory," requiring minimal clarification, notably on the omission of statistical data and the depth of discussion on certain topics. Ten percent of the responses were "unsatisfactory," requiring substantial clarification, including a failure to provide a definitive answer when necessary.*
Conclusions: *This study reveals that ChatGPT has a substantial capability to produce evidence-based, understandable responses to a majority of common questions related to the IPP procedure. While there is room for improvement, ChatGPT can serve as an advantageous tool for patient education, enhancing preoperative understanding and contributing to informed decision-making during urological consultations for IPP.*

Key Words: artificial intelligence, ChatGPT, inflatable penile prosthesis

Introduction

Patient education is crucial in medicine and to the field of urology, particularly for conditions that are of a sensitive nature, such as erectile dysfunction

(ED). Despite the high prevalence of ED, many patients feel uncomfortable seeking professional advice, which leads them to resort to the anonymity of the internet for answers.[1] In this digital quest for privacy, the inflatable penile prosthesis (IPP) procedure emerges as a significant subject of inquiry, representing a durable intervention with decades of evolution and high patient satisfaction rates.[2] The successful placement of an IPP demands meticulous patient education given the procedure's complexity and necessity for rigorous preoperative work up and appropriate patient selection.[3] Amidst the revolution of health information dissemination, artificial intelligence (AI) powered chatbots like ChatGPT have stepped to the forefront, offering an accessible and interactive platform for patient education. Most IPP-related data hails from institutional studies conducted in traditional hospital settings[4] and AI chatbots can expand a patient's access to personalized and trustworthy medical guidance, overcoming obstacles like distance and discomfort in discussing sensitive topics. This research assesses the ability of AI chatbots to bridge the divide in thorough and confidential patient education, focusing on evaluating the efficacy of ChatGPT in communicating details about IPP.[5] This study aims to evaluate the potential of AI as an educational tool that aligns with contemporary needs for privacy, accuracy, and depth in patient-centric health communication.

Materials and methods

Selection of patient questions

To ensure a comprehensive and relevant data set, this study began by identifying and collecting FAQs from ten reputable websites renowned for their patient education materials on urological procedures. These included academic medical centers, specialized urology clinics, and established healthcare organizations like the Mayo Clinic and Cleveland Clinic. Each website's frequently asked questions (FAQ) section was meticulously reviewed to cross-reference the ten most common questions by patients considering the IPP procedure.

Criteria for question inclusion

The selection criteria for the questions included frequency of appearance across different sources, signifying common patient concerns, and the potential of each question to elicit informative responses regarding the IPP procedure. Additionally, particular attention was given to questions that focused on unique or complex aspects of the procedure, thus

providing insights into the depth and scope of ChatGPT's knowledge base.

FAQ compilation process

Following the review, a list of the most prevalent questions was compiled. Each question was cross-referenced across all sources to ensure it represented a broad spectrum of patient inquiries, avoiding site-specific biases. The final list encompassed various topics, from procedural details and postoperative care to potential risks and success rates, aiming to cover a comprehensive picture of the patient's journey through IPP process.

Interaction with ChatGPT

The finalized questions were sequentially presented to the ChatGPT chatbot in individual sessions. Each session was treated independently to prevent any learning effect that could influence the chatbot's responses. Additionally, to mimic a realistic patient chatbot interaction, no follow up questions were asked, and each query was posed without any context or additional information that could lead to the chatbot's response.

Response evaluation

An evidence-based framework was employed to appraise the accuracy and clarity of ChatGPT's responses. A panel of urologists familiar with patient education standards and current IPP procedural guidelines independently reviewed each response. The responses were then categorized into one of four predefined classifications as defined in prior AI medical literature:[6]

1. Excellent response: Responses in this category were deemed to provide comprehensive, accurate information requiring no further clarification.

2. Satisfactory with minimal clarification: These responses were accurate but lacked specific details, such as statistical data, which could be supplemented with minimal additional information.

3. Satisfactory with moderate clarification: Responses under this classification contained the core correct information but required moderate clarification to fill in gaps in explanation or context.

4. Unsatisfactory requiring substantial clarification: These were responses that missed essential information or were vague, necessitating significant further clarification to be considered patient ready.

Results

This study presented the ChatGPT chatbot with ten common questions patients typically ask regarding

the inflatable penile prosthesis (IPP) procedure. Each question was posed a single time to the chatbot without context or repetition to avoid introducing bias through artificial learning. The responses were evaluated and categorized based on their completeness, accuracy, and the need for further clarification. The results are summarized as follows:

Question 1: What are the risks associated with IPP?

Response rated: 2

ChatGPT's response to the question about the risks of IPP was comprehensive, but it did not include specific percentages related to risk prevalence. Therefore, it was rated as satisfactory, with minimal clarification to provide the statistical context that patients may find helpful.

Question 2: How long does a penile implant last?

Response rated: 1

The response to the durability of penile implants was deemed excellent, offering detailed information that would not require further clarification for the patient to have a thorough understanding.

Question 3: When can I have sex after an inflatable penile prosthesis procedure?

Response rated: 1

ChatGPT provided a clear timeline for when sexual activity could be resumed following the IPP procedure. This response was also classified as excellent.

Question 4: Will I lose any length after getting a penile implant?

Response rated: 2

The chatbot's reply regarding penile length after implantation was accurate but required additional detail on the concept of stretched length and nuances surrounding individual variance, hence rated as satisfactory with minimal clarification.

Question 5: Will I be able to have spontaneous erections with a penile implant?

Response rated: 4

The chatbot's response did not directly address the nuances of a prosthesis that cause loss of natural erection or the concept of erection around the device, necessitating further explanation. This was the only response rated as unsatisfactory, requiring substantial clarification.

Question 6: Will anyone notice that I have a penile implant?

Response rated: 1

The question concerning the detectability of the penile implant by others was answered thoroughly, earning an excellent rating.

Question 7: What is the cost of penile implant surgery?

Response rated: 1

Information regarding the cost of the procedure was complete and transparent, with the response receiving an excellent classification.

Question 8: What is the recovery like after an inflatable penile prosthesis procedure?

Response rated: 1

ChatGPT's explanation of the recovery process post-IPP was considered excellent, as it adequately detailed what patients could expect without requiring additional information.

Question 9: What will my partner feel after having an inflatable penile prosthesis procedure?

Response rated: 1

The response to what a partner might feel following the procedure was informative and precise, resulting in an excellent rating.

Question 10: Will an inflatable penile prosthesis procedure impact how my penis feels to me?

Response rated: 1

Lastly, the chatbot successfully addressed concerns about personal sensation post-procedure, and this response was also rated as excellent.

The ChatGPT chatbot delivered mostly excellent or satisfactory responses with minimal clarification. Specifically, 70% of the responses were classified as excellent, 20% as satisfactory with minimal clarification, and 10% as unsatisfactory requiring substantial clarification. These findings indicate that the AI chatbot can provide accurate and understandable information for most IPP-related patient questions. However, the need for improvement in certain areas, such as providing quantitative statistical data and addressing the nuances of specific concerns, was also identified. We have included a complete list of ChatGPT's responses for question 5 and 8 (see Table 1).

Discussion

This investigation into the efficacy of ChatGPT as a source of patient education for those considering an IPP yields insight about the role of artificial intelligence in the communication of healthcare information. The study's results indicate that ChatGPT can deliver comprehensive and understandable responses to most patient inquiries regarding IPP. A remarkable 70% of responses were classified as 'excellent,' signifying a strong alignment with the evidence-based, patient-centric information crucial for informed decision-making.

Notably, the responses that required minimal clarification highlight an opportunity for AI chatbots to incorporate more nuanced discussions, particularly in areas where patients may seek more detailed

TABLE 1. **Question 5 and 8 ChatGPT's responses**

Question 5
Will I be able to have spontaneous erections with a penile implant?

ChatGPT response
Penile implants are devices used to treat erectile dysfunction (ED) and are typically placed inside the penis through surgery. There are two main types of penile implants: inflatable implants and malleable (bendable) implants.
With inflatable penile implants, the device consists of inflatable cylinders that are placed inside the penis. When you want to have an erection, you can manually inflate the cylinders using a pump that is implanted in the scrotum. After the erection, you can deflate the cylinders to return to a flaccid state.
Malleable implants, on the other hand, are rods that are surgically implanted into the erectile chambers of the penis. These implants allow you to manually position your penis into an erect or flaccid state.
The length of your penis after a penile implant surgery should be similar to your natural length when erect. However, the length in the flaccid state might be slightly shorter due to the presence of the implant. The extent of any change in length can vary based on the type of implant and the individual's anatomy. It's essential to discuss these concerns with a qualified healthcare professional or a urologist who specializes in these procedures. They can provide you with personalized information based on your specific situation.

Question 8
What is the recovery like after an inflatable penile prosthesis procedure?

ChatGPT response
Recovery after an inflatable penile prosthesis (IPP) procedure typically involves several stages. It's important to note that the recovery process can vary from person to person, and your healthcare provider will provide specific instructions tailored to your situation. Here's a general overview of what you might expect:
1. Hospital stay: IPP surgery is usually performed in a hospital or surgical center. After the surgery, you may need to stay in the hospital for a day or two for monitoring.
2. Immediate postoperative period: After the surgery, you'll be taken to a recovery room where medical staff will monitor your vital signs. You may experience some pain and discomfort, which will be managed with pain medications.
3. Postoperative care: Your healthcare provider will provide instructions on how to care for the surgical site, including cleaning and dressing changes. It's crucial to follow these instructions to prevent infection.
4. Activity restrictions: You'll likely be advised to avoid strenuous activities, heavy lifting, and sexual activity for a few weeks to allow the surgical site to heal properly. Your healthcare provider will provide specific guidelines based on your condition and the type of surgery performed.
5. Pain management: Pain and discomfort are common after surgery. Your healthcare provider will prescribe pain medications to manage this discomfort. It's important to take the medications as prescribed and report any unusual pain or symptoms to your healthcare provider.
6. Follow up appointments: You'll have follow up appointments with your healthcare provider to monitor your recovery progress. During these appointments, your provider will assess the surgical site, address any concerns you might have, and provide further instructions on resuming normal activities, including sexual activity.
7. Resuming sexual activity: Your healthcare provider will guide you on when it's safe to resume sexual activity. It's essential to follow their recommendations to prevent complications and ensure a successful outcome.
8. Emotional support: Coping with the emotional aspects of IPP surgery, such as adjusting to the device and addressing concerns about sexual function, is crucial. It's important to have open communication with your healthcare provider, partner, or a mental health professional if needed.
Recovery experiences can vary, and some individuals may resume regular activities sooner than others. It's essential to communicate openly with your healthcare provider about your progress, concerns, and any unusual symptoms you may experience during the recovery process.

statistical information. For instance, while the chatbot addressed the general risks associated with IPP, the lack of specific percentages identifies a need for integrating quantitative data within AI algorithms to provide a more comprehensive picture. Moreover, the area of spontaneous erections post-IPP—the only question to receive an 'unsatisfactory' rating—underscores the necessity for AI to be programmed with the sensitivity and specificity to address all facets of patient concerns, including those that involve dispelling misconceptions or providing direct answers to complex questions.

A primary limitation of the study is the use of a predefined set of questions, which might not cover the full range of patient inquiries. Notably, the phrasing of a question can subtly alter the response, a nuance that this study could not capture due to its structured format. ChatGPT's responses were assessed solely on content, without the interactive dynamics typical in real-world patient-provider interactions, such as follow up questions or patient-specific contextual influences. Additionally, it appears to take its responses from generalized medical information without incorporating the wider body of medical literature and interpreting this as a human physician would.

Another limitation is the subjective nature of the response categorization, which relies on the evaluators' judgment and could potentially lead to inter-rater variability. The study also did not include a control group for comparison, such as responses from healthcare professionals, which could have provided a benchmark for evaluating ChatGPT's effectiveness. Furthermore, the absence of direct patient feedback on the responses limits the understanding of how patients might perceive and understand the information provided by the chatbot.

The study does not account for AI technology's dynamic and constantly evolving nature. As AI systems like ChatGPT are continuously updated and improved, the responses generated could vary significantly over time, implying that the findings of this study might represent a specific time point in the chatbot's developmental trajectory rather than its full potential or limitations. The study also did not investigate long term outcomes of using AI in patient education, such as patient satisfaction, retention of information, or impact on surgical decisions and outcomes.

Addressing these limitations in future research would involve expanding the range and format of questions, incorporating interactive elements, engaging directly with patients for feedback, and conducting comparative analyses with traditional educational methods. Despite these limitations, this study provides foundational insights into the potential role of AI in enhancing patient education. It underscores the need for ongoing research as these technologies continue to advance.

The potential for AI to serve as a first point of contact for patient education is promising, but this study also identifies areas for future development. Enhancing the AI's ability to provide quantitative data, query and interpret medical literature and handle nuanced medical topics with greater precision should be a priority. Future research could also explore the integration of AI chatbots within actual clinical settings, measuring their impact on patient outcomes and satisfaction in a real-world context. Additionally, investigating the role of AI in addressing the informational needs of diverse populations, including non-English speakers and those with varying levels of health literacy, would be a valuable extension of this work.

Conclusion

This study contributes to the growing field of AI in patient education within the scope of urology. ChatGPT's performance in providing accurate and intelligible responses to a range of patient questions about IPP is commendable, marking it as a potentially valuable educational tool. Its effectiveness in enhancing preoperative understanding and supporting informed decision-making during urological consultations is apparent. Nonetheless, there is discernible room for enhancement, particularly in offering detailed statistical data and refining the depth of responses to complex questions.

The clinical significance of this study is manifold. It demonstrates the feasibility of utilizing AI as an adjunct to traditional patient education methods, potentially increasing the accessibility and reach of reliable medical information. Moreover, it provides a framework for future AI tools that could further streamline and personalize patient education. By leveraging the capabilities of AI, healthcare providers can augment the preoperative education process, ultimately empowering patients as they navigate the complexities of treatment decisions for ED. As we stand on the cusp of a new era in patient education, this study serves as a steppingstone toward a more informed and engaged patient population in the field of urology.

Note: The authors acknowledge ChatGPT for its contribution to this paper.[7] □

Assessing artificial intelligence responses to common patient questions regarding inflatable penile prostheses using a publicly available natural language processing tool (ChatGPT)

References

1. Xiong Y, Zhang Y, Zhang F, Wu C, Qin F, Yuan J. Applications of artificial intelligence in the diagnosis and prediction of erectile dysfunction: a narrative review. *Int J Impot Res* 2023;35(2):95-102.
2. Rodriguez KM, Kohn TP, Davis AB, Hakky TS. Penile implants: a look into the future. *Transl Androl Urol* 2017;6 (Suppl 5):S860-S866.
3. Chung E, Mulhall J. Practical considerations in inflatable penile implant surgery. *J Sex Med* 2021;18(8):1320-1327.
4. Segal RL, Siegelbaum MH, Lerner BD, Weinberg AC. Inflatable penile prosthesis implantation in the ambulatory setting: a systematic review. *Sex Med Rev* 2020;8(2):338-347.
5. Simon DP, Alter K, Bajic P, Levine LA. Device profile of the Ambicor two-piece inflatable penile prosthesis for treatment of erectile dysfunction: overview of its safety and efficacy. *Expert Rev Med Devices* 2021;18(1):9-14.
6. Mika AP, Martin JR, Engstrom SM, Polkowski GG, Wilson JM. Assessing ChatGPT responses to common patient questions regarding total hip arthroplasty. *J Bone Joint Surg Am* 2023;105(19):1519-1526.
7. OpenAI. ChatGPT. San Francisco: OpenAI; 2023. Available from: https://openai.com/chatgpt [2023 November 27].