ChatGPT in Healthcare: Assessing its Role in Educating Hernia Surgery Patients

Andrés **Torres-Olivieri**, MD[a]; Isabel **Mayorga-Perez**, MD[a]; Luis **Alamo-Irizarry**, MD[a]; Ricardo **Vargas-Figueroa**, BS[c]; Carlos **Crespo-Borges**, BS[b]; Cesar **Luque-Fontanez**, MD[a]

a. Department of Surgery, University of Puerto Rico Medical Sciences Campus, San Juan, PR
b. School of Medicine, University of Puerto Rico Medical Sciences Campus, San Juan, PR
c. School of Medicine, Universidad Central del Caribe, Bayamón, PR

Abstract

Background

Artificial intelligence (AI) has gained increasing relevance in medicine, particularly with the emergence of ChatGPT, an AI-driven language model. Prior studies have evaluated ChatGPT's accuracy in medical information retrieval, demonstrating potential as a patient education tool. However, its reliability in providing accurate and comprehensive information on one of the most common surgical procedures worldwide, inguinal hernia repair, remains unexplored.

Methods

A set of 11 patient-focused questions covering basic knowledge, diagnosis, treatment, and complications related to inguinal hernia repair was queried using ChatGPT-3.5. Responses were assessed by 10 hernia surgeons using a four-tier grading scale: comprehensive, correct but inadequate, partially incorrect, or completely incorrect. Statistical analysis was performed using ANOVA to assess differences in response accuracy across categories.

Results

ChatGPT provided comprehensive responses for the majority of questions, particularly in the treatment category (75%). The second most frequent classification was correct but inadequate, indicating incomplete but generally accurate information. Notably, no responses were graded as completely incorrect, and no significant differences in accuracy were observed between question categories (p = 1.0).

Conclusion

ChatGPT demonstrates strong potential as a reliable patient education tool for inguinal hernia repair, providing predominantly accurate and comprehensive information. While most responses were sufficiently informative, some lacked depth that a surgeon would provide in a clinical setting. This study offers preliminary evidence supporting ChatGPT's utility in bridging the gap in patient education in surgical fields, though further research is needed to validate its role in clinical decision-making and preoperative counseling.

Keywords: Artificial intelligence; ChatGPT; inguinal hernia; patient education; surgical counseling; AI in surgery

Introduction

The integration of artificial intelligence (AI) into healthcare has revolutionized the ways medical professionals interact with patients and make clinical decisions. Among various AI technologies, ChatGPT, a sophisticated language model developed by OpenAI, has emerged as a powerful tool capable of providing accurate and contextually relevant information to both healthcare providers and patients. This article explores the utilization of ChatGPT in surgical contexts, particularly its role in addressing patients' inquiries and facilitating informed decision-making in the patient-centered medical practice.

As healthcare increasingly embraces digital solutions, the reliance on AI-driven platforms like ChatGPT is being studied more frequently. Its application has been studied among several surgical fields, such as otorhinolaryngology, urology, and more [1-5]. These breakthroughs have shown the potential of ChatGPT in serving as a reliable patient information source. However, surgeons have expressed their concerns on the platform's accuracy [3]. This highlights the need for substantial evidence on whether these systems demonstrate consistent accuracy in responding to medical questions. This trust is paramount in surgical cases where patients often seek reassurance and clarity regarding complex procedures and potential outcomes. Moreover, the platform has shown great promise in providing up to date and reliable information in multiple languages [6], thus posing as an invaluable tool for patients facing language barriers in accessing healthcare information.

The application of ChatGPT extends beyond mere patient information; it could support clinical decision-making by offering evidence-based insights that align with established clinical guidelines [7]. By assisting healthcare practitioners in navigating the complexities of surgical protocols, ChatGPT can enhance the quality of care while ensuring adherence to best practices. As AI continues to evolve, the potential for technologies like ChatGPT to bridge communication gaps and support clinical excellence appears boundless. This study focuses on ChatGPT's application in general surgery, particularly inguinal hernia repair.

Methodology

Institutional review board approval was not sought for this prospective study due to the absence of patient-level data. Authors have no affiliation or involvement with OpenAI Inc., the developer of ChatGPT.

Questions included in the study were elaborated by our group of investigators. Our aim was to select a set of questions, which would represent commonly asked questions by patients during a surgeon appointment. Topics addressed were pathophysiology, diagnosis, complications, and risks and benefits of conservative versus surgical treatment, in both open and laparoscopic approaches. A total of 11 questions were eventually compiled.

In this study, we used the ChatGPT 3.5 model, a free version that is accessible to the public. The questions were submitted to ChatGPT within the same conversation. No feedback was provided at any point. All questions asked were in English.

	QUESTIONS
1	What is an inguinal hernia?
2	What types of inguinal hernias exist?
3	What are symptoms related to an inguinal hernia?
4	Is it necessary to use a mesh for my inguinal hernia repair?
5	How is an inguinal hernia diagnosed?
6	Do I need a CT scan prior to getting my inguinal hernia repaired?
7	What are indications for an inguinal hernia repair?
8	What are complications of an inguinal hernia repair?
9	Is surgery absolutely needed for an inguinal hernia?
10	What happens if I don't get my inguinal hernia repaired?
11	What is the difference between laparoscopic and open inguinal hernia repair?

Table 1. Questions selected for inclusion in the study to represent common patient inquiries during surgical consultations about inguinal hernias. These questions were asked to ChatGPT to analyze the responses provided.

Questions, along with their ChatGPT-generated answers, were entered into a Microsoft Form. This form was later sent to 10 hernia surgeons from the University of Puerto Rico General Surgery Department who evaluated the accuracy of the responses to questions using a grading scale, which consists of four options:

A) Comprehensive- defined as accurate, nothing more a general surgeon can add if asked by a patient
B) Correct but inadequate- all information is correct but incomplete; a general surgeon would have more important information to add if asked by a patient
C) Some correct and some incorrect- parts of the answer are correct, while another is completely false
D) Completely incorrect- the entire answer is false

Grading responses obtained for each question were compiled as percentages. These percentages were then entered into R-software, where an ANOVA test was used to determine p-value among the categories.

Results

10 laparoscopic surgeons evaluated 11 question prompts, along with their ChatGPT-generated responses to determine their accuracy. These questions were further categorized into basic knowledge (4 questions), diagnosis (2 questions), treatment (2 questions), and complications/risks (3 questions). **Figure 1** provides a distribution of responses based on these categories.

The highest rate of 'comprehensive' responses was obtained under the treatment category (75%). On the other hand, there were no responses obtained for 'completely incorrect' among all questions. No significant difference was observed between categories regarding response grades (p-value=1).

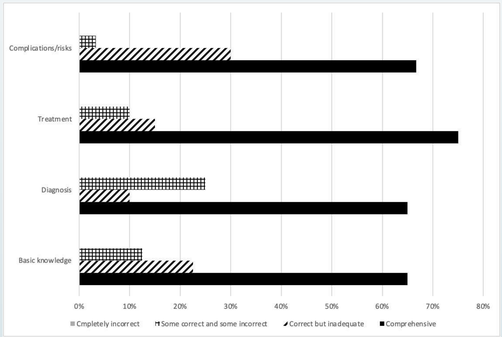


Figure 1. Distribution of accuracy from ChatGPT responses related to inguinal hernias within each category.

Be a part of our study!

We'd love for you to be a part of our study to help us improve and expand our research! If you're a surgeon who frequently performs hernia repair surgeries, simply scan the QR code and complete the short questionnaire. It'll only take a few minutes!

Conclusion

Inguinal hernia repair is one of the most commonly performed surgeries worldwide [8]. However, despite its prevalence, multiple studies have highlighted the lack of quality online patient information on inguinal hernia repair [9-10]. While office visits serve as an ideal time to clarify all potential questions related to surgical procedures, patients often forget or develop new questions after their surgeon appointment. The internet is often navigated in an attempt to find answers, however, search results may be overwhelming, and it may be difficult to identify reliable information [9-10]. The emergence of the AI model ChatGPT serves as a potential platform containing up to date and accurate health related information, which patients could obtain in a timely manner.

We examined the accuracy of ChatGPT in responding to patient-based questions related to inguinal hernia surgery. Questions used covered topics of basic knowledge, diagnosis, treatment, and complications/risks. Our hernia surgeons found that the majority of ChatGPT responses were 'comprehensive'. This trend was observed among all topics covered (**Figure 1**). In all categories, except for diagnosis, the second most common grade was 'correct but inadequate', meaning most of the information was correct, however a surgeon could provide more information. Completeness of this information could potentially be clarified in a follow up office visit. More notably, no surgeon classified any response as completely incorrect.

Our findings suggest that ChatGPT may pose as a reliable source for patients who seek out information related to inguinal hernia repair. These results follow the trend seen in other studies, in which ChatGPT has shown comprehensive responses in other surgical fields. To our knowledge, this is the first study investigating the application of ChatGPT in inguinal hernia. Our results could serve as an early evidence-based study demonstrating the potential of this platform to bridge the gap on patient information among inguinal hernia repairs. While ChatGPT shows substantial promise as an accurate and reliable patient tool, further studies need to be conducted to be able to validate its use in the medical field.

References