Using Artificial Intelligence to Increase Patient Awareness: Evaluating ChatGPT's Responses to Common Patient Questions About Lateral Epicondylitis

Valentin Petrov (✉)[1], Keti Tokmakova[1], Petar Molchovski[1]

[1]Department of Orthopedics and Traumatology, Faculty of Medicine, Medical University of Plovdiv, Plovdiv, Bulgaria

Valentin.Petrov@mu-plovdiv.bg

DOI: 10.58542/jbota.v62i1.174

Abstract— Patients usually have multiple questions related to the diagnosis, treatment options, and expected prognosis. In view of the growing popularity and widespread use of ChatGPT, it is logical to assume that more and more patients will independently seek answers to their medical queries through the chatbot. The current study aims to evaluate the potential of ChatGPT as an additional tool for raising awareness in patients with lateral epicondylitis. **Methods:** Ten hypothetical questions were asked to the free version of ChatGPT-3.5, with the aim of simulating a consultation between a doctor and a patient diagnosed with lateral epicondylitis. Each response generated by ChatGPT was evaluated by two independent orthopedics and traumatology specialists on two separate four-point Likert scales – one for factual accuracy and the other for content relevance. **Results:** The results show that in most cases, ChatGPT generates content with a high degree of factual accuracy and relevance, especially in questions related to symptomatology, causes, and non-operative treatment. Responses related to more specific or expert topics – such as differential diagnosis and surgical treatment – showed some limitations in both accuracy and relevance. Especially important is the result that ChatGPT adapts its responses depending on the level of education indicated by the patient. **Conclusions**: The study shows that ChatGPT can be a useful source of basic medical information on the topic of lateral epicondylitis. However, the use of AI chatbots in a medical context requires clearly stated goals, content verification, and strict adherence to privacy and personal data protection requirements. Additional studies with a broader scope, involving real patients and a variety of clinical trials, are needed cases in order to assess the actual applicability of such technologies in the field of healthcare.

Keywords—ChatGPT, artificial intelligence in medicine, patient education, lateral epicondylitis

1. Introduction

Over the past two decades, the use of online resources for searching for medical and health information has grown significantly, becoming a common practice among patients. Research shows that patients visiting outpatient orthopedic clinics use online sources to search for orthopedic information, reflecting a trend towards greater patient participation in decisions related to their health and treatment. Medical information can increase patients' health culture and awareness of their illness and treatment, while increasing their propensity to follow doctors' prescriptions and recommendations. However, medical information available on the Internet is often characterized by inaccuracies, low relevance, and insufficient scientific value[1].

Immediately after its release for free use in November 2022, ChatGPT (an AI-driven language model) aroused huge interest from users and in less than two months became the fastest-growing application in history, reaching over 100 million active users[2]. Thanks to its ability to generate logical, personalized, and easily understandable responses, ChatGPT finds applications in multiple fields, including education, science, journalism, public health, and medicine. And although research on its benefits for medical education is still in its infancy, there is no doubt that its potential is enormous[3]. Chatbots like ChatGPT can be successfully used by patients as virtual assistants to provide information about diseases and treatments by answering frequently asked questions in different languages and directing to relevant resources. On the other hand, if the content generated by such systems is not systematically evaluated, they can provide inaccurate or misleading information, posing a risk to patients[4]. That is why they are more in-depth studies aimed at assessing the accuracy, reliability and reliability of the, the relevance and clinical usefulness of chatbot-generated responses to real-world medical questions[5].

Goal of the study is to evaluate the potential of ChatGPT as an additional tool for raising awareness of patients with lateral epicondylitis, which is one of the most common diseases in the elbow area caused by Overload and repetitive, uniform hand movements.

The research question that we will seek an answer to in the study is: **Does ChatGPT provide factually accurate and relevant answers to frequently asked questions from patients about lateral epicondylitis?**

2. Materials and Methods

Within the framework of the current study, 10 hypothetical questions were asked to the free version of ChatGPT (model 3.5), with the aim of simulating a consultation between a doctor and a patient diagnosed with lateral epicondylitis (Table 1). The purpose of the questions was to cover all the main aspects of the condition that may be of interest to the patient – symptoms, diagnosis, treatment options and expected development. Each question was entered into ChatGPT with a preliminary phrase modeling the patient's role: "I have been diagnosed with lateral epicondylitis. Can you tell me...?"

It is important to note that ChatGPT uses probabilistic language model, which implies the ability to generate different answers when repeatedly asking identical questions. Therefore, for the purposes of this study, we used the first response generated by the model, without further editing. All questions were entered within one day, from the same user account, and each question was asked in a separate chat session in order to prevent previous conversations from influencing the response.

All responses generated by ChatGPT were evaluated by two independent specialists in orthopedics and traumatology. Each chatbot response was analyzed and evaluated on two separate four-point Likert scales – one for factual accuracy (from 1 – very inaccurate to 4 – very accurate) and the other for content relevance (from 1 – very irrelevant to 4 – very relevant). Particular attention was paid to the relevance of the answers – it was evaluated whether ChatGPT provides targeted and relevant information that effectively answers the specific question without overloading the patient with unnecessary or insignificant details.

Table 1. Questions about lateral epicondylitis asked to ChatGPT

N	**Question**
1.	I have been diagnosed with lateral epicondylitis. Can you tell me what lateral epicondylitis is?
2.	I have been diagnosed with lateral epicondylitis. Can you tell me what causes lateral epicondylitis?
3.	I have been diagnosed with lateral epicondylitis. Can you tell me what are the risk factors for lateral epicondylitis?
4.	I have been diagnosed with lateral epicondylitis. Can you tell me what are the symptoms of lateral epicondylitis?
5.	I have been diagnosed with lateral epicondylitis. Can you tell me other diseases with symptoms similar to lateral epicondylitis?
6.	I have been diagnosed with lateral epicondylitis. Can you tell me specific tests for diagnosing lateral epicondylitis?
7.	I have been diagnosed with lateral epicondylitis. Can you tell me non-surgical treatment of lateral epicondylitis?
8.	I have been diagnosed with lateral epicondylitis. Can you tell me surgical treatment of lateral epicondylitis?
9.	I have been diagnosed with lateral epicondylitis. Can you tell me how long it will take to return to normal activities after lateral epicondylitis treatment?
10.	I have been diagnosed with lateral epicondylitis. Can you tell me if it is possible for lateral epicondylitis to reappear after treatment?

3. Results

Table 2 contains descriptive statistics of the answers of evaluating doctors regarding factual accuracy. Frequency (F), percentage frequency (%) and average (M) are included. The averages of all answers range between 2.5 and 4, indicating that ChatGPT is able to provide correct information in response to typical clinical questions asked by patients. The highest score (M = 4) was reported for the answer to question 4 (symptoms

of lateral epicondylitis). Good results were also reported in question 2 (causes leading to the disease) and question 7 (options for non-operative treatment), both of which received an average score of M = 3.5.

The answer to question 5 (differential diagnosis) received the lowest score (M = 2.5), with reviewers highlighting the inclusion of inaccurate or deviating information, which casts doubt on the effectiveness of ChatGPT in the context of differential diagnosis.

Table 2. Summary of Factual Accuracy Assessments Given by Orthopedics and Traumatology Physicians

Question number	1 - Very inaccurate		2 - Inaccurate		3 – Exactly		4 – Very accurate		M
	F	%	F	%	F	%	F	%	
Question 1					2	100			3
Question 2					1	50	1	50	3.5
Question 3					2	100			3
Question 4							2	100	4
Question 5			1	50	1	50			2.5
Question 6					2	100			3
Question 7					1	50	1	50	3.5
Question 8					2	100			3
Question 9					2	100			3
Question 10					2	100			3

Relevance scores largely coincide with those for factual accuracy (Table 3). Question 4 (symptoms of lateral epicondylitis) and question 9 (recovery time) were given a maximum value (M = 4), and the answers were determined to be highly appropriate for patients without medical education.

The lowest relevance (M = 2) was reported again in question 5 (differential diagnosis) due to the inclusion of some irrelevant examples. The answer to question 8 (surgical treatment) also received a lower relevance score (M = 2.5), due to insufficient specificity and predominantly general wording.

In addition to evaluating standard questions, ChatGPT's ability to adapt the style of its answers was also analyzed according to the patient's educational level, indicated in the formulation of the questions. For this purpose, three different formulations of the questions were used:

1. "I am a patient with primary education and I have been diagnosed with lateral epicondylitis. Can you tell me..."

2. "I am a patient with secondary education and I have been diagnosed with lateral epicondylitis. Can you tell me..."

3. "I am a patient with lateral epicondylitis and I have no medical education. Can you tell me..."

The results show that ChatGPT exhibits linguistic adaptability by modifying the volume, terminology, and style of its responses. In the first option (primary education), the answers were significantly shorter, with simplified language and without specialized terminology. In the second (secondary education) moderate complexity was observed, and the third option led to the most detailed and explanatory answers, including definitions of medical terms and explanations aimed at a complete beginner user.

Table 3. Summary of Relevance Ratings Given by Doctors of Orthopedics and Traumatology

Question number	1 – Very inappropriate		2 - Inappropriate		3 – Appropriate		4 – Very relevant		M
	F	%	F	%	F	%	F	%	
Question 1					2	100			3
Question 2					1	50	1	50	3.5
Question 3					1	50	1	50	3.5
Question 4							2	100	4
Question 5			2	100					2
Question 6					2	100			3
Question 7					1	50	1	50	3.5
Question 8			1	50	1	50			2.5
Question 9							2	100	4
Question 10					2	100			3

3. Discussion

The current study explores ChatGPT's ability to provide accurate and relevant medical information to patients diagnosed with lateral epicondylitis. The results show that in most cases, ChatGPT manages to generate content with a high degree of factual accuracy and relevance, especially in issues related to symptomatology, causes, and non-operative treatment. Similar conclusions are supported in other studies evaluating the effectiveness of ChatGPT in various medical fields. Zhang et al.[1] share that ChatGPT performs well in providing accurate and relevant answers related to knee replacement. A study by Aliyeva et al.[6] analyzed the potential of ChatGPT as an aid in postoperative care for patients with cochlear implants. According to them, the chatbot demonstrates a high degree of accuracy, clarity, and relevance in presenting complex medical infor-

mation in a way that is accessible and understandable to patients. AlShehri et al.[7] conclude that ChatGPT can answer frequently asked questions from patients about hip arthroscopy with satisfactory accuracy.

Responses related to more specific or expert topics – such as differential diagnosis and surgical treatment – showed some limitations in both accuracy and relevance. Similar results have been reported in other studies. Gwak et al.[8] note that ChatGPT encounters difficulties in describing specific symptoms associated with diagnosing impingement shoulder syndrome. Similarly, Shao et al.[9] found inaccuracies in ChatGPT's responses related to both the diagnosis of lung cancer and the description of symptoms associated with surgical complications after thoracic surgery.

Especially important is the result that ChatGPT adapts its responses depending on the level of education indicated by the patient. When explicitly indicating a lack of medical knowledge or a lower educational level, the chatbot generates more accessible, simplified, and explanatory answers that make it easier for the user to understand. This is also confirmed by Guo et al.[10], according to which ChatGPT provides consistent responses that are short and easy to understand. This ability for linguistic adaptation represents a significant advantage in the use of such technologies in a medical context, especially when it is aimed at effective communication with patients who have difficulties with medical terminology.

The results confirm the potential of the chatbot to be used as a tool to increase patients' awareness of their disease, providing largely accurate and relevant information. But they also emphasize the need for such information to be interpreted with caution by the patient, especially when it comes to more complex clinical cases and decisions.

4. Conclusion

The study shows that ChatGPT can be a useful tool for providing basic medical information on the topic of lateral epicondylitis. The chatbot demonstrates the ability to generate relatively accurate and relevant responses related to symptoms, causes, and recovery, while differential diagnosis and surgical treatment questions require more careful expert verification. The ability to adapt their answers to the patient's indicated educational level is a promising feature that improves the accessibility and usefulness of information.

Despite these advantages, the use of AI chatbots in a medical context must be accompanied by clearly defined restrictions and purposes, verification of content, as well as strict adherence to the principles of confidentiality and protection of personal data. Further research with a broader scope, involving real patients and diverse clinical cases, is needed to realistically assess the applicability of such technologies in the field of healthcare.

5. Acknowledgement

This research is supported by the Bulgarian Ministry of Education and Science under the National Program "Young Scientists and Postdoctoral Students – 2".

6. References

1. Zhang, S, Liau, ZQG, Tan, KLM, et al. Evaluating the accuracy and relevance of ChatGPT responses to frequently asked questions regarding total knee replacement. Knee Surgery & Related Research. 2024;36(1):15. https://doi.org/10.1186/s43019-024-00218-5
2. Wu, J, Ma, Y, Wang, J, et al. The application of ChatGPT in medicine: a scoping review and bibliometric analysis. Journal of Multidisciplinary Healthcare. 2024;17:1681-1692. https://doi.org/10.2147/JMDH.S463128
3. Sparks, CA, Fasulo, SM, Windsor, JT, et al. ChatGPT is moderately accurate in providing a general overview of orthopaedic conditions. JB JS Open Access. 2024;9(2): e23.00129 DOI: 10.2106/JBJS. OA.23.00129
4. Thia I, Saluja M. ChatGPT: Is This Patient Education Tool for Urological Malignancies Readable for the General Population? Res Rep Urol. 2024;16:31-37. https://doi.org/10.2147/RRU.S440633
5. Wang, C, Liu, S, Yang, H, et al. Ethical considerations of using ChatGPT in health care. Journal of Medical Internet Research. 2023;25:e48009. doi:10.2196/48009
6. Aliyeva A, Sari E, Alaskarov E, et al. Enhancing Postoperative Cochlear Implant Care With ChatGPT-4: A Study on Artificial Intelligence (AI)-Assisted Patient Education and Support. Cureus. 2024;16(2):e53897. doi: 10.7759/cureus.53897
7. AlShehri Y, McConkey M, Lodhia P. ChatGPT Provides Satisfactory but Occasionally Inaccurate Answers to Common Patient Hip Arthroscopy Questions. Arthroscopy. 2025;41(5):1337-1347. doi: 10.1016/j.arthro.2024.06.017
8. Zhao J, Shao C, Li H, et al. Appropriateness and Comprehensiveness of Using ChatGPT for Perioperative Patient Education in Thoracic Surgery in Different Language Contexts: Survey Study. Interactive journal of medical research. 2023;12:e46900. DOI: https://doi.org/10.2196/preprints.46900
9. Gwak GT, Hwang UJ, Jung, SH, et al. Search for Medical Information and Treatment Options for Musculoskeletal Disorders through an Artificial Intelligence Chatbot: Focusing on Shoulder Impingement Syndrome. J Musculoskelet Sci Technol. 2023;7(1):8-16. https://doi.org/10.29273/jmst.2023.7.1.8
10. Guo AA, Razi B, Kim P, et al. The Role of Artificial Intelligence in Patient Education: A Bladder Cancer Consultation with ChatGPT. Société Internationale d'Urologie Journal. 2024; 5(3):214-224. https://doi.org/10.3390/siuj5030032

Authors:
Dr. Valentin Petrov – Assistant at Medical University Plovdiv, Department of Orthopedics and Traumatology. Resident at University Hospital "Saint George" – Plovdiv, clinic of Orthopedics and Traumatology.

Associate Prof. Keti Tokmakova, PhD - Associate Prof. at Medical University Plovdiv, Department of Orthopedics and Traumatology. Works at University Hospital "Saint George" – Plovdiv, clinic of Orthopedics and Traumatology.

Dr.Petar Molchovski - Assistant at Medical University Plovdiv, Department of Orthopedics and Traumatology. Works at Department of emergency medicine - Orthopedics and Traumatology.