Seizure: European Journal of Epilepsy 114 (2024) 1–8

Contents lists available at ScienceDirect

Seizure: European Journal of Epilepsy

journal homepage: www.elsevier.com/locate/seizure

Assessing the performance of ChatGPT's responses to questions related to epilepsy: A cross-sectional study on natural language processing and medical information retrieval

Hyun-Woo Kim [a], Dong-Hyeon Shin [a], Jiyoung Kim [b,c], Gha-Hyun Lee [b,c], Jae Wook Cho [a,c,*]

[a] Department of Neurology, Pusan National University Yangsan Hospital, 50612 Geumoro 20, Yangsan, South Korea
[b] Department of Neurology, Pusan National University Hospital, Busan, South Korea
[c] Pusan National University School of Medicine, Research Institute for Convergence of Biomedical Science and Technology, Yangsan, South Korea

ARTICLE INFO

Keywords:
Epilepsy
Artificial intelligence
ChatGPT
Natural language processing

ABSTRACT

Background: Epilepsy is a neurological condition marked by frequent seizures and various cognitive and psychological effects. Reliable information is essential for effective treatment. Natural language processing models like ChatGPT are increasingly used in healthcare for information access and data analysis, making it crucial to assess their accuracy.
Objective: This study aimed to investigate the accuracy of ChatGPT in providing educational information related to epilepsy.
Methods: We compared the answers from ChatGPT-4 and ChatGPT-3.5 to 57 common epilepsy questions based on the Korean Epilepsy Society's "Epilepsy Patient and Caregiver Guide." Two epileptologists reviewed the responses, with a third serving as an arbiter in cases of disagreement.
Results: Out of 57 questions, 40 responses from ChatGPT-4 had "sufficient educational value," 16 were "correct but inadequate," and one was "mixed with correct and incorrect" information. No answers were entirely incorrect. GPT-4 generally outperformed GPT-3.5 and was often on par with or better than the official guide.
Conclusions: ChatGPT-4 shows promise as a tool for delivering reliable epilepsy-related information and could help alleviate the educational burden on healthcare professionals. Further research is needed to explore the benefits and limitations of using such models in medical contexts.

1. Introduction

Epilepsy, a neurological disorder affecting approximately 1 % of the global population, disrupts the brain's electrical activity and manifests as recurrent seizures of varying severity [1,2]. Although many patients manage their symptoms with various treatments including medication, diets, immunotherapy, surgery, and neuromodulatory devices, the condition can significantly impair quality of life, affecting factors such as employment, social relationships, and psychological well-being [3–6]. Given the prevalence and impact of epilepsy, accurate and accessible information regarding the condition, including its causes, symptoms, diagnosis, and treatment options, is critical [7,8]. Education about the potential consequences of epilepsy and skill development for self-management in daily life are also important [9]. However, a significant gap exists between the information patients seek and what they can readily access [10]. This often leads individuals to alternative sources like web search engines for information [11,12].

As natural language processing (NLP) models such as ChatGPT (generative pre-training transformer) gain popularity for accessing medical information, more healthcare providers and individuals with epilepsy, are likely to rely on these tools to obtain information regarding the symptoms and conditions of the disease. ChatGPT is a state-of-the-art NLP model that uses deep learning techniques to produce human-like responses to natural language inputs. It is currently one of the largest publicly available NLP models. ChatGPT is based on the GPT-3.5 architecture, which was developed by OpenAI [13]. The model was trained on a massive corpus of text from the internet, including books, articles, and websites, spanning several decades up until 2021. It has found broad acceptance in healthcare among other sectors since its public release in November 2020 [14,15].

* Corresponding author.
E-mail address: sleepcho@pusan.ac.kr (J.W. Cho).

https://doi.org/10.1016/j.seizure.2023.11.013
Received 28 September 2023; Received in revised form 18 November 2023; Accepted 22 November 2023
Available online 23 November 2023

However, ChatGPT's capabilities are pattern-based, lacking the nuanced medical expertise that clinical situations often require [14]. Various studies have produced mixed results regarding its accuracy in medical fields. For instance, a study that examined ChatGPT's responses to inquiries concerning cirrhosis and liver cancer found that the answers provided by ChatGPT regarding diagnosis and prevention were lacking information. However, the responses demonstrated a relatively strong performance in the domains of fundamental knowledge, lifestyle recommendations, and treatment options for the diseases [16]. Another study in the field of pathology demonstrated that ChatGPT achieved satisfactory scores in addressing higher-order reasoning questions [17]. Furthermore, in a study conducted in the United States on cancer information, ChatGPT demonstrated a 96.9 % accuracy in answering questions from the "Common Cancer Myths and Misconceptions" page, indicating that it generally provides satisfactory information on common cancer myths and misconceptions [18]. Other studies have shown that ChatGPT can provide relatively accurate information on the COVID-19 vaccine and play a role in correcting false conspiracy theories [19]. Moreover, ChatGPT has shown promise in radiologic decision-making in breast cancer screenings [18]. However, not all studies showed favorable results. In two simulated exams using questions selected from the American ophthalmology board preparation question banks, ChatGPT-3.5 demonstrated only a modest accuracy of 42.7 %–55.8 % [20]. Moreover, ChatGPT is characterized by providing slightly different responses when asked the same question multiple times, even with the same prompt [21].

While ChatGPT cannot replace professional healthcare providers due to various limitations, it is undeniable that the use of this NLP model for obtaining medical information is likely to increase in the future. Therefore, we aimed to assess the accuracy and reliability of ChatGPT's responses to questions commonly asked by patients with epilepsy or their caregivers. Additionally, we examined its knowledge base and problem-solving skills by assessing its performance when providing educational information related to epilepsy.

2. Methods

2.1. Data source - question selection

For this study, we selected questions from the 2023 edition of the "Epilepsy Patient and Caregiver Guide," published annually by the Korean Epilepsy Society. The guidebook contains 57 questions covering a wide range of topics related to epilepsy, including common symptoms, diagnosis, medications, and management strategies. These questions were classified into the following categories: "Overview of Epilepsy," "Diagnosis," "Treatment," "Seizure Management," "Pediatric Epilepsy," "Women's Health and Epilepsy," "Geriatric Epilepsy," "Advice for Daily Living," and "Myths and Facts about Epilepsy." The guidebook is publicly available for download on the Korean Epilepsy Society website (https://www.kes.or.kr) and was translated into English for this study.. The English translation was performed by a bilingual professional in English education

2.2. ChatGPT

We employed ChatGPT version 4 for this study, even though version 3.5 is currently more prevalent. ChatGPT version 4 is a paid service expected to supersede the earlier version. ChatGPT version 4.0, developed by OpenAI based in the United States, is accessible through the URL https://openai.com/gpt-4. Users can find detailed instructions for its use on this webpage. This version is available via OpenAI's official website or through their API. ChatGPT version 4 operates on a subscription model, with the cost set at $20 per month. This service offers enhanced features and capabilities compared to its predecessors, providing users with a more advanced tool for natural language processing tasks. The questions were typed in English and grouped by category using the "New Chat" feature to consider the interrelatedness of the questions. Each question was entered twice to test the consistency of ChatGPT's responses. After confirming that there were no significant differences in the content of the answers, the first answer was selected for analysis. For comparison, we also tested the responses generated by GPT-3.5. All questions were entered on March 27, 2023.

2.3. Evaluation of responses

The responses were evaluated by two epileptologists with over 10 years of experience in treating patients with epilepsy in tertiary care hospitals. The responses generated by ChatGPT were evaluated and categorized into four grades: 1) "sufficient educational value," 2) "correct but inadequate," 3) "mixed with correct/incorrect/outdated information," and 4) "incorrect." If the opinions of the two epileptologists differed, a third epileptologist with more than 10 years of experience in treating patients with epilepsy served as the final arbiter.

The responses from ChatGPT-3.5 and the content of the "Epilepsy Patient and Caregiver Guide" were also compared and assessed against the responses generated by ChatGPT-4. The performance of ChatGPT-4 was rated as "much better," "better," "similar," "worse," or "much worse" than that of the other two sources. In cases where the evaluations of the two raters differed, the final assessment was derived using the method presented in Table 1. If the evaluations were split between "better" and "worse," the decision was made based on the judgment of a third epileptologist.

2.4. Ethical considerations

Since this study involved the use of publicly available programs and did not involve human participants, it was deemed exempt by the Institutional Review Board.

3. Results

3.1. Evaluation of responses from ChatGPT-4

ChatGPT's responses to all the questions are listed in Supplementary Tables S1–S9. Of 57 questions, 40 responses were classified as having "sufficient educational value," and 16 responses were considered "correct but inadequate." Among the 40 responses with "sufficient educational value," 39 were agreed on by both reviewers, while only one required evaluation by a third reviewer. Only one answer was considered "mixed with correct/incorrect/outdated" information, and none were "incorrect" (Table 2).

Specifically, in the "Diagnosis" category, all four answers were considered to have "sufficient educational value." In the "Seizure Management" category, both answers were classified as having "sufficient educational value." In "Pediatric Epilepsy," five out of six, and in "Women's Health and Epilepsy," 13 out of 14 responses were classified as having "sufficient educational value." In "Myths and Facts about Epilepsy," six out of seven responses were found to have "sufficient

Table 1
Final ratings on the comparison of ChatGPT-4′s responses with ChatGPT-3.5′s responses and "The Epilepsy Patient and Caregiver Guide" in cases of divergent evaluations by two reviewers.

Reviewer 1	Reviewer 2	Final rating
Much better	Better	Better
Much better	Similar	Better
Much better	Worse or much worse	Decided by the third reviewer
Better	Similar	Similar
Better	Worse or much worse	Decided by the third reviewer
Similar	Worse	Similar
Similar	Much worse	Worse
Worse	Much worse	Worse

Table 2
The performance rating of ChatGPT-4′s responses to questions related to epilepsy.

	Questions	Responses from ChatGPT-4			
		Sufficient educational value	Correct but inadequate	Mixed with correct/ incorrect/ outdated data	Completely incorrect
Overview of Epilepsy	What is an epileptic seizure?		✓		
	Are there any famous people in history who had epilepsy?	✓			
	Is epilepsy a common disease?		✓		
	What are the causes of epilepsy?		✓		
	How are epileptic seizures classified?		✓		
Diagnosis	How is epilepsy diagnosed?	✓			
	What is an epileptic aura?	✓			
	What are the essential tests for the diagnosis of epilepsy?	✓			
	What are some other tests that can be helpful in diagnosing epilepsy?	✓			
Treatment	How is medication therapy for epilepsy administered?	✓			
	When should medication therapy for epilepsy be started?		✓		
	How long should medication therapy for epilepsy be continued?	✓			
	In what cases should surgery be considered a treatment for epilepsy?		✓		
	What types of surgery are available for epilepsy treatment?		✓		
	What is ketogenic diet therapy?	✓			
	What is medical marijuana oil?	✓			
Seizure Management	What is the first aid treatment for seizures?	✓			
	Is it dangerous to be next to someone who is having a seizure?	✓			
Pediatric Epilepsy	What are the characteristics of pediatric epilepsy?	✓			
	What should I do if my child has a seizure?	✓			
	What is the typical course of action when someone with epilepsy goes to the hospital after having a seizure?	✓			
	Is the administration of anticonvulsants the same for pediatric patients as it is for adults?	✓			
	What are some treatment options for refractory epilepsy?	✓			
	Is febrile seizure considered a type of epilepsy?		✓		
Women's Health and Epilepsy	Is there a relationship between epilepsy, abnormal menstrual cycles, and changes in female hormones?	✓			
	What are the things to know about contraception when managing epilepsy?	✓			
	I'm considering getting pregnant. What should I do if I have epilepsy?	✓			
	I'm concerned about the risk of birth defects.		✓		
	What prenatal diagnostic tests should be done if I have epilepsy?	✓			
	Is there a risk of intellectual disability for the fetus due to my epilepsy?	✓			
	Should I continue taking anticonvulsant medication during pregnancy?	✓			
	How should anticonvulsant medication be taken during pregnancy?	✓			
	What if there are seizures during pregnancy?	✓			
	What about folic acid and vitamin K supplementation during pregnancy for someone with epilepsy?		✓		
	How about labor pains and childbirth with epilepsy?	✓			
	What about postpartum care and breastfeeding if I have epilepsy?	✓			
	How can I take care of my child if I have epilepsy?	✓			
	What about menopause and osteoporosis if I have epilepsy?	✓			
Geriatric Epilepsy	Is long-term use of anticonvulsants associated with an increased risk of dementia?		✓		
	Is it okay to take the antiepileptic medication with other drugs?		✓		
Advice for Daily Living	How should I get vaccinated as someone with epilepsy?		✓		
	Can I drink alcohol if I have epilepsy?	✓			
	I have a cold. What should I do if I have epilepsy?	✓			
	Does taking anticonvulsants affect cognitive ability?	✓			
	Are there any specific foods to avoid while taking anticonvulsants?	✓			
	Is it safe for someone with epilepsy to exercise?		✓		
	Can someone with epilepsy drive a car?		✓		
	Are alternative therapies effective in treating epilepsy?	✓			
	I am concerned about workplace issues related to my epilepsy.	✓			
	Can someone with epilepsy join the military?		✓		
Myths and Facts about Epilepsy	Is epilepsy a curable condition?			✓	
	Are there forms of epilepsy that are difficult to treat?	✓			
	Do epileptic seizures cause changes in brain injuries?	✓			
	Can epilepsy cause intellectual disability or mental illness?	✓			

(*continued on next page*)

Table 2 (*continued*)

	Questions	Responses from ChatGPT-4			
		Sufficient educational value	Correct but inadequate	Mixed with correct/ incorrect/ outdated data	Completely incorrect
	What types of epilepsy are more easily treatable and have a better prognosis?	✓			
	Is epilepsy contagious to other people?	✓			
	Is epilepsy an inherited genetic disease that can be passed down to one's children?	✓			

educational value." In the "Treatment" category, four out of seven responses were classified as having "sufficient educational value," while three were "correct but inadequate." In the "Overview of Epilepsy" category, only one out of five responses had "sufficient educational value," and four were "correct but inadequate," which was considered relatively insufficient in terms of information provided compared to that provided for other categories.

3.2. Comparison of ChatGPT-4′s responses with ChatGPT-3.5′s responses and the "Epilepsy patient and caregiver guide"

When comparing the responses of ChatGPT-4 with ChatGPT-3.5, of the 57 questions, there were no answers produced by ChatGPT-4 that were "much better" than those generated by ChatGPT-3.5. However, 10 answers were "better," 45 were "similar," and two were "worse". Therefore, ChatGPT-4 exhibited similar or improved performance compared to that of ChatGPT-3.5. ChatGPT-4 performed worse than ChatGPT-3.5 on the following two questions: "Is it okay to take antiepileptic medication with other drugs?" and "How should I get vaccinated as someone with epilepsy?" (Table 3).When compared to the official guide, ChatGPT-4 provided a "much better" response to the question, "I am concerned about workplace issues related to my epilepsy." Additionally, it provided 32 "better" responses and 20 "similar" responses. Four responses were "worse" compared to the official guide, but none were "much worse." Overall, ChatGPT-4′s responses were either better or at a similar level than those of the official guide. In the "Overview" category, the information from the official guide was either similar or better than that provided by ChatGPT-4, while the remaining categories had a comparable distribution of "better" and "similar" responses from ChatGPT-4. The four questions to which the official guide provided better answers than ChatGPT-4 were: "Is epilepsy a common disease?", "When should medication therapy for epilepsy be started?", "Is febrile seizure considered a type of epilepsy?", and "Can someone with epilepsy join the military?".

4. Discussion

ChatGPT can comprehend natural language and provide tailored responses based on user input. In this study, the AI was trained on a database of epilepsy-related information and commonly asked questions. The results showed that ChatGPT was able to provide accurate answers to questions about epilepsy, such as the causes, symptoms, and treatment of the disease. It also provided practical advice on how to manage seizures and handle emergency situations. Overall, the study suggests that generative AI, such as ChatGPT, can be a valuable tool in providing information and support to epilepsy patients and their caregivers.

For the "Overview" category, which provides general information on the disease, the responses received a slightly weaker evaluation compared to those of the other categories. Although in-depth content is not necessary for the general public, the information was somewhat insufficient due to its simple, listing-style presentation. When asked about famous individuals known to have had epilepsy, the response was noteworthy, as it listed names and explained the limitations of relying on historical descriptions, stating that: "It is important to remember that diagnosing historical figures with epilepsy can be speculative, as our understanding of the disorder has evolved over time and accurate medical records may not be available." Although the responses generated by ChatGPT-4 were similar in quality to those produced by its predecessor, ChatGPT-3.5, the official guide provided the most adequate information, likely because the guide includes illustrations that make it easier for readers to understand.

The "Diagnosis" section received a positive evaluation, with all responses providing accurate information. Although there is always a risk of outdated or incorrect information when using any source, including AI, the fact that ChatGPT was effective in providing useful and accurate information despite being trained on data up to 2019 is promising. This content was of higher quality than the information in the official guide. Answers related to "Treatment" were generally good as well. However, regarding some of the most common concerns, such as when to start medication and when to consider surgery, ChatGPT-4 elaborated on various possibilities but lacked a clear explanation of the general concept. In this regard, the official guide was more concise and easier to understand than ChatGPT-4, likely owing to its use of charts and figures in its explanations. Nevertheless, ChatGPT-4 provided more detailed answers for specific treatments such as the ketogenic diet and marijuana oil.

In the "Seizure Management" section, ChatGPT-4 provided detailed and accurate information on how to respond to a patient having a seizure and addressed concerns about the safety of being near a person having a seizure. This response was considered superior to that of the official guide. ChatGPT-4 also provided useful information for specific populations, such as children, women, and elderly patients. Although the explanation on whether febrile seizures are a type of epilepsy was relatively detailed, it was considered inadequate compared to the content provided in the official guide.

ChatGPT-4′s responses to questions related to "Advice for Daily Living" and "Myths and Facts about Epilepsy" were particularly impressive. Patients with epilepsy often have many questions about everyday issues such as alcohol consumption, exercise, diet, vaccination eligibility, and driving [9] but may not always obtain sufficient information from healthcare providers during consultations. ChatGPT-4 provided specific information on these topics, demonstrating that chatbots can be helpful tools to alleviate the burden on healthcare providers. Although ChatGPT-4′s response that epilepsy is not a curable condition was unsatisfactory, it provided appropriate information on various prejudices and misconceptions regarding epilepsy. Overall, the responses generated by ChatGPT-4 were either better than or comparable to those provided in the official guide.

In recent years, there has been an increasing reliance on the Internet as a source of health information. Individuals with epilepsy often use Google to search for information, such as diagnosis, treatment options, and medication side effects [22–24]. YouTube has also become a popular source of information for epilepsy-related content, including personal experiences, educational videos, and patient advocacy [12,25,26].

The increasing popularity of ChatGPT worldwide is expected to lead

Table 3
Comparison of ChatGPT-4′s responses with ChatGPT-3.5′s responses and the "Epilepsy Patient and Caregiver Guide".

	Questions	Compared to									
		GPT 3.5					Official guide				
		Much better	Better	Similar	Worse	Much worse	Much better	Better	Similar	Worse	Much worse
Overview of Epilepsy	What is an epileptic seizure?			✓					✓		
	Are there any famous people in history who had epilepsy?			✓					✓		
	Is epilepsy a common disease?			✓						✓	
	What are the causes of epilepsy?			✓					✓		
	How are epileptic seizures classified?		✓						✓		
Diagnosis	How is epilepsy diagnosed?			✓					✓		
	What is an epileptic aura?			✓				✓			
	What are the essential tests for the diagnosis of epilepsy?			✓				✓			
	What are some other tests that can be helpful in diagnosing epilepsy?			✓					✓		
Treatment	How is medication therapy for epilepsy administered?		✓					✓			
	When should medication therapy for epilepsy be started?		✓							✓	
	How long should medication therapy for epilepsy be continued?			✓				✓			
	In what cases should surgery be considered a treatment for epilepsy?			✓					✓		
	What types of surgery are available for epilepsy treatment?			✓					✓		
	What is ketogenic diet therapy?			✓					✓		
	What is medical marijuana oil?			✓				✓			
Seizure Management	What is the first aid treatment for seizures?			✓					✓		
	Is it dangerous to be next to someone who is having a seizure?			✓				✓			
Pediatric Epilepsy	What are the characteristics of pediatric epilepsy?			✓				✓			
	What should I do if my child has a seizure?			✓					✓		
	What is the typical course of action when someone with epilepsy goes to the hospital after having a seizure?			✓				✓			
	Is the administration of anticonvulsants the same for pediatric patients as it is for adults?		✓					✓			
	What are some treatment options for refractory epilepsy?		✓					✓			
	Is febrile seizure considered a type of epilepsy?		✓							✓	
Women's Health and Epilepsy	Is there a relationship between epilepsy, abnormal menstrual cycles, and changes in female hormones?		✓					✓			
	What are the things to know about contraception when managing epilepsy?			✓				✓			
	I'm considering getting pregnant. What should I do if I have epilepsy?			✓				✓			
	I'm concerned about the risk of birth defects.			✓					✓		
	What prenatal diagnostic tests should be done if I have epilepsy?			✓				✓			
	Is there a risk of intellectual disability for the fetus due to my epilepsy?			✓				✓			
	Should I continue taking anticonvulsant medication during pregnancy?			✓				✓			
	How should anticonvulsant medication be taken during pregnancy?			✓				✓			
	What if there are seizures during pregnancy?			✓					✓		
	What about folic acid and vitamin K supplementation during pregnancy for someone with epilepsy?			✓					✓		
	How about labor pains and childbirth with epilepsy?			✓				✓			
	What about postpartum care and breastfeeding if I have epilepsy?			✓				✓			
	How can I take care of my child if I have epilepsy?			✓				✓			
	What about menopause and osteoporosis if I have epilepsy?		✓					✓			
Geriatric Epilepsy	Is long-term use of anticonvulsants associated with an increased risk of dementia?			✓				✓			

(*continued on next page*)

Table 3 (*continued*)

	Questions	Compared to									
		GPT 3.5					Official guide				
		Much better	Better	Similar	Worse	Much worse	Much better	Better	Similar	Worse	Much worse
	Is it okay to take the antiepileptic medication with other drugs?				✓			✓			
Advice for Daily Living	How should I get vaccinated as someone with epilepsy?				✓				✓		
	Can I drink alcohol if I have epilepsy?			✓				✓			
	I have a cold. What should I do if I have epilepsy?			✓				✓			
	Does taking anticonvulsants affect cognitive ability?			✓				✓			
	Are there any specific foods to avoid while taking anticonvulsants?			✓				✓			
	Is it safe for someone with epilepsy to exercise?			✓				✓			
	Can someone with epilepsy drive a car?			✓					✓		
	Are alternative therapies effective in treating epilepsy?			✓				✓			
	I'm concerned about workplace issues related to my epilepsy.			✓			✓				
	Can someone with epilepsy join the military?			✓						✓	
Myths and Facts about Epilepsy	Is epilepsy a curable condition?			✓					✓		
	Are there forms of epilepsy that are difficult to treat?		✓					✓			
	Do epileptic seizures cause changes in brain injuries?			✓					✓		
	Can epilepsy cause intellectual disability or mental illness?			✓				✓			
	What types of epilepsy are more easily treatable and have a better prognosis?		✓						✓		
	Is epilepsy contagious to other people?			✓				✓			
	Is epilepsy an inherited genetic disease that can be passed down to one's children?			✓					✓		

to a greater demand for obtaining medical information through similar chatbot tools in the future [27]. This includes not only patients but also healthcare professionals seeking to engage in scientific research [14]. However, the accuracy and reliability of online information on epilepsy can vary greatly, and patients should be cautious when relying on these sources for medical advice [12,28,29].

Language models like ChatGPT may not always provide accurate or nuanced information because the goal of the AI training was primarily to achieve general-purpose cognitive abilities and not for use in medical or healthcare applications. Furthermore, ChatGPT has certain limitations, such as the inability to evolve continuously in real-time since it was trained on a limited dataset up to 2021, which may contain inaccuracies [30]. Therefore, it is important to carefully evaluate the accuracy and reliability of the information provided by these tools and encourage users to seek additional sources of information and advice as needed. Although there were some shortcomings in answering various questions about epilepsy, our study showed that ChatGPT-4 provided accurate answers in most cases, with no "incorrect" answers. The results demonstrated that ChatGPT-4 can be a valuable tool for providing disease-related information and daily living advice, and for addressing various prejudices related to epilepsy.

Additionally, GPT-4 provided responses that were either better or at par with those of version 3.5 for most of the questions. Improvements were observed not only in the level of detail of the information provided but also in the way it was expressed. Specifically, there was a distinct change in how the AI began its answers with phrases like "I am not a medical professional, but…" or "I am not a doctor, but…" These changes reflect that ChatGPT-4 has improved its ability to recognize its limitations and communicate them clearly to the user. Additionally, when a question stated, "I am concerned about workplace issues related to my epilepsy," GPT-4 empathetically began its response with "It is understandable to have concerns about workplace issues related to epilepsy," which was not seen in version 3.5. ChatGPT-4 is characterized by a substantially larger model size, an increased number of parameters, and enhanced training techniques, which collectively contribute to its improved accuracy, fluency, and context understanding. Consequently, ChatGPT-4 demonstrates a marked improvement in overall performance, providing users with a more engaging and coherent conversational experience [31,32].

While this study highlights the positive aspects of ChatGPT-4, it should not be considered a substitute for professional medical advice. The source of the AI training was text-based data from open sources on the Internet, such as publicly available medical documents, research articles, health-related websites, and health-related podcasts and videos [27]. Therefore, while it can provide general medical information, it may lack the clinical expertise required to provide tailored advice for a patient's individual condition. Future research may focus on examining the appropriateness and value of responses when asked about individual situations of patients with epilepsy.

The educational efficacy of ChatGPT's responses is likely due to its training on official guidelines from various associations and medical sites on the internet. While ChatGPT's ability to utilize official, publicly accessible documents for its responses is a considerable advantage, the absence of explicit references presents a noteworthy limitation. This lack of transparency about sources emphasizes the need for caution when relying exclusively on AI tools for acquiring medical information.

While this study sheds light on ChatGPT's ability to address questions, we acknowledge the need for future investigations into more individualized and less-mainstream queries. However, assessing the accuracy of responses to these nuanced questions will require rigorous study designs to facilitate medical consensus, underlining the complexity of integrating AI tools in healthcare information dissemination.

As the performance of AI-powered chatbots continues to improve

and the amount of information available for training increases, it will be important to explore how well they can provide customized answers to the unique circumstances of each patient.

4.1. Limitations

This study has several limitations. First, it relies on subjective judgment by reviewers. Since ChatGPT's responses are composed of narratives and cannot be quantified, the assessment could vary if conducted by a different evaluator. While this study acknowledges the subjective nature of evaluating ChatGPT responses, future research could enhance objectivity by employing several methodologies. These include integrating quantitative metrics like accuracy and relevance scores, utilizing text mining techniques for thematic analysis, and adopting a consensus approach with multiple evaluators. Additionally, employing machine learning algorithms for response classification and conducting longitudinal studies to observe response consistency over time may offer further insights. Such multifaceted approaches could significantly reduce subjective bias in evaluating AI-generated medical information. Second, ChatGPT's responses may vary depending on factors such as the patient's country, location, social status, and medical infrastructure. To compensate for this, we compared the answers with the official published guide. Practically, exploring how ChatGPT responds to similar medical queries starting with 'I live in Korea…' versus 'I live in the United States/South Africa/Australia…' offers a promising direction for future research. Such studies could reveal how ChatGPT's answers are shaped by regional contexts and its adaptability to various healthcare environments. This exploration will be a focal point in our future research endeavors, aiming to enhance our understanding of AI in providing culturally relevant and geographically specific medical information. Lastly, due to its cross-sectional design, ChatGPT's responses might vary when the same questions are reiterated after the model gains more information or with future updates.

Conclusions

This study effectively highlights ChatGPT-4′s potential as a valuable tool for conveying general medical knowledge related to epilepsy to the general public. Although concerns about accuracy, copyright, and the AI's capacity to provide tailored information persist, these models can be regarded as tools that alleviate the educational burdens on healthcare professionals in terms of patient education. In the foreseeable future, chatbot-style tools like ChatGPT-4 are expected to gain increased traction in the sphere of medical information delivery. Consequently, further research is essential to fully comprehend the potential advantages and limitations of employing language models in healthcare and medicine.

Funding information

This research was supported by a grant from the Korean Health Technology R&D Project through the Korean Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI21C0852), and supported by Research institute for Convergence of biomedical science and technology, Pusan National University Yangsan Hospital (grant number : 30-2021-13).

CRediT authorship contribution statement

Hyun-Woo Kim: Methodology, Writing – original draft, Writing – review & editing. **Dong-Hyeon Shin:** Data curation. **Jiyoung Kim:** Conceptualization, Data curation, Methodology, Writing – review & editing. **Gha-Hyun Lee:** Methodology. **Jae Wook Cho:** Conceptualization, Data curation, Methodology, Methodology, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

We express our gratitude to Splash English for their assistance in the English translation of the guidebook.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.seizure.2023.11.013.

References

[1] Shorvon SD, Andermann F, Guerrini R. The causes of epilepsy: common and uncommon causes in adults and children. Cambridge University Press; 2011.
[2] Beghi E, Giussani G, Nichols E, et al. Global, regional, and national burden of epilepsy, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol 2019;18(4):357–75.
[3] Berto P. Quality of life in patients with epilepsy and impact of treatments. Pharmacoeconomics 2002;20:1039–59.
[4] Smeets VM, van Lierop BA, Vanhoutvin JP, et al. Epilepsy and employment: literature review. Epilepsy Behav 2007;10(3):354–62.
[5] Elwes R, Marshall J, Beattie A, et al. Epilepsy and employment. A community based survey in an area of high unemployment. J Neurol, Neurosurg Psychiatry 1991;54 (3):200–3.
[6] Jacoby A. Stigma, epilepsy, and quality of life. Epilepsy Behav 2002;3(6):10–20.
[7] Couldridge L, Kendall S, March A. A systematic overviewa decade of research'. The information and counselling needs of people with epilepsy. Seizure 2001;10(8): 605–14.
[8] Lewis SA, Noyes J, Mackereth S. Knowledge and information needs of young people with epilepsy and their parents: mixed-method systematic review. BMC Pediatr 2010;10(1):1–17.
[9] Henning O, Alfstad KA, Nakken KO, et al. A call for better information about epilepsy: the patients' perspective—an online survey. Seizure 2019;69:173–9.
[10] Henning O, Nakken KO, Lossius MI. People with epilepsy and their relatives want more information about risks of injuries and premature death. Epilepsy Behav 2018;82:6–10.
[11] Arora VS, McKee M, Stuckler D. Google Trends: opportunities and limitations in health and health policy research. Health Policy (New York) 2019;123(3):338–41.
[12] Madathil KC, Rivera-Rodriguez AJ, Greenstein JS, et al. Healthcare information on YouTube: a systematic review. Health Informat J 2015;21(3):173–94.
[13] OpenAI. Models GPT-3.5, URL : https://openai.com/blog/chatgpt/ [*accessed* 2023-04-15].
[14] Aydın Ö., Karaarslan E. OpenAI ChatGPT generated literature review: digital twin in healthcare. Available at SSRN 4308687 2022.
[15] Vaishya R, Misra A, Vaish A. ChatGPT: is this version good for healthcare and research? Diabetes Metab Syndr: Clin Res Rev 2023:102744.
[16] Yeo YH, Samaan JS, Ng WH, et al. Assessing the performance of ChatGPT in answering questions regarding cirrhosis and hepatocellular carcinoma. medRxiv 2023. 2023.02. 06.23285449.
[17] Sinha RK, Roy AD, Kumar N, et al. Applicability of ChatGPT in assisting to solve higher order problems in pathology. Cureus 2023;15(2).
[18] Johnson SB, King AJ, Warner EL, et al. Using ChatGPT to evaluate cancer myths and misconceptions: artificial intelligence and cancer information. JNCI Cancer Spectrum 2023;7(2).
[19] Sallam M, Salim NA, Ala'a B, et al. ChatGPT output regarding compulsory vaccination and COVID-19 vaccine conspiracy: a descriptive study at the outset of a paradigm shift in online search for information. Cureus 2023;15(2).
[20] Antaki F, Touma S, Milad D, et al. Evaluating the performance of ChatGPT in ophthalmology: an analysis of its successes and shortcomings. medRxiv 2023. 2023.01. 22.23284882.
[21] Hopkins AM, Logan JM, Kichenadasse G, et al. Artificial intelligence chatbots will revolutionize how cancer patients access information: chatGPT represents a paradigm-shift. JNCI Cancer Spectrum 2023;7(2).
[22] Moalong KMC, Jamora RDG, Roberto KT, et al. Patterns of Google search behavior for epilepsy and seizures in the Philippines: an infodemiological study. Epilepsy Behav 2021;125:108377.
[23] Brigo F, Igwe SC, Ausserer H, et al. Why do people Google epilepsy?: an infodemiological study of online behavior for epilepsy-related search terms. Epilepsy Behav 2014;31:67–70.
[24] Roberto KT, Jamora RDG, Moalong KMC, et al. Infodemiology of autoimmune encephalitis, autoimmune seizures, and autoimmune epilepsy: an analysis of online search behavior using Google Trends. Epilepsy Behav 2022;132:108730.
[25] Oh J, You SY. Febrile seizure: what information can caregivers access through YouTube? Seizure 2021;91:91–6.

[26] Muhammed L, Adcock JE, Sen A. YouTube as a potential learning tool to help distinguish tonic–clonic seizures from nonepileptic attacks. Epilepsy Behav 2014; 37:221–6.
[27] Lee P, Bubeck S, Benefits Petro J. Limits, and Risks of GPT-4 as an AI Chatbot for Medicine. N Engl J Med 2023;388(13):1233–9.
[28] Brna P, Dooley J, Esser M, et al. Are YouTube seizure videos misleading? Neurologists do not always agree. Epilepsy Behav 2013;29(2):305–7.
[29] Lo AS, Esser MJ, Gordon KE. YouTube: a gauge of public perception and awareness surrounding epilepsy. Epilepsy Behav 2010;17(4):541–5.
[30] Sallam M. ChatGPT utility in health care education, research, and practice: systematic review on the promising perspectives and valid concerns. MDPI; 2023.
[31] OpenAI. Gpt-4 technical report, URL : https://cdn.openai.com/papers/gpt-4.pdf [Last accessed 2023-05-03].
[32] OpenAI. Gpt-4 system card, URL : https://cdn.openai.com/papers/gpt-4-system-card.pdf [Last accessed 2023-05-03].