Evaluation of the quality and readability of ChatGPT responses to frequently asked questions about myopia in traditional Chinese language

Li-Chun Chang[1,2,3]**, Chi-Chin Sun**[4,5]**, Ting-Han Chen**[6]**, Der-Chong Tsai**[7,8,9]**, Hui-Ling Lin**[1,2,3,10] **and Li-Ling Liao**[11,12] (iD)

Abstract

Introduction: ChatGPT can serve as an adjunct informational tool for ophthalmologists and their patients. However, the reliability and readability of its responses to myopia-related queries in the Chinese language remain underexplored.

Purpose: This study aimed to evaluate the ability of ChatGPT to address frequently asked questions (FAQs) about myopia by parents and caregivers.

Method: Myopia-related FAQs were input three times into fresh ChatGPT sessions, and the responses were evaluated by 10 ophthalmologists using a Likert scale for appropriateness, usability, and clarity. The Chinese Readability Index Explorer (CRIE) was used to evaluate the readability of each response. Inter-rater reliability among the reviewers was examined using Cohen's kappa coefficient, and Spearman's rank correlation analysis and one-way analysis of variance were used to investigate the relationship between CRIE scores and each criterion.

Results: Forty-five percent of the responses of ChatGPT in Chinese language were appropriate and usable and only 35% met all the set criteria. The CRIE scores for 20 ChatGPT responses ranged from 7.29 to 12.09, indicating that the readability level was equivalent to a middle-to-high school level. Responses about the treatment efficacy and side effects were deficient for all three criteria.

Conclusions: The performance of ChatGPT in addressing pediatric myopia-related questions is currently suboptimal. As parents increasingly utilize digital resources to obtain health information, it has become crucial for eye care professionals to familiarize themselves with artificial intelligence-driven information on pediatric myopia.

Keywords

ChatGPT, myopia, ophthalmologists, health education, quality

[1]School of Nursing, Chang Gung University of Science and Technology, Gueishan
[2]School of Nursing, College of Medicine, Chang Gung University, Tao-Yuan
[3]Department of Nursing, Linkou Chang Gung Memorial Hospital, Linkou
[4]Department of Ophthalmology, Chang Gung Memorial Hospital, Keelung
[5]School of Medicine, College of Medicine, Chang Gung University, Taoyuan
[6]Chang Gung University of Science and Technology, Taoyuan City
[7]Faculty of Medicine, National Yang Ming Chiao Tung University, School of Medicine, Taipei
[8]Community Medicine Research Center and Institute of Public Health, National Yang Ming Chiao Tung University, Taipei

[9]Department of Ophthalmology, National Yang Ming Chiao Tung University Hospital, Yilan
[10]Taipei Medical University, Taipei
[11]Department of Public Health, College of Health Science, Kaohsiung Medical University, Kaohsiung City
[12]Department of Medical Research, Kaohsiung Medical University Hospital, Kaohsiung City

Corresponding author:
Li-Ling Liao, Department of Public Health, College of Health Science, Kaohsiung Medical University, Kaohsiung City; Department of Medical Research, Kaohsiung Medical University Hospital, Kaohsiung City.
Email: liling@kmu.edu.tw

Introduction

ChatGPT, a large language model (LLM) developed by artificial intelligence (AI) using supervised and reinforcement learning techniques, is a tool that generates human-like text responses to human input with high accuracy, accessibility, and low cost.[1] Due to its accurate and well-formulated responses to various queries,[2] ChatGPT has rapidly gained global attention. Of particular interest is its ability to grasp natural language in the field of healthcare.[3] ChatGPT has been used to support medical examinations, confirm clinical diagnoses, support clinical decision-making, and provide professional advice for patient education.[4] Due to its wide recognition and acceptance, ChatGPT is widely used to disseminate medical information to the general public. It can serve as a valuable tool for healthcare professionals in providing patient education and comprehensively addressing patient concerns. The evaluation of the responses of ChatGPT to health queries through grading by an independent physician for accuracy and reproducibility has gained considerable attention, paving the way for the potential integration of ChatGPT with patient education strategies.[5,6]

The global prevalence of myopia, a condition that causes visual impairment, is expected to increase significantly to approximately five billion affected individuals in 2050.[7] Within educational systems emphasizing high academic achievement expectations, myopia is the most prevalent condition encountered in pediatric ophthalmology practices and has emerged as a public health issue among schoolchildren in Chinese-speaking societies with a prevalence far exceeding that observed in western countries.[7,8]

Parents often lack knowledge about the diagnosis and treatment of myopia, and this presents challenges that may not always be fully addressed within the busy clinical settings of ophthalmologists.[9] Online discussion fora and websites are an easily accessible source of information for parents of myopic children.[10] Understanding parents' discussions about myopia on the internet helps focus on their needs within the broader context of myopia issues. The comprehensibility and accuracy of the answers are significant parameters of appropriateness to address.[11] Accurate online medical information enhances the comprehension of conditions and guides patients toward suitable treatments. Similarly, incorrect or misleading information can lead to confusion, delayed treatment, or the choice of detrimental "therapies." Therefore, appropriateness and readability are key factors in health information provision.[12] Health education information read by patients should be accurate, reliable, and current, whereas highly readable information is more accessible to readers. Consequently, the ability to apply information easily to address specific health issues is a core necessity for information usability.[13]

A study of the responses of ChatGPT to 28 questions related to childhood myopia showed an acceptability rate of 87.5%.[14] Biswas et al.[6] demonstrated that ChatGPT generates responses to 11 myopia FAQs with an acceptability rate of >80%. In addition to professional assessment, the readability of responses as judged by analyzing their vocabulary and grammatical features using readability formulas is an important evaluation criterion.[15] The number of investigations evaluating the reliability of ChatGPT responses for patient education is increasing but the number of studies on ChatGPT responses to health issues in the Chinese language is limited. We aimed to evaluate the reliability and readability of childhood myopia-associated information provided to parents by ChatGPT in the Chinese language.

Methods

Data collection

This study collected public posts from a private Facebook group on pediatric myopia over one year from March 2022 to March 2023. Using keywords such as "myopia," "diagnosis," "treatment," "atropine," "deterioration," "spectacle," "improvement," and "follow-up," the top 20 most frequently asked questions in the published articles were identified according to the number of responses. The inclusion criteria were as follows: (1) posts in traditional Chinese; (2) issues related to the diagnosis, prevention, and treatment of childhood myopia; and (3) posts with responses by more than 100 individuals. The exclusion criteria included advertisements related to myopia, inquiries about treatment costs, and product recommendations (Table 1).

Twenty questions standardized by the researchers were selected and classified into six categories: (1) diagnosis; (2) atropine; (3) spectacles; (4) orthokeratology; (5) effective control, and (6) others (Table 2).

ChatGPT response generation

ChatGPT (version GPT-3.5, OpenAI) was queried three times (i.e. in three separate "trials")[16] for each question in traditional Chinese language on 16 April 2023. An example of a prompt entered into the ChatGPT interface is "I am a parent of an elementary-school student, and I would like to know how myopia is diagnosed?" To prevent the effect of interaction between successive inputs, each query was run on a fresh ChatGPT session. Moreover, each prompt was regenerated three times, with each answer rated separately to assess the variation in response quality.

Quality evaluation based on expert reviews

We invited ophthalmologists who met the following criteria: (1) having more than eight years of experience in

Table 1. Appropriateness, usability, and clarity of ChatGPT responses to FAQs related to myopia.

	Question	Appropriateness			Usability			Clarity			All
		≥4	Median	IQR	≥4	Median	IQR	≥4	Median	IQR	≥4
1	Does a visual acuity of 1.0 indicate the absence of myopia?	10	5	5–5	10	5	5–5	10	5	5–5	10
2	Is cycloplegic refraction necessary during an eye examination?	10	4	4–5	10	5	4–5	9	4	4–5	10
3	Can short-acting mydriatic agents control the progression of myopia?	2	3	2–3	2	3	2–3	1	2	1–3	1
4	When does my child need to wear glasses for myopia?	3	3	3–4	3	4	2–4	3	3	2–4	2
5	Do atropine agents cause cataracts?	1	3	1–3	1	2	1–3	0	1	1–2	0
6	What is the recommended age for using atropine eye drops for the treatment of myopia?	9	5	4–5	7	4	3–5	9	4	4–5	6
7	What are the evidence-based methods for treating myopia in children effectively?	7	4	3–4	5	4	3–5	4	4	3–4	4
8	Is Orthokeratology (Ortho-k) corneal reshaping lens safe?	9	5	4–5	9	5	4–5	9	5	4–5	9
9	Can laser treatment be used to treat myopia?	3	4	3–4	3	4	2–4	1	3	2–3	2
10	Can wearing glasses control myopia?	6	4	4–5	5	5	4–5	5	4	4–4	4
11	What precautions should I take when using atropine eye drops?	9	5	4–5	6	5	3–5	10	5	5–5	6
12	Will discontinuing atropine eye drops cause a rebound in myopia progression?	6	4	4–5	9	5	4–5	7	5	3–5	6
13	Is blue light filtering eyewear effective in controlling myopia?	7	5	4–5	8	5	4–5	6	5	4–5	7
14	Is it necessary to use atropine eye drops daily?	10	5	5–5	10	5	4–5	10	5	5–5	10
15	Does receiving orthokeratology eliminate the need for atropine eye drops for treatment?	7	4	5–5	7	5	4–5	7	4	4–5	6
16	Can vision training treat childhood myopia?	9	5	5–5	9	5	5–5	9	5	4–5	7
17	What data should I look for to determine if myopia is not worsening?	7	4	4–5	8	5	4–5	7	4	3–5	8
18	What are the effective methods for preventing myopia?	10	5	5–5	10	5	4–5	10	5	4–5	10
19	Is it necessary to use low-dose atropine eye drops for the prevention of myopia?	9	4	4–5	7	5	3–5	6	4	4–5	9
20	Does not wearing glasses worsen the degree of myopia?	2	3	3–4	3	3	2–4	3	3	2–4	3

Note "all" refers to the number of experts who rated appropriateness, usability, and clarity all above 4.

Table 2. Performance and comparison of ChatGPT responses across categories.

Categories	No.	CRIE scores	CRIE Mean (SD)	Appropriateness ANE (SD)	Usability ANE (SD)	Clarity ANE (SD)
1. Diagnosis	1	9.78	9.71 (0.10)	10.00 (0.00)	10.00 (0.00)	9.50 (0.70)
	2	9.63				
2. Atropine	5	9.21	8.76 (0.87)	6.67 (3.314)	7.33 (3.38)	7.00 (3.79)
	6	9.25				
	11	8.05				
	12	9.88				
	14	7.50				
	19	11.67				
3. Spectacles	4	7.29	7.34 (0.47)	3.67 (1.15)	3.67 (2.08)	3.37 (1.15)
	10	7.36				
	20	7.38				
4. Orthokeratology	8	8.00	8.13 (0.18)	3.50 (2.12)	4.50 (3.53)	3.55 (2.12)
	15	8.26				
5. Effective treatment	3	9.38	9.46 (1.08)	8.00 (1.41)	8.00 (1.14)	8.00 (1.41)
	7	9.53				
6. Others	9	11.46	10.76 (1.06)	7.60(2.70)	7.20(2.68)	6.60(3.50)
	13	9.29				
	16	10.67				
	17	12.07				
	18	10.31				
F(p)	–	–	5.71(0.004)	2.462(0.080)	1.69(0.200)	1.67(0.200)

Note no: the item number of the FAQs; ANE: the average number of experts who rated the questions above 4; SD: standard deviation; CRIE: Chinese Readability Index Explorer; F(p) indicates the ANOVA results for CRIE scores and expert review scores.

pediatric myopia; (2) dedicating at least one pediatric myopia outpatient clinic per week; and (3) having published case reports or studies related to pediatric myopia. Ten experts participated and rated the appropriateness, usability, and clarity of each response using a five-point Likert-type scale from 1 (very inappropriate/very difficult to use/extremely unclear) to 5 (very appropriate/usable/clear), as previously described.[17] The inter-rater agreement between the 10 ophthalmologists was excellent with a Kappa value of 0.820 (95% confidence interval, 0.721–0.920).

We calculated the median, interquartile range (IQR), and range of scores for each question. For scores less than 4, ophthalmologists provided explanations for the answers deemed inappropriate. Only responses that obtained scores of ≥4 on all three indicators (appropriateness,

usability, and clarity) by at least eight ophthalmologists were considered satisfactory.

Analysis of readability using the Chinese readability index explorer

The readability of each response was assessed using the Chinese Readability Index Explorer (CRIE, version 3.0, available at http://www.chinesereadability.net/CRIE), an online automated tool for analyzing Chinese text.[18] The CRIE integrates four subsystems and 82 multilevel linguistic features to analyze word count, lexical richness, semantic complexity, syntactic diversity, and cohesion. Recent studies have confirmed the reliability of the CRIE in assessing the readability of texts written in Chinese.[19,20] In this study, we entered 20 responses into the CRIE system to create readability prediction models and assign scores to categorize text into the corresponding grade reading levels: levels 1–6, elementary school (grades 1–6); levels 7–9, middle school (grades 7–9); and levels 10–12, high school (grades 10–12).

Statistical analysis

We conducted a descriptive analysis of the scores by 10 ophthalmologists, including the median and the IQR. Cohen's kappa coefficient was used to measure inter-rater reliability. One-way analysis of variance (ANOVA) was used to test for differences in the readability and expert ratings of the categorized responses. Spearman's rank correlation analysis was performed for appropriateness, usability, clarity, and CRIE scores, with the significance level set at $p < 0.05$.

Results

As shown in Table 1, the median appropriateness, usability, and clarity scores of the responses to the 20 FAQs ranged from 3 to 5, 2 to 5, and 1 to 5, respectively. Five of the responses (25%) met the satisfactory response criteria as rated by more than eight ophthalmologists. These satisfactory responses covered topics such as "significance of visual acuity (#1)," "necessity of cycloplegic refraction to diagnose myopia (#2)," "corneal reshaping mechanisms of orthokeratology (#8)," "indicators of myopia progression (#17)," and "effective methods for preventing myopia (#18)." For example, the highly rated response for topic #8 was "Orthokeratology (Ortho-K) corneal reshaping lenses are a type of specialized contact lenses that can correct myopia by reshaping the cornea. Compared to traditional contact lenses, Ortho-K lenses have several advantages, such as effectively controlling the progression of myopia and eliminating the need for daytime glasses or contact lenses."

Responses related to treatment efficacy (#3, #4, #9, and #20) and side effects (#5) were rated low for all three criteria (no more than three ophthalmologists provided scores of >4) and were considered inaccurate or flawed by the ophthalmologists. These responses included suggestions of short-term efficacy for myopia treatments, wearing spectacles, recommendations for prolonged use despite pupil dilation, and endorsement of laser for myopia. For example, a poorly rated response was for topic #9: "Laser treatment can be used to treat myopia, primarily through two methods: LASIK and LASEK/PRK. Both procedures require a longer recovery period and care. Regardless of the method used, laser treatment for myopia should be performed by experienced professionals in appropriate medical facilities. Detailed evaluation and examination are necessary to determine the suitability and safety of the treatment plan."

Table 2 presents the mean scores of items and categories for CRE, along with the average number of experts who rated the questions above 4. The CRIE readability scores for the ChatGPT responses ranged from 7.29 to 12.07. The readability of 35% of the responses (to seven questions) was at the middle-school level, while that of the remaining questions was above the high-school level (>9). Responses related to laser surgery (#9), low-concentration atropine for myopia prevention (#19), and determination of myopia progress (#17) had high readability scores (>11). When categorized by the nature of the questions, responses about spectacles and orthokeratology had relatively lower readability scores (mean = 7.34 and 8.13, respectively), while those discussing effective treatments and diagnoses had higher readability scores (mean = 9.46 and 9.71, respectively).

As shown in Table 2 and Figure 1, the results of one-way ANOVA showed significant differences in the readability of the ChatGPT responses to different categories of FAQs (p = 0.004), whereas no significant differences were observed among the expert evaluations on all three criteria. Table 3 shows the correlations between the CRIE readability scores and expert review scores. Spearman's rank correlation analysis showed no significant correlations between the expert ratings of appropriateness, usability, and clarity of the responses and the CRIE readability scores.

Discussion

This study is the first to demonstrate that a novel LLM-derived conversational AI program can provide appropriate, usable, and clear responses in traditional Chinese language to myopia-related FAQs, as determined by ophthalmologists and the CRIE. ChatGPT is an AI model based on deep learning and natural language processing, and it is trained on vast amounts of text to generate human-like text responses by predicting the most probable
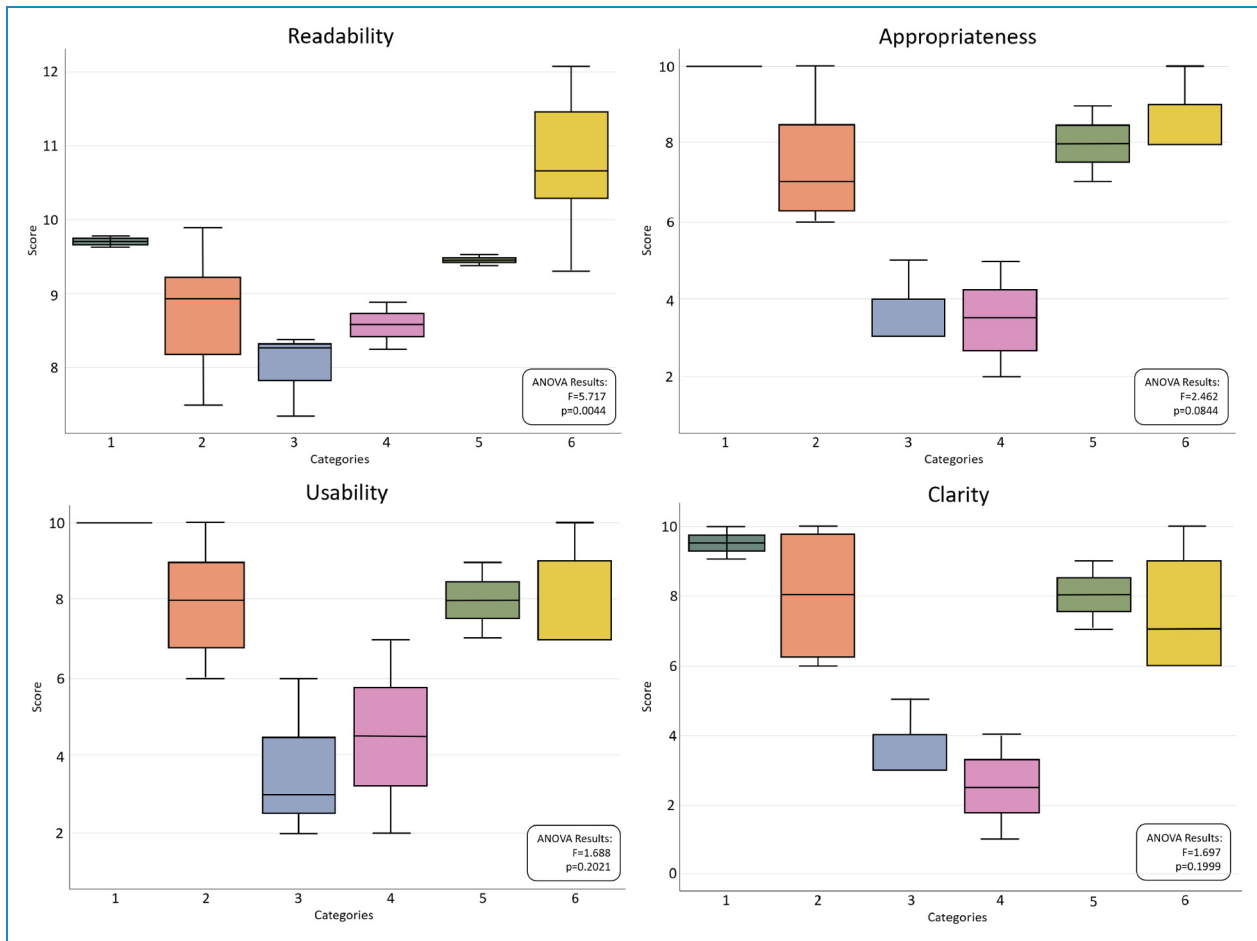
Figure 1. One-way of analysis of variance of different treatment categories for myopia management.
Note: Category: 1. Diagnosis, 2. Atropine, 3. Spectacles, 4. Orthokeratology, 5. Effective treatment, 6. Others.

next word or sequence of words when given a prompt or question.[17]

This study found that less than half of the ChatGPT myopia-related responses in the Chinese language met the criteria for appropriateness, usability, and clarity, as evaluated by ophthalmologists. Several studies have highlighted the suboptimal clarity of online or printed materials related to ophthalmic diseases, which impairs patient comprehension.[21,22] The reports on the quality of ChatGPT responses to FAQs by previous studies have been inconsistent.[14,23] ChatGPT provided high-quality responses to questions from the general public, as evaluated by physicians,[3] while its performance in addressing ophthalmologic treatment questions was low.[22,24,25] This is consistent with the findings of the present study.

Nikdel et al.[14] reported that more than 80% of the responses of ChatGPT 4.0 to questions about myopia were deemed appropriate by two ophthalmologists; however, the percentage of appropriate responses in this study was lower. In addition to using different versions of ChatGPT, this difference may be attributed to the number

of ophthalmologists invited to evaluate the responses and the language used. In the study of Biswas et al.,[6] the response of ChatGPT to the question of whether a single treatment could effectively control myopia was assigned high accuracy scores by experts. Similar to Lim and Pushpanathan,[20] we asked ChatGPT about each specific myopia treatment, and the responses to each question in our study received lower appropriateness ratings. The use of atropine eye drops is more common in Asian countries with high prevalence of myopia, such as Hong Kong, Taiwan, and mainland China, whereas it is less commonly studied in the United States.[26] The quality of the responses of ChatGPT about the use of atropine was less satisfactory in our study. This indicates that the primary training texts for ChatGPT are in English,[27] and its responses tend to align with the American socio-cultural environment. This may limit its adaptability to other cultural contexts, especially in understanding and addressing treatments more commonly used in Asia.[28]

The readability of health information is assessed based on factors such as sentence length and average number of

Table 3. Correlations between CRIE, appropriateness, usability, clarity.

Variables	CRIE	Appropriateness	Usability	Clarity
CRIE	1	0.184	0.201	0.244
Appropriateness		1	0.711**	0.854**
Usability			1	0.897**
Clarity				1

Note ** $p < 0.01$.

syllables per word to determine the reading level required to understand the text effectively.[29] Based on the CRIE readability scores, the responses related to spectacles and orthokeratology were rated at middle-school level. This is consistent with the findings of King et al.[30] who demonstrated that the readability of ChatGPT responses exceeds the American Medical Association's recommended reading level of fifth to sixth grade. Consistent with the findings of Tao and Hua (2022),[31] the readability level of ophthalmology-related ChatGPT responses also exceeded that of eighth grade. The expert ratings for spectacle-related responses were low, whereas responses concerning orthokeratology were rated more favorably in terms of expertise. The readability of the responses of ChatGPT is influenced by the quality of its training sources[28] and the prompts provided, which should be tailored to the reader's literacy level.[32] Given that the training data of ChatGPT originate from existing sources, it is plausible that materials regarding spectacles were not composed by ophthalmologists. This may contribute to perceptions of insufficient professional depth from an ophthalmological viewpoint. Additionally, ophthalmologists may maintain conservative views on the role of spectacles in treating myopia, potentially leading to disagreements with responses linking spectacles to myopia treatment.[33] Orthokeratology is supported by extensive evidence-based studies,[34] practical guidelines,[35] and comprehensive health education materials,[36,37] providing ChatGPT with robust textual resources for more accurate responses.

The responses of ChatGPT to inquiries related to the diagnosis and effective treatment of myopia have a high readability level, often necessitating a reading proficiency beyond that of the middle school. This high level of readability highlights the need to pay more attention to the lower level of knowledge displayed by parents concerned with these topics.[10] The study identified a lack of significant correlation between readability and the quality of information provided by ChatGPT, corroborating earlier findings that the quality of health information does not necessarily correspond with its readability.[38]

The performance of ChatGPT in addressing mydriatic treatments for myopia was inadequate and potentially misleading. The appropriateness was significantly lower, warranting further attention. This resonates with the notion that ChatGPT does not possess genuine comprehension or knowledge of the text it generates, and that a lack of deeper understanding of the context or meaning behind the words it uses could lead to misleading responses.

This study has some limitations. First, the 20 FAQs were obtained from websites, rather than directly gathering queries from parents during clinical visits. This may have limited the representation of their genuine concerns. Second, interaction with ChatGPT was confined to a single prompt, potentially constraining the precision of the responses obtained. Future research should employ diverse prompts and other versions of ChatGPT or Gen AI machine to verify responses. Third, this study evaluated the quality of the responses of ChatGPT based solely on ophthalmologists' opinions. Future analyses should include other medical resources and revise any inaccuracies in the responses of ChatGPT based on expert feedback to provide more accurate and useful information.

Eye care practitioners may not be easily accessible, and their services may be costly and time-consuming. Conversely, online chatbots have gained widespread popularity and attention among parents. AI is viewed as a promising tool, offering opportunities to explore its utility in patient interactions and providing valuable information on eye diseases. In the future, families, schools, and pediatric nurses should collaborate to assist parents in evaluating information related to the prevention and treatment of myopia using ChatGPT.

Given the significance of online information as a pivotal channel for the parents of children with myopia in guiding treatment decisions,[10,39] it is essential for ECPs to comprehend the role and functionality of ChatGPT in health education for myopia control. Based on the findings of this study, several recommendations can be made. First, to better utilize ChatGPT in healthcare services, healthcare professionals need to receive training on effective prompting techniques to engage with ChatGPT across various topics, focusing on formulating appropriate questions to obtain more reliable information through multi-step interactions.

Additionally, professionals should be aware of the language and cultural differences in ChatGPT's training data to avoid potential biases or misleading information. This study utilized ChatGPT-3.5 with one-shot prompting to generate responses. Future studies should compare the appropriateness, usability, and clarity of responses from different ChatGPT versions, other generative AI models, and prompting techniques to optimize the use of ChatGPT as a tool for supporting healthcare professionals.

Conclusions

In summary, this study evaluated the performance of ChatGPT in responding to the most commonly asked myopia-related questions by parents online. ChatGPT performed better in addressing general myopia-related questions than in responding to treatment-related questions, particularly those concerning the use of atropine eye drops and spectacles for children. The responses of ChatGPT in this study demonstrated high appropriateness, usability, and clarity regarding atropine treatment, making ChatGPT a valuable supplementary tool for patient and health education. Importantly, education materials should remind parents to ask additional questions to verify the accuracy of ChatGPT's responses concerning eyeglasses. However, the readability of the responses provided by ChatGPT to myopia-related questions still exceeded the American Medical Association's standard of being below the sixth-grade level. Healthcare professionals need to be trained in the use of generative artificial intelligence to ensure they can appropriately identify the accuracy of myopia-related health information generated by ChatGPT and align it with the comprehension levels of individual patients.

ORCID iD: Li-Ling Liao 🆔 https://orcid.org/0000-0003-0867-2699

References

1. Abramson A. How to use ChatGPT as a learning tool. *Monit Psychol* 2023; 54: 67. Retrieved May 1, 2024 from https://www.apa.org/monitor/2023/06/chatgpt-learning-tool.
2. Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. *Healthcare* 2023; 11: 887.
3. Ayers JW, Poliak A, Dredze M, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA Intern Med* 2023; 183: 589–596.
4. Yılmaz İE and Doğan L. Talking technology: exploring chatbots as a tool for cataract patient education. *Clin Exp Optom* 2024; 9: 1–9.
5. Samaan JS, Yeo YH, Rajeev N, et al. Assessing the accuracy of responses by the language model ChatGPT to questions regarding bariatric surgery. *Obes Surg* 2023; 33: 1790–1796.
6. Biswas S, Logan NS, Davies LN, et al. Assessing the utility of ChatGPT as an artificial intelligence-based large language model for information to answer questions on myopia. *Ophthalmic Physiol Opt* 2023; 43: 1562–1570.
7. Holden BA, Fricke TR, Wilson DA, et al. Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. *Ophthalmol* 2016; 123: 1036–1042.
8. Ding X, Morgan I, Hu Y, et al. The causal effect of education on myopia: evidence that more exposure to schooling, rather than increased age, causes the onset of myopia. *Invest Ophthalmol Vis Sci* 2023; 64: 25.
9. Chang LC, Li FJ, Sun CC, et al. Trajectories of myopia control and orthokeratology compliance among parents with myopic children. *Cont Lens Anterior Eye* 2021; 44: 101360.
10. Hung LL, Liao LL, Chen HJ, et al. Factors associated with follow-up visits in parents with myopic children wearing orthokeratology lens. *J Nurs Res* 2022; 30: e244.
11. Shao CY, Li H, Liu XL, et al. Appropriateness and comprehensiveness of using ChatGPT for perioperative patient education in thoracic surgery in different language contexts: survey study. *Interact J Med Res* 2023; 12: e46900.
12. King RC, Samaan JS, Yeo YH, et al. Appropriateness of ChatGPT in answering heart failure related questions. *Heart Lung Circ* 2024: S1443-9506(24)00165-3. doi:10.1016/j.hlc.2024.03.005.
13. Miller S, Gilbert S, Virani V, et al. Patients' utilization and perception of an artificial intelligence-based symptom assessment and advice technology in a British primary care waiting

room: exploratory pilot study. *JMIR Hum Factors* 2020; 7: e19713.

14. Nikdel M, Ghadimi H, Tavakoli M, et al. Assessment of the responses of the artificial intelligence-based chatbot ChatGPT-4 to frequently asked questions about amblyopia and childhood myopia. *J Pediatr Ophthalmol Strabismus* 2024; 61: 86–89.

15. Polat E, Polat YB, Senturk E, et al. Evaluating the accuracy and readability of ChatGPT in providing parental guidance for adenoidectomy, tonsillectomy, and ventilation tube insertion surgery. *Int J Pediatr Otorhinolaryngol* 2024; 181: 111998. 2024/06/03.

16. Haver HL, Ambinder EB, Bahl M, et al. Appropriateness of breast cancer prevention and screening recommendations provided by ChatGPT. *Radiology* 2023; 307: e230424.

17. Lee TC, Staller K, Botoman V, et al. ChatGPT answers common patient questions about colonoscopy. *Gastroenterol* 2023; 165: 509–511.e7.

18. Sung Y-T, Chang T-H, Lin W-C, et al. CRIE: an automated analyzer for Chinese texts. *Behav Res Methods* 2016; 48: 1238–1251.

19. Tseng H-C, Chen B, Chang T-H, et al. Integrating LSA-based hierarchical conceptual space and machine learning methods for leveling the readability of domain-specific texts. *Nat Lang Eng* 2019; 25: 331–361.

20. Sung Y-T, Chen J-L, Cha J-H, et al. Constructing and validating readability models: the method of integrating multilevel linguistic features with machine learning. *Behav Res Methods* 2015; 47: 340–354.

21. John AM, John ES, Hansberry DR, et al. Analysis of online patient education materials in pediatric ophthalmology. *J AAPOS* 2015; 19: 430–434.

22. Patel AJ, Kloosterboer A, Yannuzzi NA, et al. Evaluation of the content, quality, and readability of patient accessible online resources regarding cataracts. *Semin Ophthalmol* 2021; 36: 384–391.

23. Lim ZW, Pushpanathan K, Yew SME, et al. Benchmarking large language models' performances for myopia care: a comparative analysis of ChatGPT-3.5, ChatGPT-4.0, and Google bard. *EBioMedicine* 2023; 95: 104770.

24. Potapenko I, Boberg-Ans LC, Stormly Hansen M, et al. Artificial intelligence-based chatbot patient information on common retinal diseases using ChatGPT. *Acta Ophthalmol* 2023; 101: 829–831.

25. 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 Spectr* 2023; 7: pkad010.

26. Wei XL, Wu T, Dang KR, et al. Efficacy and safety of atropine at different concentrations in prevention of myopia progression in Asian children: a systematic review and meta-analysis of randomized clinical trials. *Int J Ophthalmol* 2023; 16: 1326–1336.

27. Ting Y-T, Hsieh T-C, Wang Y-F, et al. Performance of ChatGPT incorporated chain-of-thought method in bilingual nuclear medicine physician board examinations. *Digit Health* 2024; 10: 20552076231224074.

28. Cao Y, Zhou L, Lee S, et al. Assessing cross-cultural alignment between ChatGPT and human societies: an empirical study. *arXiv* 2023; 2303: 17466.

29. Thia I and Saluja M. ChatGPT: is this patient education tool for urological malignancies readable for the general population? *Res Rep Urol* 2024; 16: 31–37.

30. King RC, Samaan JS, Yeo YH, et al. A multidisciplinary assessment of ChatGPT's knowledge of amyloidosis: observational study. *JMIR Cardiogr* 2024; 8: e53421.

31. Tao BK, Hua N, Milkovich J, et al. GPT-3.5 and Bing Chat in ophthalmology: an updated evaluation of performance, readability, and informative sources. *Eye* 2024; 38: 1897–1902.

32. Valencia OA G, Thongprayoon C, Miao J, et al. Empowering inclusivity: improving readability of living kidney donation information with ChatGPT. *Front Digit Health* 2024; 6: 1366967.

33. Bist J, Kaphle D, Marasini S, et al. Spectacle non-tolerance in clinical practice – A systematic review with meta-analysis. *Ophthalmic Physiol Opt* 2021; 41: 610–622.

34. Tang K, Si J, Wang X, et al. Orthokeratology for slowing myopia progression in children: a systematic review and meta-analysis of randomized controlled trials. *Eye Contact Lens* 2023; 49: 404–410.

35. Cho P, Cheung SW, Mountford J, et al. Good clinical practice in orthokeratology. *Cont Lens Anterior Eye* 2008; 31: 17–28.

36. Chan B, Cho P and Cheung SW. Orthokeratology practice in children in a university clinic in Hong Kong. *Clin Exp Optom* 2008; 91: 453–460.

37. Sun CC, Liao GY, Liao LL, et al. A cooperative management app for parents with myopic children wearing orthokeratology lenses: mixed methods pilot study. *Int J Environ Res Public Health* 2021; 18: 10316.

38. Li Y, Zhou X, Zhou Y, et al. Evaluation of the quality and readability of online information about breast cancer in China. *Patient Educ Couns* 2021; 104: 858–864.

39. Cheung S-W, Lam C and Cho P. Parents' knowledge and perspective of optical methods for myopia control in children. *Optom Vis Sci* 2014; 91: 634–641.