AJO International 1 (2024) 100032

Contents lists available at ScienceDirect

AJO International

journal homepage: www.sciencedirect.com/journal/ajo-international

Pilot study evaluating the usability of MonŒil, a ChatGPT-based education tool in ophthalmology

Claudio XOMPERO, Walid BENETTAYEB, Eric H. SOUIED, Carl-Joe MEHANNA [*]

Department of Ophthalmology, Centre Hospitalier Intercommunal de Créteil, Créteil, France

ARTICLE INFO

Keywords:
Age-related macular degeneration
Ophthalmology
Patient education
ChatGPT
Artificial intelligence
chatbot

ABSTRACT

Importance: There is an increasing use of artificial intelligence (AI) in ophthalmology to respond to the needs of patients to access reliable, easy-to-understand medical information.

Objective: To assess patient satisfaction with the usability of MonŒil, an AI-based platform designed for patient education in ophthalmology.

Design: This was a pilot cross-sectional study of the usability of MonŒil by patients followed for advanced age-related macular degeneration (AMD). MonŒil is based on ChatGPT-4 technology with specific ophthalmology-focused customizations and accessibility enhancements, and is freely available at *monoeil.help*. Patients were given 20 min of unsupervised interaction with no prior training or guidance, after which their feedback was collected.

Setting: The study was performed in the ophthalmology department at the Creteil University Hospital.

Participants: Participants included 54 patients diagnosed with advanced AMD defined presenting for follow-up. Patients had to be older than 50 years of age, capable of giving informed consent, and able to understand and interact with MonŒil. Exclusion criteria were severe visual and cognitive impairment that prevented interaction with MonŒil.

Main Outcome(s) and Measure(s): The primary outcome was the usability of MonŒil as measured by the System Usability Scale (SUS) questionnaire.

Results: Of the 54 participants, 34 were female (62.96 %). The mean age of the cohort was 77.76±8.14 years (range 58 to 97 years). The mean SUS score was 90.23±12.04 with a median of 92.50 (range 42.50 to 100.00), indicating excellent usability. There was a positive relationship between visual acuity and SUS score (regression coefficient 0.30 (95 % CI 0.08 to 0.51), r2=0.19, $p = 0.0077$).

Conclusions and Relevance: MonŒil demonstrated excellent usability and satisfaction in a sample population of elderly patients with advanced AMD. These results suggest that AI-based tools like MonŒil can enhance patient education with minimal oversight in a complex field like ophthalmology, supporting its use as an adjunct to the physician-patient discussion. Further studies may be necessary to establish the applicability of MonŒil to a broader user base, and to assess its usefulness and clinical impact on patient outcomes such as patient knowledge and vision-related quality of life.

Introduction

The integration of artificial intelligence (AI) in medicine has grown exponentially over the last few years.[1] In ophthalmology in particular, the applications of AI have been mainly geared towards improving our current technologies' diagnostic accuracy.[2–6] Recently, with the advent and improvement of large language models (LLM) such as ChatGPT (OpenAI, San Francisco, CA, USA), AI's ability to answer standardized questions and aid in patient education has been evaluated. In particular, GPT-4, the latest and most advanced LLM from OpenAI,[7] was proven to be effective in answering patient inquiries in ophthalmology.[8–11]

In ophthalmology, several diseases present unique challenges in patient communication and information dissemination. In particular, age-related macular degeneration (AMD), a leading cause of vision impairment among the elderly,[12] necessitates not only timely and effective treatment but also comprehensive patient understanding of the condition and its management strategies.[13] Research indicates that many patients struggle to completely grasp or remember this

* Corresponding author.
E-mail address: carljoe.mehanna@gmail.com (C.-J. MEHANNA).

https://doi.org/10.1016/j.ajoint.2024.100032
Received 3 February 2024; Received in revised form 18 May 2024; Accepted 18 May 2024
Available online 19 May 2024

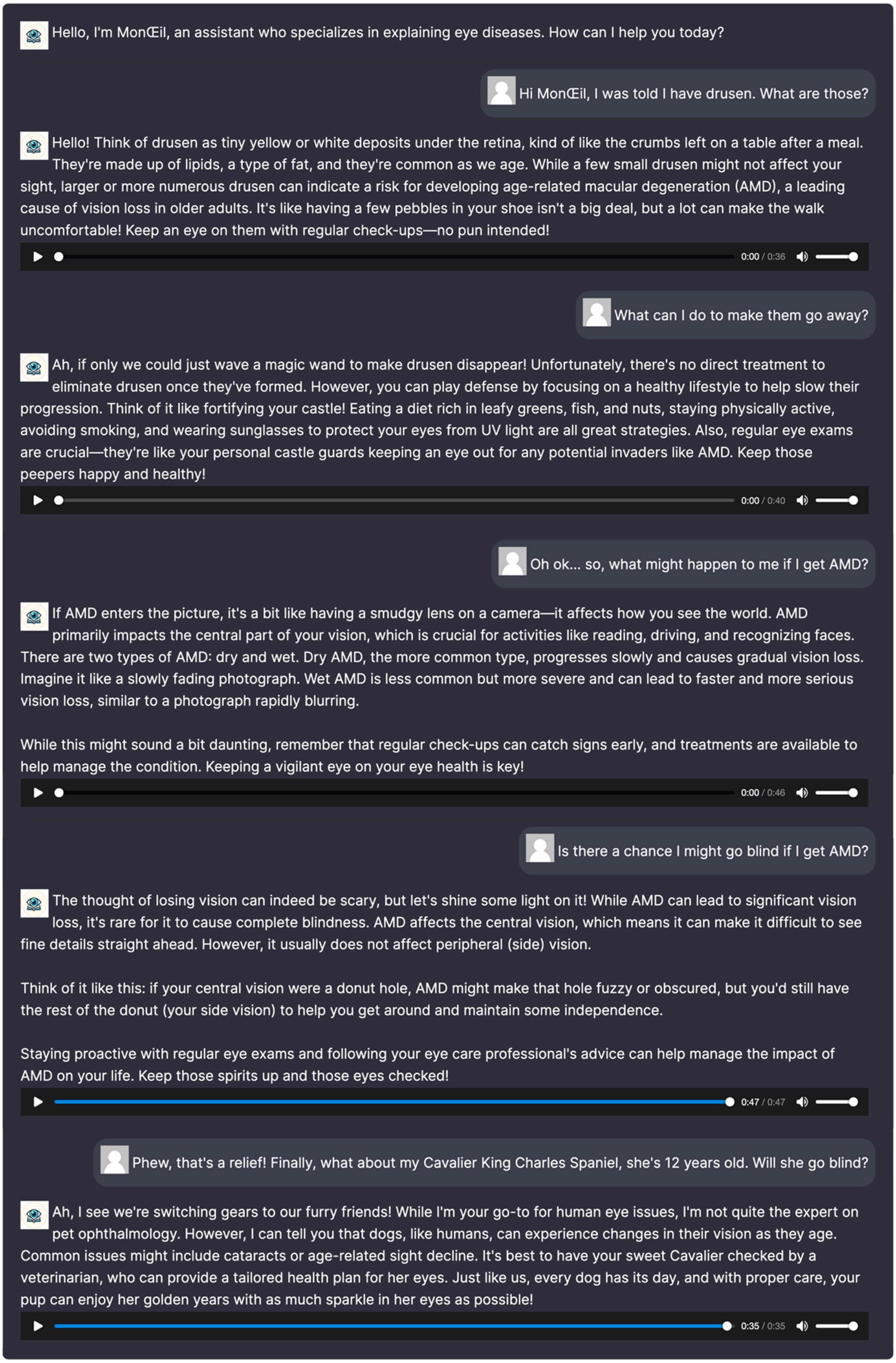

Fig. 1. Example of a series of interactions with MonŒil (in English).

information, and this gap in understanding and memory can lead to decreased patient satisfaction and compliance, as well as worse treatment outcomes.[14,15] Patients often turn to the internet to find details about their medical condition, its symptoms, and potential treatments.[16,17] However, the average online ophthalmic resource only fulfills about 26 % of the standards for high-quality informative content in AMD.[18] AI-based LLM were shown to offer accurate and satisfactory responses addressing AMD patients' questions.[19] Nevertheless, not all AI-generated patient education material (PEM) are equivalent, but with appropriate prompts, ChatGPT-generated PEM can outperform other chatbots and even brochures from major ophthalmological societies.[20] There is thus a need for educational tools in ophthalmology that consistently provide accessible, high-quality, and understandable information for patients beyond PEM such as leaflets. These tools would serve to complement and empower the physician-patient discussion outside the consultation.

We therefore developed "MonŒil", a ChatGPT-4-based LLM focused on ophthalmology-related material. MonŒil is a tool that aims at improving patient education and patient access to reliable, concise, and accurate information for free. It offers a unique user experience by simulating an educational conversation that is friendly and easily understandable, while proposing enhanced accessibility features. In this pilot study, we aim to evaluate the usability of MonŒil in a group of patients with advanced AMD.

Materials and methods

This cross-sectional study was conducted to evaluate the usability of MonŒil by patients with advanced AMD who were using it for the first time. The study was conducted in accordance with the tenets of the Declaration of Helsinki and with the approval from the Université Paris-Est Creteil institutional review board. Informed consent was provided by all participants.

Participants

Between October 2023 and December 2023, patients and their accompanying relatives attending the ophthalmology department at the Creteil University Hospital, in Creteil, France, were approached. The hospital is a tertiary care center with a high load of AMD patients attending specialized clinics.

Inclusion criteria were: 1) age 50 years or older; 2) have a proven diagnosis of advanced AMD complicated by either macular neovascularization or geographic atrophy; 3) be presenting for a follow-up visit or a planned intravitreal injection; and 4) be capable of giving informed consent.

Exclusion criteria were 1) severe visual impairment preventing interaction with MonŒil, defined as visual acuity <20/200 in the better seeing eye; 2) insufficient cognitive ability to understand and interact with *MonŒil* defined as any diagnosis of dementia or cognitive disorder (e.g., stroke with residual cognitive deficit, Parkinson disease, etc.) reported by the patient, their accompanying relative, or documented in the medical chart; and 3) co-existing significant ocular pathology (e.g., advanced glaucoma, corneal scars or ectasia, uveitis, etc.).

MonŒil system

MonŒil is a customized ChatGPT model fine-tuned for ophthalmology-related queries. MonŒil builds on GPT-4, the latest and most advanced LLM from OpenAI,[7] that was validated in answering patient inquiries in ophthalmology.[8–11] GPT-4 is available through paid subscription plans and can be further modified to standardly follow a set of prompts and rules.

Specifically, we have customized *MonŒil* according to the following rules: 1) limit the scope of responses to ophthalmology-related material, and decline answering non-ophthalmological queries; 2) do not suggest a specific diagnosis based on symptoms given; however, it is possible to suggest a differential diagnosis; 3) do not suggest a specific medical treatment based on a diagnosis or a set of symptoms; however, it is possible to discuss broad therapeutic guidelines; 4) limit the answers to less than 200 words, equivalent to about one minute of reading or speech by the text-to-speech software; 5) answer in a simple, friendly, and slightly humorous tone, "as if one is talking to his close friend who is an ophthalmologist"; 6) always end by recommending a consultation with an ophthalmologist for further information or stating that *MonŒil* is not a substitute to it. While *MonŒil* will not give specific diagnoses or treatments, it will unrestrictedly answer questions about a particular diagnosis or treatment the patient might have received from their physician. This approach ensures that *MonŒil* serves as an informational and educational tool, complementing the expertise of ophthalmologists without attempting to supplant it. Fig. 1 illustrates an example of an interaction with MonŒil.

The chatbot was then integrated into a website, and made freely available for use at *https://monoeil.help/*. The website does not ask for any personal information nor does it require signing up. The sessions are stored on the user's device for future consultation, but can be deleted by clearing the cache or using private browsing. No data other than website traffic is collected from the interactions, ensuring privacy and security. Additionally, it is possible to use dictation to ask the questions, and we have added a built-in text-to-speech generator that automatically speaks out the answers. However, these two options were disabled to offer a standard experience using the core version of the website. Lastly, the website, the chatbot, and the speech output are all available in two languages: French and English. In our study, all of the patients spoke French and hence used the French version.

Procedure

Initially, patients with advanced AMD presenting to their consultation alone or with their accompanying relative were introduced to MonŒil by their treating physician. This initial step helped to explain the purpose of the application and the background that motivated its development. After this introduction, participants were given the opportunity to use the application independently and for the first time. This session lasted 20 min and was conducted in a separate room without the presence or supervision of the investigators in order to mimic a real-world scenario. There were no rules or restrictions regarding the nature or number of questions asked. In fact, patients were encouraged to ask the questions in their own words and across all domains (e.g., diagnosis, prognosis, information on their treatments, dietary recommendations, activities of daily living, etc.). Following the independent interaction, each patient was requested to complete the System Usability Scale (SUS) questionnaire to record their feedback on MonŒil's usability.

Before concluding the session, the retina specialists (ES and CJM) reviewed MonŒil's responses for gross inaccuracies, of which there were none. While this study does not aim to assess the accuracy of GPT-4′s responses, as this had already been established in other studies,[21–23] this review process ensured the reliability of the information provided by the AI system. It also allowed the ophthalmologists to gather feedback and address any residual doubts or questions the patients might have had, ensuring clarity and reinforcing the educational value of MonŒil.

Data collection

Data collected included patient age, gender, years since first presentation, and best corrected visual acuity (BCVA). In our practice, BCVA is routinely taken using standard ETDRS charts, and was converted to letters for the purpose of analysis. In addition, patients were requested to answer the 10 questions of the System Usability Scale questionnaire using a 5-point Likert scale.

Table 1
Patient characteristics.

$N = 54$	Mean (SD)	Median (Min; Max)
Age [years]	77.76 (8.14)	76 (58; 97)
BCVA OD [ETDRS letters]	65.93 (18.33)	72.50 (25; 85)
BCVA OS [ETDRS letters]	65.09 (21.00)	77.50 (20; 85)
BCVA better eye [ETDRS letters]	73.24 (14.64)	80.00 (35; 85)
Duration of follow-up [years]	6.00 (4.36)	5 (1; 14)
SUS score	90.23 (12.04)	92.50 (42.50; 100.00)

BCVA = best corrected visual acuity; OD = right eye; OS = left eye; SUS = system usability scale.

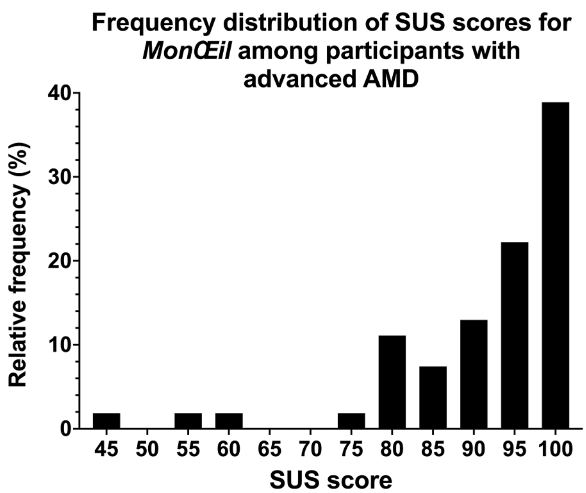

Fig. 2. Distribution of SUS scores for the usability of MonŒil in the sample of participants with advanced AMD. In this histogram, SUS scores were rounded up to the nearest multiple of 5.

System usability scale (SUS)

The SUS questionnaire is a quick, simple, validated 10-item questionnaire with five response options based on Likert scale. It gives a subjective assessment of the usability of new computer systems or websites.[24,25]

To analyze SUS data effectively, each item is scored and then aggregated to form the overall SUS score. This scoring involves a specific method: for the odd-numbered questions, we subtract 1 from the user's response. Conversely, for the even-numbered questions, we subtract the user's response from 5. After tallying all these scores, we multiply the total by 2.5, which effectively converts it into a scale ranging from 0 to 100. It is established that a SUS score above 68 is regarded as above average, whereas a score surpassing 80 is deemed excellent.[26,27]

Statistical analysis

Data were summarized using descriptive statistics and frequency distributions. Multiple linear regression analyses were performed between SUS score and age, gender, BCVA of the better seeing eye, and years of follow-up. All statistical analyses were performed using GraphPad Prism version 10.1.1 (GraphPad Software, San Diego, California USA, www.graphpad.com). P values inferior to 0.05 were considered statistically significant.

Results

This study involved 54 participants with a diagnosis of advanced AMD in at least one eye of whom 34 were female (62.96 %). The mean age of the cohort was 77.76±8.14 years (range 58 to 97 years), with an average of 6.00±4.36 years of follow-up. BCVA (in ETDRS letters) was similar in both eyes at 65.93±18.33 and 65.09±21.00 for the right and left eyes, respectively ($p = 0.99$), which corresponds to a Snellen equivalent of around 20/50 (Table 1).

The analysis of System Usability Scale scores from the questionnaire revealed a mean SUS score of 90.23±12.04 with a median of 92.50 (range 42.50 to 100.00). Out of the 54 participants, 47 (87.04 %) gave a score of 80 or higher, 37 (68.52 %) gave a score of 90 or higher, and 17 participants (31.48 %) gave a score of 100 (Fig. 2). In particular, assessing the need for the system through the first question of the SUS "I think that I would like to use this system frequently", we found patients scored 4.35±0.99, with a median of 5 (range 1 to 5).

During linear regression analysis, age and BCVA of the better seeing eye were significantly associated with SUS scores ($p < 0.05$). There were no significant associations between SUS scores and gender or duration of follow-up. However, on multiple regression analysis, only BCVA of the better seeing eye remained significant with a regression coefficient of 0.30 (95 % CI 0.08 to 0.51), despite a low $r^2 = 0.19$ ($p = 0.0077$).

Discussion

MonŒil, the custom AI-based educational chatbot specialized in ophthalmology, showed excellent usability with mean and median SUS scores above 90. These results were obtained in participants with advanced AMD without cognitive or severe visual impairment. The SUS score is a widely recognized tool for assessing the user-friendliness of systems, and its application in this context is pivotal to evaluate patient interaction with digital health aids.[24,25] MonŒil scores similarly to some of the highest-rated websites such as Google search (SUS score of 93)[28] from which it differs fundamentally: in practice, the patient types or dictates a question, but rather than getting multiple links as in traditional search engines, a single, easy to understand, short response is displayed. This eliminates the additional steps of having to search for the information and the risks of falling down the rabbit hole of self-diagnosis or self-treatment. The positive reception of MonŒil is consistent with a growing body of research emphasizing the central role of artificial intelligence in patient education and engagement.[29–31]

This study was conducted in an academic hospital specialized in the management of AMD that has dedicated AMD consultations as well as injection clinics with trained nurses. In fact, as of late 2022, our institution offers a multi-disciplinary approach to the management of patients with treatment-naïve neovascular AMD: therapeutic patient education workshops, low-vision rehabilitation, and an in-house psychologist for the management of the emotional burden associated with the disease. While none of the patients in our cohort benefited from these additional services, they might still be more prepared and educated than the average AMD patient. Regardless, MonŒil performed excellently, indicating that there are still unmet needs that MonŒil can answer, corroborated by the high score on the first question of the SUS. It can thus be envisaged that MonŒil be used as a complement during patient education workshops or as a standalone tool for patients in underserved areas, for whom there may be even greater benefits.

The ease of use and contextual information provided by MonŒil addresses a critical need in ophthalmology, especially in diseases as complex as AMD, where patient understanding can significantly impact outcomes and adherence to treatment protocols.[14,15] Patients' understanding is not always complete, which may be explained by suboptimal information about AMD diagnosis, treatment, prevention, or prognosis.[32] In addition, patients with AMD present at an advanced age, and the ability to recall information declines with age.[33] The possibility of having simple and accessible education tools outside the consultation is thus very important in this population. This study also highlights the importance of a hybrid approach to healthcare, using AI-based tools as adjuncts to the existing discussion with healthcare professionals. In fact, patients gave positive feedback overall and were pleasantly surprised particularly with the signature way in which MonŒil phrases its answers. In contrast to other applications of AI in medicine, MonŒil

Table 2
Characteristics of patients with the lowest SUS scores.

Participant	Age [years]	BCVA OD [ETDRS letters] (Snellen equivalent)	BCVA OS [ETDRS letters] (Snellen equivalent)	Gender [M/F]	Duration of follow-up [years]	SUS Score
1	88	50 (20/100)	25 (20/320)	F	3	42,5
2	75	35 (20/200)	20 (20/400)	F	12	55
3	82	55 (20/80)	50 (20/100)	M	5	60

BCVA = best corrected visual acuity; OD = right eye; OS = left eye; SUS = system usability scale.

operates exclusively in patient education and does not provide any specific diagnostic or treatment recommendations. It thus complies with current ethical guidelines for the autonomous use of AI in healthcare. To our knowledge, this is the first study assessing the usability of a customized GPT-4-based tool for patient education.

The visual acuity of our cohort was around 20/50 Snellen equivalent in either eye, which is consistent with average visual acuities of treated patients with AMD.[34] While data on binocular distance or near visual acuities were not available, BCVA of the better seeing eye was used as surrogate for binocular visual acuity. We found a significant positive correlation between BCVA of the better seeing eye and SUS score. This means that better visual acuities were associated with higher SUS scores and an overall better user experience. However, the low r^2 value indicates that BCVA alone does not adequately explain the SUS scores. This could be due to nonlinearity of the relationship, or more likely to the contribution of other variables which we did not evaluate such as educational level, presence of systemic comorbidities, degree of cognitive abilities, or degree of health- or computer literacy. Conversely, lower visual acuities would be associated with lower SUS scores, which is what we found in our study with the three lowest SUS scores found in the patients with the lowest visual acuities (Table 2). In addition, in our patient selection, we excluded patients with severe visual impairment despite having accessibility options available. We have purposefully disabled these options so that all patients would have the same experience using a standard "stock" version of the website. Individually altering the website to make it more accessible would have required investigator interference and might have falsely inflated the SUS scores.

Our pilot study nevertheless presents several important limitations. First, our study population was limited to patients with advanced AMD without cognitive or severe visual impairment. Additionally, they were followed for several years at an academic hospital specialized in AMD care and might have better education or be more computer literate. These represent the optimal conditions for patients with AMD, and could thus limit the generalizability of our results. However, on the one hand, with MonŒil's accessibility options and free remote access, this problem can be overcome by involving a caregiver that is more computer literate. On the other hand, MonŒil is not limited to patients with AMD and can be used by anyone comfortable operating a website. Second, patients with low health literacy might find it challenging to even ask relevant questions,[35] and may find leaflets to be more beneficial. One solution would be to offer both options, possibly during patient education workshops, where the booklet would explain the most important facts in layman terms, and *MonŒil* could answer any remaining questions. Third, an improved assessment of patient demographics by evaluating variables such as degree of cognitive abilities or degree of health or computer literacy might yield a better regression model to explain the SUS scores. Finally, while a validation of the individual chatbot responses was not performed, the primary aim of this study was to evaluate the usability *MonŒil*. Its core competencies reflect the strengths and limitations of the GPT-4 technology it is built upon, which has been extensively tested in ophthalmology.[19–23,36–38] In fact, GPT-4 was rated as very accurate by specialists,[21] and outperformed GPT-3.5[11,23] as well as ophthalmologists and residents in answering StatPearl[8] and French European Board of Ophthalmology questions.[9]

In conclusion, our study demonstrates the excellent usability of MonŒil, a ChatGPT-4-based education tool in patients with advanced AMD with exciting directions for the future. One priority is to continuously improve on the website by enhancing its accessibility features, incorporating images and visual aids, expanding language and security options, and accommodating varying levels of health literacy to reach a wider audience. Furthermore, after evaluating its usability, an important step would be to assess the usefulness and clinical impact of MonŒil on patient outcomes such as patient knowledge and vision-related quality of life. It is important to view MonŒil as part of the continuity of care process, ensuring its use remains tightly incorporated in the ophthalmologists' work. Ultimately, this could lead to more effective patient management and improved outcomes, especially in underserved areas.

Financial disclosures

No financial disclosures

Table of contents statement

The study evaluates patient satisfaction with *MonŒil*, an AI-based platform for ophthalmology education, among elderly individuals with advanced age-related macular degeneration. Participants interacted with *MonŒil* for 20 min, providing feedback afterward. Results show excellent usability (mean SUS score of 90.23) and a positive correlation between visual acuity and usability. *MonŒil* proves promising for patient education in ophthalmology, suggesting its potential to supplement physician-patient discussions. Further research is needed for broader application and clinical impact assessment.

Declaration of competing interest

None of the authors has any conflicts of interest to disclose.

Acknowledgements

None

Funding/support

None

References

1. Aung YYM, Wong DCS, Ting DSW. The promise of artificial intelligence: a review of the opportunities and challenges of artificial intelligence in healthcare. *Br Med Bull*. 2021;139(1):4–15.
2. Ting DSW, Cheung CYL, Lim G, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *JAMA*. 2017;318(22):2211–2223.
3. Ting DSJ, Foo VH, Yang LWY, et al. Artificial intelligence for anterior segment diseases: emerging applications in ophthalmology. *Br J Ophthalmol*. 2021;105(2):158–168.
4. Keenan TDL, Chen Q, Agrón E, et al. DeepLensNet: deep learning automated diagnosis and quantitative classification of cataract type and severity. *Ophthalmology*. 2022;129(5):571–584.
5. Grassmann F, Mengelkamp J, Brandl C, et al. A deep learning algorithm for prediction of age-related eye disease study severity scale for age-related macular degeneration from color fundus photography. *Ophthalmology*. 2018;125(9):1410–1420.

6. Girard MJA, Schmetterer L. Artificial intelligence and deep learning in glaucoma: current state and future prospects. *Prog Brain Res*. 2020;257:37–64.
7. OpenAI. GPT-4. https://openai.com/gpt-4.
8. Moshirfar M, Altaf AW, Stoakes IM, Tuttle JJ, Hoopes PC. Artificial intelligence in ophthalmology: a comparative analysis of GPT-3.5, GPT-4, and human expertise in answering StatPearls questions. *Cureus*. 2023;15(6):e40822.
9. Panthier C, Gatinel D. Success of ChatGPT, an AI language model, in taking the French language version of the European Board of Ophthalmology examination: a novel approach to medical knowledge assessment. *J Fr Ophtalmol*. 2023;46(7):706–711.
10. Cai LZ, Shaheen A, Jin A, et al. Performance of generative large language models on ophthalmology board-style questions. *Am J Ophthalmol*. 2023;254:141–149.
11. Cheong KX, Zhang C, Tan TE, et al. Comparing generative and retrieval-based chatbots in answering patient questions regarding age-related macular degeneration and diabetic retinopathy. *Br J Ophthalmol*. 2024;15. Published online May.
12. Bourne RRA, Jonas JB, Flaxman SR, et al. Prevalence and causes of vision loss in high-income countries and in Eastern and Central Europe: 1990-2010. *Br J Ophthalmol*. 2014;98(5):629–638.
13. Burgmüller M, Cakmak N, Weingessel B, Vécsei C, Vécsei-Marlovits PV. Patient knowledge concerning age-related macular degeneration: an AMD questionnaire. *Wien Klin Wochenschr*. 2017;129(9–10):345–350.
14. Ley P. Memory for medical information. *Br J Soc Clin Psychol*. 1979;18(2):245–255.
15. Watson PW, Mckinstry B. A systematic review of interventions to improve recall of medical advice in healthcare consultations. *J R Soc Med*. 2009;102(6):235–243.
16. Shuyler KS, Knight KM. What are patients seeking when they turn to the internet? Qualitative content analysis of questions asked by visitors to an orthopaedics web site. *J Med Internet Res*. 2003;5(4):e24.
17. Chen YY, Li CM, Liang JC, Tsai CC. Health information obtained from the internet and changes in medical decision making: questionnaire development and cross-sectional survey. *J Med Internet Res*. 2018;20(2):e47.
18. Yoo P, Carlone D, Ren LY, Lam WC. Assessment of online health resources for ophthalmology patients with age-related macular degeneration or diabetic retinopathy. *Can J Ophthalmol*. 2016;51(1):e1–e2.
19. Ferro Desideri L, Roth J, Zinkernagel M, Anguita R. Application and accuracy of artificial intelligence-derived large language models in patients with age related macular degeneration. *Int J Retina Vitreous*. 2023;9(1):71.
20. Eid K, Eid A, Wang D, Raiker RS, Chen S, Nguyen J. Optimizing ophthalmology patient education via chatbot-generated materials: readability analysis of AI-Generated patient education materials and the American society of ophthalmic plastic and reconstructive surgery patient brochures. *Ophthalmic Plast Reconstr Surg*. 2023;16. Published online November.
21. Potapenko I, Boberg-Ans LC, Stormly Hansen M, Klefter ON, van Dijk EHC, Subhi Y. Artificial intelligence-based chatbot patient information on common retinal diseases using ChatGPT. *Acta Ophthalmol*. 2023;101(7):829–831.
22. Pushpanathan K, Lim ZW, Er Yew SM, et al. Popular large language model chatbots' accuracy, comprehensiveness, and self-awareness in answering ocular symptom queries. *iScience*. 2023;26(11), 108163.
23. Taloni A, Borselli M, Scarsi V, et al. Comparative performance of humans versus GPT-4.0 and GPT-3.5 in the self-assessment program of American Academy of Ophthalmology. *Sci Rep*. 2023;13(1):18562.
24. Brooke J. SUS: a "quick and dirty' usability scale. In: *Proceedings of the Usability Evaluation in Industry*. 3. 1995:189–194.
25. Brooke J. SUS: a retrospective. *J Usability Stud*. 2013;8(2):29–40.
26. Sauro J, Lewis J. *Quantifying the User Experience: Practical Statistics For User Research*. Morgan Kaufmann; 2016.
27. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. *Int J Med Inform*. 2019;126:95–104.
28. Kortum PT, Bangor A. Usability ratings for everyday products measured with the system usability scale. *Int J Hum Comput Interact*. 2013;29(2):67–76.
29. Hernandez CA, Vazquez Gonzalez AE, Polianovskaia A, et al. The future of patient education: AI-driven guide for type 2 diabetes. *Cureus*. 2023;15(11):e48919.
30. Kuşcu O, Pamuk AE, Sütay Süslü N, Hosal S. Is ChatGPT accurate and reliable in answering questions regarding head and neck cancer? *Front Oncol*. 2023;13, 1256459.
31. Gurnani B, Kaur K. Leveraging ChatGPT for ophthalmic education: a critical appraisal. *Eur J Ophthalmol*. 2023;16, 11206721231215862. Published online November.
32. Schalnus R, Aulmann G, Hellenbrecht A, Hägele M, Ohrloff C, Lüchtenberg M. Content quality of ophthalmic information on the internet. *Ophthalmologica*. 2010;224(1):30–37.
33. Kessels RPC. Patients' memory for medical information. *JRSM*. 2003;96(5):219–222.
34. CATT Research Group, Martin DF, Maguire MG, et al. Ranibizumab and bevacizumab for neovascular age-related macular degeneration. *N Engl J Med*. 2011;364(20):1897–1908.
35. Menendez ME, van Hoorn BT, Mackert M, Donovan EE, Chen NC, Ring D. Patients with limited health literacy ask fewer questions during office visits with hand surgeons. *Clin Orthop Relat Res*. 2017;475(5):1291–1297.
36. Nanji K, Yu CW, Wong TY, et al. Evaluation of postoperative ophthalmology patient instructions from ChatGPT and Google Search. *Can J Ophthalmol*. 2023;24. Published online October.
37. Nikdel M, Ghadimi H, Tavakoli M, Suh DW. 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*. 2023:1–4. Published online October 25.
38. Fowler T, Pullen S, Birkett L. Performance of ChatGPT and Bard on the official part 1 FRCOphth practice questions. *Br J Ophthalmol*. 2023. Published online November 6.