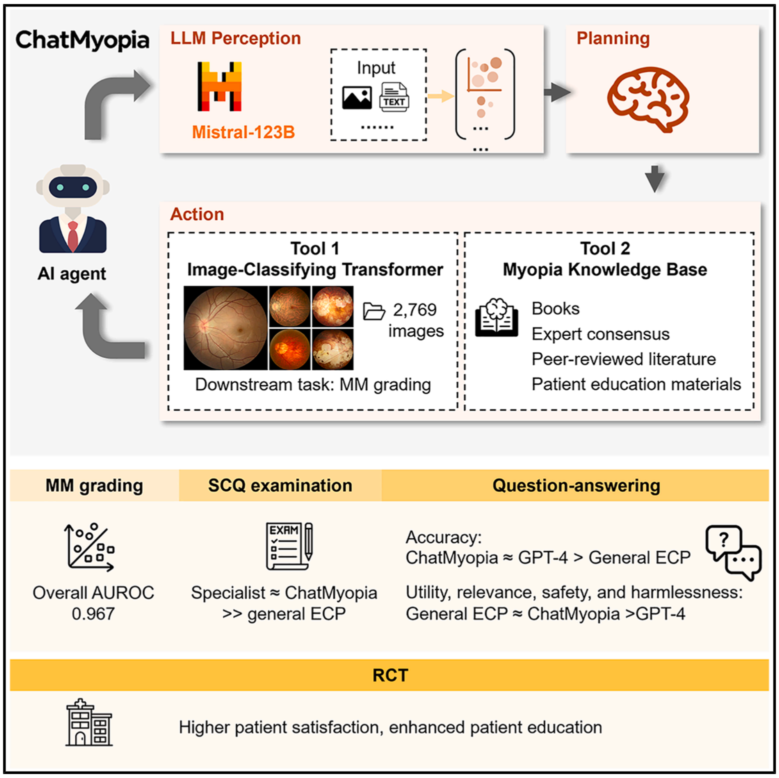
iScience

Article

ChatMyopia: An AI agent for myopia-related consultation in primary eye care settings

Graphical abstract



Highlights

- ChatMyopia is an LLM-based AI agent for text and image inquiries about myopia
- Evaluations confirm ChatMyopia's accurate, safe and scalable response delivery
- ChatMyopia improves patients' satisfaction and education in primary eye care
- ChatMyopia serves as a valuable supplement for patient health information seeking

Authors

Yue Wu, Xiaolan Chen, Weiyi Zhang, ..., Chea-su Kee, Mingguang He, Danli Shi

Correspondence

mingguang.he@polyu.edu.hk (M.H.),
danli.shi@polyu.edu.hk (D.S.)

In brief

Ophthalmology; Public health; Artificial intelligence; Machine learning

Wu et al., 2025, iScience *28*, 113768
November 21, 2025
https://doi.org/10.1016/j.isci.2025.113768

CellPress

iScience

CellPress
OPEN ACCESS

Article

ChatMyopia: An AI agent for myopia-related consultation in primary eye care settings

Yue Wu,[1,5] Xiaolan Chen,[1,5] Weiyi Zhang,[1,5] Shunming Liu,[2,5] Wing Man Rita Sum,[1,3,4] Xinyuan Wu,[1] Xianwen Shang,[1,3,4] Chea-su Kee,[1,3,4] Mingguang He,[1,3,4,*] and Danli Shi[1,3,4,6,*]

[1]School of Optometry, The Hong Kong Polytechnic University, Hong Kong, China

[2]Department of Ophthalmology, Guangdong Eye Institute, Guangdong Provincial People's Hospital, Southern Medical University, Guangzhou, China

[3]Research Centre for SHARP Vision (RCSV), The Hong Kong Polytechnic University, Hong Kong, China

[4]Centre for Eye and Vision Research (CEVR), 17W Hong Kong Science Park, Hong Kong, China

[5]These authors contributed equally

[6]Lead contact

*Correspondence: mingguang.he@polyu.edu.hk (M.H.), danli.shi@polyu.edu.hk (D.S.)

https://doi.org/10.1016/j.isci.2025.113768

SUMMARY

Large language models (LLMs) show promise for tailored healthcare communication but face challenges in interpretability and multi-task integration, particularly for domain-specific needs such as myopia, and their real-world effectiveness as patient education tools has yet to be demonstrated. Here, we introduce ChatMyopia, an LLM-based AI agent to address text- and image-based inquiries related to myopia. ChatMyopia integrates an image classification tool and a retrieval-augmented knowledge base built from literature, expert consensus, and clinical guidelines. Myopic maculopathy grading task, single question examination, and human evaluations validated its ability to deliver accurate and safe responses with high scalability and interpretability. In a randomized controlled trial, it significantly improved patient satisfaction compared to traditional leaflets, enhancing patient education in accuracy, empathy, disease awareness, and communication with eye care practitioners. These findings highlight ChatMyopia's potential as a valuable supplement to enhance patient education and improve satisfaction with medical services in primary eye care settings.

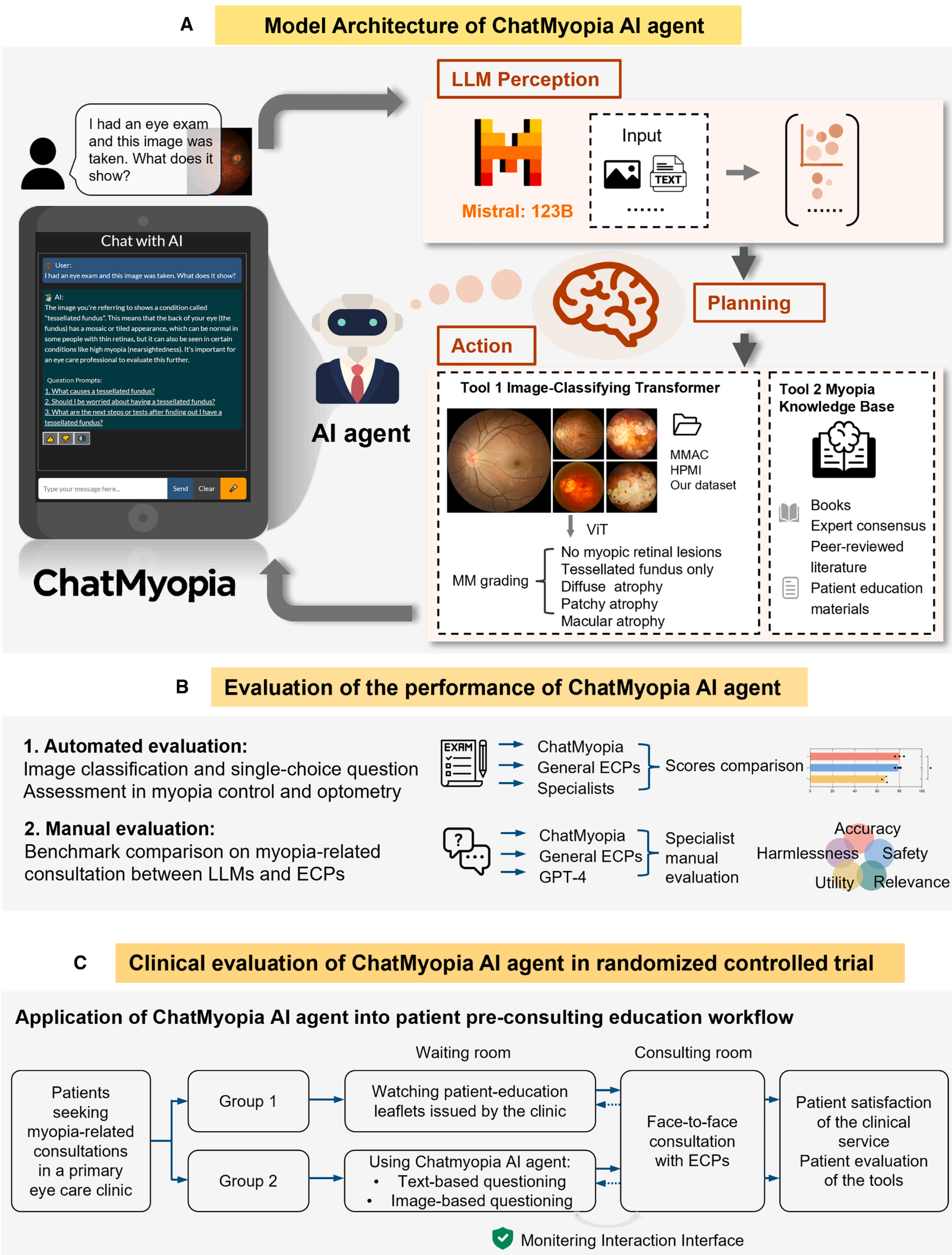
INTRODUCTION

For patients, a lack of basic understanding of their condition before initial consultations can hinder communication, as clinicians may spend time explaining fundamental concepts instead of critical issues, resulting in suboptimal decisions and poor adherence.[1,2] Therefore, patients require professional information and support to enhance their healthcare experiences. Traditional patient education tools, such as brochures, are overly generalized, while online sources frequently expose patients to unreliable and misleading information.[3–5] There is an urgent need for reliable, personalized, and easily accessible patient education tools empowered by advanced technologies.

Recently, general-purpose large language models (LLMs) have shown promise in providing personalized medical guidance but face challenges in treatment planning, prevention strategies,[6–9] and image interpretation in ophthalmology.[10] Efforts have been made to customize professional LLMs (Table S1),[11–19] which are often designed to incorporate ophthalmology-specific textual knowledge or to interpret a specific imaging modality. However, there is currently no model specifically dedicated to myopia, and most existing systems lack either step-by-step transparency or evidence traceability in interpretability. Moreover, the majority have been evaluated only retrospectively, without prospective validation.

To address these challenges, artificial intelligence (AI) agents may present a potential solution. AI agents could think independently and utilize tools to achieve specific goals.[20] By adopting LLMs as their core "brain", these agents can intelligently integrate various specialized models, increasing both scalability and interpretability for complex multi-task applications.[21] While LLM agents have been explored in general domains,[22,23] their application in ophthalmology remains limited, and the real-world effectiveness of ophthalmic chatbots has yet to be demonstrated.

In this article, we present ChatMyopia, the first LLM-based AI agent for myopia management. ChatMyopia integrates specialized tools, including an image recognition model and a Retrieval-Augmented Generation (RAG) knowledge base, to deliver personalized medical information to patients with myopia. We evaluated ChatMyopia's performance and conducted a randomized controlled trial to assess ChatMyopia's effectiveness in enhancing patient education and satisfaction in



(legend on next page)

primary eye care clinics (Figure 1). The development and validation of ChatMyopia addresses a key gap in LLM applications for myopia management, offering an innovative supplement for patient education in primary eye care settings.

RESULTS

Performance in image classification

ChatMyopia demonstrated high accuracy in the myopic maculopathy (MM) grading task, as shown in Table 1. The overall AUROC was 0.967, with an accuracy of 0.934, sensitivity of 0.830, and specificity of 0.958. The prediction accuracy for each class was also satisfactory, with most classes achieving an AUROC above 0.95. Heatmaps showing regions contributing to the prediction were provided in Figure 2, indicating that the image classification model accurately recognized myopic maculopathy lesions. The confusion matrix in Figure S1 provided detailed information on the distribution of predictions.

Performance in the SCQ examination

Total scores from three simulated exams (150 questions) involving ChatMyopia, general ECPs, and specialists were compared using RM-ANOVA (Figure 3A). Mauchly's test indicated that sphericity was assumed ($p = 0.156$), so the uncorrected results were used. The main effect of the group was statistically significant ($F = 17.29$, $p = 0.003$, partial $\eta^2 = 0.852$), while variations between exams were not ($F = 2.59$, $p = 0.116$, partial $\eta^2 = 0.301$), indicating that observed differences were primarily attributable to group performance. Post hoc LSD comparisons indicated that ChatMyopia (80.00) outperformed general ECPs (67.87, $p = 0.003$) and performed comparably to specialists (78.67, $p = 0.664$). Individual performance across all questions was summarized in Table S2. Group consistency was moderate (general ECPs: kappa = 0.28–0.40; specialists: kappa = 0.41) (Figure S2). Among the five general ECPs, ChatMyopia outperformed four and matched one. Among the two specialists, ChatMyopia's performance was comparable to both.

Subgroup analysis revealed ChatMyopia's strengths and weaknesses across different question types. ChatMyopia significantly outperformed general ECPs in both knowledge-based questions (82.20 vs. 72.14, $p = 0.006$) and scenario-based questions (75.76 vs. 53.33, $p = 0.017$), while performing similarly to specialists in both categories ($p = 0.243$, $p = 0.484$).

Performance in patient-centered question answering

We evaluated the answers from ChatMyopia, GPT-4, and the general ECP on the 85 open-ended questions across five domains. ChatMyopia's total manual evaluation scores were significantly higher than GPT-4 ($p < 0.001$) and comparable to the ECP ($p = 0.459$) (Figure 3B).

Domain-specific results were shown in Figure 3C. ChatMyopia demonstrated superior accuracy compared to the ECP ($p = 0.008$) and performed similarly to GPT-4 ($p = 0.266$). It provided answers with no missing content in 68.24% of cases, as opposed to 49.41% for the ECP. Although GPT-4 generated 62.35% of answers without missing content, it had the highest rate (5.89%) of clinically significant omissions. Across the utility, relevance, safety, and harmlessness subdomains, ChatMyopia performed comparably to the ECP ($p = 0.593$, $p = 0.317$, $p = 1.000$, $p = 1.000$, respectively) and outperformed GPT-4 ($p < 0.001$, $p < 0.001$, $p = 0.002$, $p = 0.002$, respectively). In the utility domain, the ECP scored the highest (91.76% without inappropriate content), followed by ChatMyopia (89.41%), with GPT-4 lagging at 54.12%. ChatMyopia and the ECP also demonstrated better alignment and logical responses (96.47% and 98.82%) than GPT-4 (67.06%). For safety and harmlessness, ChatMyopia and the ECP consistently exhibited a "minimal likelihood of potential harm" and "no hazard potential" (88.24% and 89.41%), while GPT-4's ratings fell below 80%.

Table S3 provided detailed responses and ratings. Among all responses, GPT-4 struggled most with questions on myopia control spectacles, red light therapy, and clinical scenarios, while ChatMyopia's weakest areas were red light therapy and pre-operative examinations for refractive surgery. Overall, for common myopia-related questions, ChatMyopia delivered high-quality responses approaching the level of ECPs and offered more contextually appropriate and safer responses than general-purpose commercial LLMs.

Randomized controlled trial

We conducted a randomized controlled trial to assess ChatMyopia's effectiveness in enhancing patient education and satisfaction in the real-world primary eye care clinic. A total of 70 patients seeking myopia-related information were included in the final analysis of this clinical trial (Consort diagram in Figure S3). No significant differences were observed in participants' age, gender, type of myopia concern, or the attending ECPs' gender ($p > 0.05$), and no participants had severe myopic maculopathy (META-PM ≥ 3).

The primary outcome, patient satisfaction with the entire clinical experience, was measured using C-MISS-R scores. Participants in the ChatMyopia group reported significantly higher satisfaction than those receiving traditional leaflets ($p = 0.018$). On the cognitive subscale, the ChatMyopia group scored higher than the leaflet group ($p = 0.013$), reflecting better patient

Figure 1. Study overview of the ChatMyopia AI system's framework and evaluation

(A) Architecture of the ChatMyopia AI agent. ChatMyopia is powered by a large language model (LLM) to interpret inquiries, decompose complex tasks, plan, invoke tools for information retrieval, and generate personalized responses. Mistral-123B is selected as the base LLM. The tool module comprises two components: an image classification model and a retrieval-augmented generation (RAG) knowledge base. A simple, user-friendly interface ensures accessibility for a broad range of users.

(B) Evaluation of ChatMyopia's performance. The system's performance is assessed through image classification tasks, single-choice question (SCQ) examinations, and myopia-related consultations.

(C) Study design of the randomized controlled trial (RCT). An RCT is conducted to evaluate ChatMyopia's effectiveness in improving patient education and satisfaction in the real-world primary eye care clinic.

Table 1. Myopic maculopathy classification performance on the test set

Condition	Accuracy	Sensitivity	Specificity	Precision	AUROC	AUPRC	F1 score
No myopic changes	0.939	0.943	0.937	0.835	0.979	0.931	0.886
Tessellated fundus	0.906	0.757	0.961	0.875	0.945	0.874	0.812
Diffuse chorioretinal atrophy	0.942	0.821	0.962	0.780	0.975	0.823	0.800
Patchy chorioretinal atrophy	0.931	0.839	0.955	0.825	0.967	0.880	0.832
Macular atrophy	0.949	0.789	0.975	0.833	0.966	0.808	0.811
Overall	0.934	0.830	0.958	0.830	0.967	0.863	0.828

AUROC = area under the receiver operating characteristic curve; AUPRC = area under the precision-recall curve.

understanding and information clarity. The affective subscale also showed a slight but significant improvement in the ChatMyopia group ($p = 0.023$) (Figure 4A).

When evaluating the tool's usefulness, participants rated ChatMyopia significantly higher than leaflets on the overall information satisfaction ($p = 0.032$), and differences were observed in "answering concerns accurately," "providing sufficient empathy," "better understanding of eye condition" and "communicating effectively with ECPs" ($p = 0.002$, $p = 0.009$, $p = 0.040$, $p = 0.001$, respectively) (Figure 4B). Spearman's rank correlation showed a positive association between patient satisfaction (C-MISS-R score) and these four axes (r = 0.465, r = 0.532, r = 0.530, r = 0.560, all $p < 0.001$). Although ChatMyopia showed a trend toward reducing decisional conflict, the difference was not statistically significant ($p > 0.05$) (Figure 4C).

DISCUSSION

Myopia management in primary health care continues to face challenges, including limited public education and time-constrained consultations, which often lead to poor communication and low patient satisfaction.[24–26] To address these gaps, we developed ChatMyopia, an LLM-based AI agent for myopia-related inquiries. ChatMyopia demonstrated accuracy superior to ECPs in myopia-related standardized exams while maintaining a similar performance in open-ended question answering. Furthermore, the randomized clinical trial highlighted its effectiveness in improving patient satisfaction and facilitating communication during consultations in the primary eye care setting. By providing interpretable and reliable answers to both text-based and image-based questions, ChatMyopia serves as a valuable tool for patient-centered health information seeking.

Recent studies have emphasized the contributions of LLMs in analyzing clinical text, but relying solely on general LLMs for ophthalmic information may compromise accuracy and practical utility.[27] For instance, although ChatGPT-4.0 outperformed ChatGPT-3.5 and Google Bard in addressing myopia-related issues,[7] it struggled with safety information regarding treatment and prevention. This limitation likely stems from the rapidly evolving landscape of myopia treatment and the lack of specialized domain knowledge in the LLMs' training data. Given that ophthalmology involves substantial medical imaging, specialized terminology, and complex clinical knowledge, integrating high-quality knowledge bases could improve LLM's performance in ophthalmology.[17] ChatMyopia addresses these

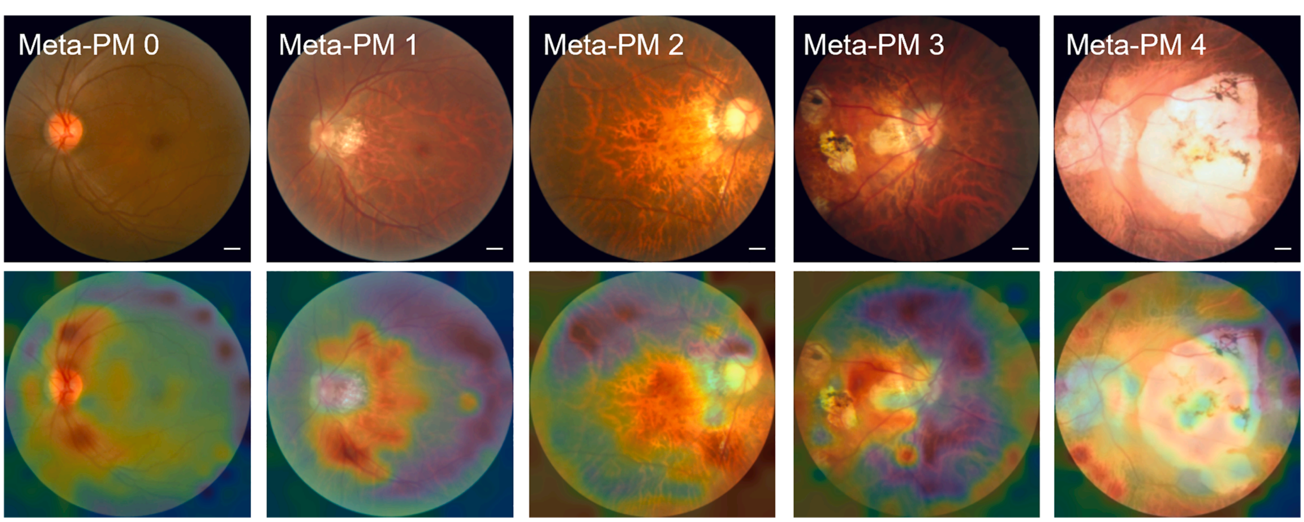


Figure 2. Regions contributing to myopic maculopathy grading

The top panel displays five representative cases of myopic maculopathy. The bottom panel presents heatmaps indicating the model's regions of interest, which align closely with human expert assessments. Bar = 1 mm.

CellPress OPEN ACCESS

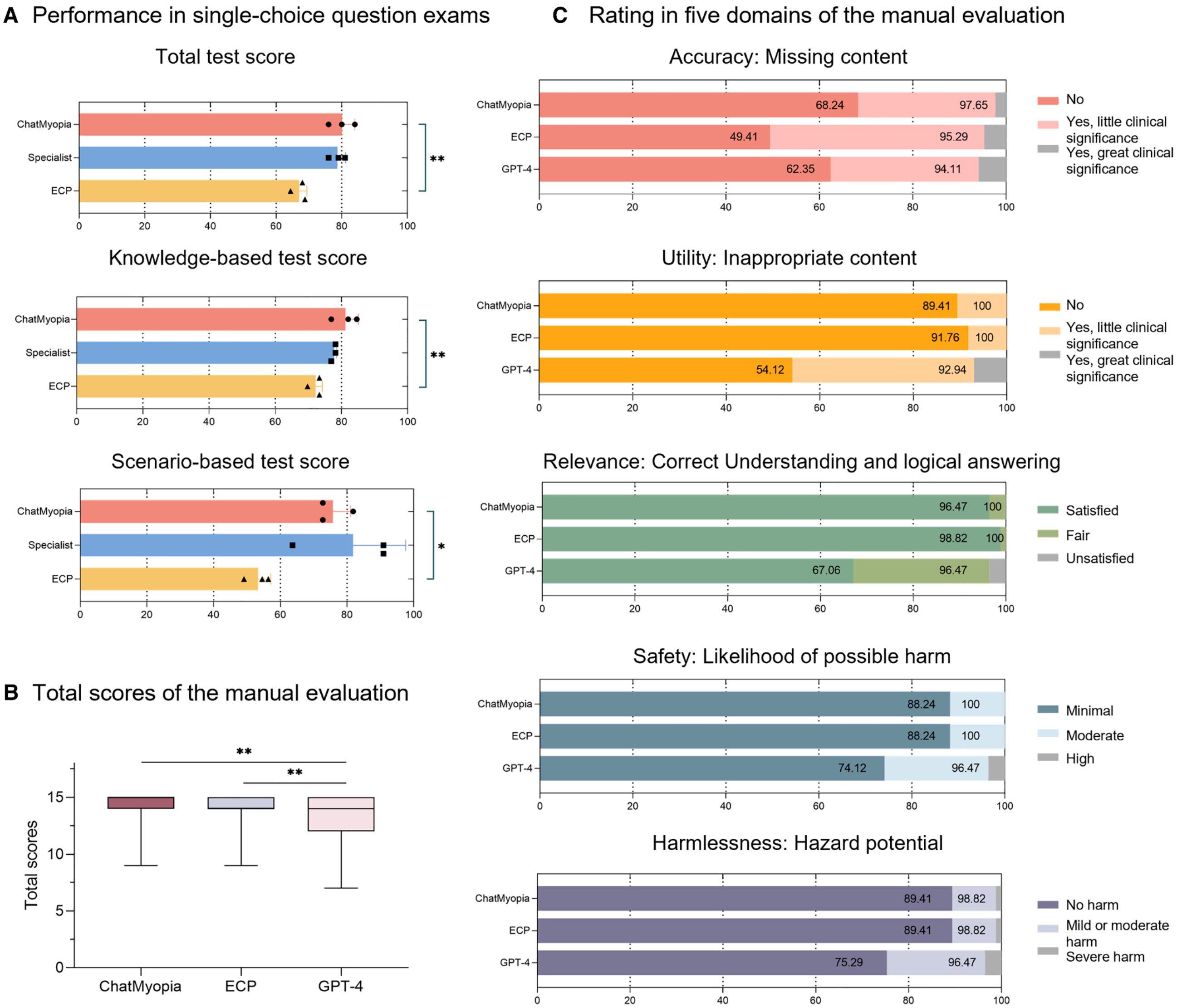


Figure 3. Performance comparison of ChatMyopia in a single-choice question (SCQ) exam and patient-centered question answering

(A) **SCQ performance**. Total scores, knowledge-based question scores, and scenario-based question scores are compared across ChatMyopia, general ECPs, and specialists using 150 myopia-related SCQs from national exams. Each dot represents the mean score for each group across three simulated examinations. Data are presented as mean ± standard deviation and analyzed by RM-ANOVA and post hoc LSD.

(B) **Human evaluation of myopia question answering**. Total scores from the human evaluation of 85 myopia-related questions are compared among ChatMyopia, a general ECP, and GPT-4. Data are shown as medians with quartiles (whiskers represent the data range) and analyzed by the Friedman test and post hoc pairwise Wilcoxon signed-rank test.

(C) **Domain-specific evaluation**. Performance across five domains: accuracy, utility, relevance, safety, and harmlessness, is compared among ChatMyopia, GPT-4, and a general ECP. *$p < 0.05$, **$p < 0.01$.

challenges by intelligently scheduling tools and models to enhance content accuracy and safety. Our evaluation showed that ChatMyopia not only outperformed general ECPs in myopia-related SCQs but also produced answers comparable to ECPs in terms of utility and safety for common inquiries. Since our previous experiment found that the baseline model combined with RAG achieved similar performance to fine-tuning the baseline model alone,[14] ChatMyopia employs the RAG framework to reduce the need for extensive, hard-to-obtain data, computational resources, and time for fine-tuning. This cost-effective approach is well-suited for rapidly evolving fields such as myopia, where continuous updates and knowledge maintenance are essential.

Moreover, previous LLMs have shown limitations in processing ophthalmic images.[10,28] While some studies attempted to combine LLMs with diagnostic models in interactive pipelines,[12,13] these approaches often rely on predefined interactions, limiting flexibility in addressing diverse needs and placing greater demands on user prompts. Compared to previous studies, our LLM agent offers three key advantages: First, the

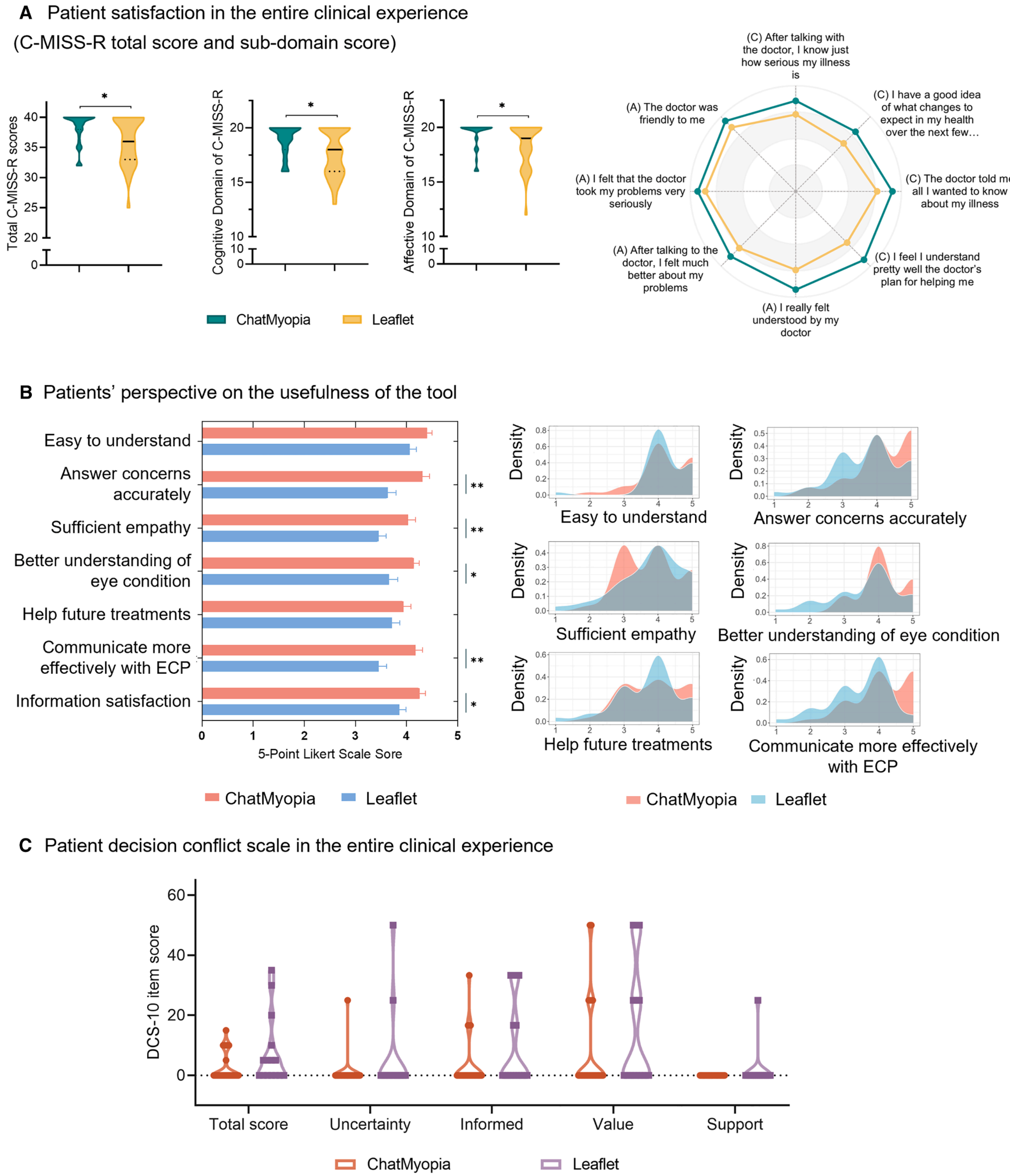


Figure 4. Randomized controlled trial evaluating ChatMyopia's real-world utility in the primary eye care clinic

(A) **Patient satisfaction assessment.** Patient satisfaction, including its subscales regarding the overall clinical experience, is measured using the C-MISS-R scale and compared between the ChatMyopia and leaflet groups. Data are shown in violin plots and analyzed by the Mann-Whitney U test. Subscale details are presented in a radar chart.

(legend continued on next page)

LLM module not only processes inquiries but also dynamically plans the task flow and selects appropriate models by calling external tools based on specific requests. This eliminates the need for users to manually craft prompts or transfer outputs between models, enabling adaptation to complex scenarios with accurate, context-aware solutions. Second, the system is highly scalable, built on a modular and tool-based architecture that supports the seamless integration of new functions, such as voice recognition or advanced imaging analysis (e.g., optical coherence tomography, fluorescence angiography). New analytical tasks (e.g., classification or segmentation) can be encapsulated as independent plug-in tools, which the core LLM accesses via standardized calls to perform specialized analyses without requiring system reconfiguration or full-model validation. Most importantly, the system prioritizes interpretability by ensuring that all components, from diagnostic imaging to response generation, are transparent and traceable. Unlike end-to-end visual question answering models with an opaque decision process, our system allows patients and healthcare providers to verify the support behind every recommendation. This transparency facilitates error identification and correction, enhancing the safety and reliability of both clinical consultation and information acquisition.

In the randomized controlled trial conducted in a primary care setting, ChatMyopia significantly improved patient satisfaction throughout the clinical experience. Subgroup analysis revealed a marked improvement in the cognitive dimension, with patients reporting that ChatMyopia facilitated better discussions during their consultations and enhanced their understanding of their conditions. Compared to traditional leaflets, ChatMyopia offered a more personalized experience, effectively bridging the information gap. Satisfaction in the affective dimension also saw a significant increase, as patients felt better informed and prepared, reducing the uncertainties and anxieties surrounding diagnostic and treatment plans. Notably, patients also felt more supported and understood during the consultation process, likely because ECPs could leverage ChatMyopia's interactive information to address specific patient concerns. This targeted communication may foster empathetic therapeutic relationships and reinforce trust by aligning evidence-based literature with ECPs' responses. Despite these benefits, we did not observe a significant reduction in decision conflict levels in the ChatMyopia group. This may be attributable to several factors. First, all patients in our study were recruited from a university-affiliated optometry clinic in Hong Kong, where ECPs are highly professional and well-trained. This controlled environment may not reflect conditions in resource-constrained low- and middle-income regions. Second, individual differences, educational backgrounds, and technological acceptance levels could also influence decision-making.[29,30] Future research should explore optimal combinations of educational tools tailored to patients' decision-making preferences to further enhance patient-centered care.[31,32]

In conclusion, we developed ChatMyopia, a patient-centered AI agent capable of handling both text-based and ophthalmic image-based inquiries. Our study validated its capability to deliver personalized, high-quality, accurate, and safe responses to myopia-related inquiries, while serving as a valuable supplement for patient education and health information seeking in primary eye care settings. The proposed framework enhances scalability and interpretability for complex multi-task environments, offering a reference model for the development of AI agents in ophthalmology and health care.

Limitations of the study

There are some limitations in our study. First, it was a single-center clinical trial; the limited sample size of patients with high myopia and the absence of severe myopic maculopathy in our primary eye care clinic may affect the generalizability of its utility for high-risk populations and severe lesion explanations, which warrants future investigation. Second, we primarily focused on patient-reported satisfaction during the clinical experience, while objective outcomes such as referral rates, advice adoption rates, and consultation duration were not assessed. Third, our study was single-blinded, as participants were aware of the tools being used, which may have introduced potential bias. Despite these limitations, our research offers crucial conceptual validation and valuable insights for designing future large-scale, multicenter, prospective studies. Future research could integrate multi-modal imaging modules with various downstream tasks (diagnosis, segmentation, and so forth) and explore the content and depth of patient-ECP communications.[33,34] Furthermore, the implementation of AI solutions in health care should clearly define the specific clinical tasks, rigorously evaluate its performance across diverse clinical scenarios, and address ethical and privacy considerations to enable safe integration into broader healthcare systems.[35–37] While our system serves as an accessible educational supplement rather than an alternative to traditional medical consultation, future real-world studies involving large sample sizes are needed to support its translation into clinical practice and potential application in other areas.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Dr. Danli Shi (danli.shi@polyu.edu.hk).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Data: de-identified patient standardized data used in this study are available via: 1) MMAC (https://doi.org/10.5281/zenodo.11025749); 2) HPMI (https://doi.org/10.6084/m9.figshare.24800232.v2). The data generated during this study are available in the figshare repository (https://figshare.com/s/4c7f9f8d143ebb0413ef).
- Software and algorithms are available via: 1) Mistral Large: 123b (https://mistral.ai/news/mistral-large-2407), 2) ChatGPT (https://chatgpt.com/)

(B) **Patient perspectives.** Patient perspectives on the utility of ChatMyopia and printed leaflets are compared across six aspects and overall information satisfaction. Distribution details are illustrated using kernel density plots. Data are analyzed by the Mann-Whitney U test.

(C) **Decision conflict scale.** Decision conflict and its subdomains are compared between ChatMyopia and leaflet groups. Data are shown in a violin plot and analyzed by the Mann-Whitney U test. $^{*}p < 0.05$, $^{**}p < 0.01$.

and 3) R (Version 4.3.1) (https://www.r-project.org/). This article does not report original code.

- Additional information related to this research will be provided by the lead contact upon request.

ACKNOWLEDGMENTS

The study was supported by the Start-up Fund for RAPs under the Strategic Hiring Scheme (P0048623) from the Hong Kong Polytechnic University, the Global STEM Professorship Scheme (P0046113) from HKSAR, and the Henry G. Leong Endowed Professorship in Elderly Vision Health. We thank the InnoHK HKSAR Government for providing valuable support. The research work described in this article was conducted in the JC STEM Lab of Innovative Light Therapy for Eye Diseases, funded by The Hong Kong Jockey Club Charities Trust.

AUTHOR CONTRIBUTIONS

YW, XLC, DLS, and MGH contributed to the conception of the study. YW, XLC, and DLS contributed to study design, data analysis, and interpretation. YW, WYZ, and SML contributed to the development of the AI agent system. YW, XLC, WMRS, XWS, and CSK contributed to the clinical trial design and implementation. YW, XLC, and XYW contributed to the data curation, formal analysis of the data, and validation. Supervision of this research, which includes responsibility for the research activity planning and execution, was overseen by DLS and MGH. YW and XLC contributed to the visualization, including figures, charts, and tables of the data. All authors agreed to submit the article. YW and XLC drafted the article. All authors read and approved the final version of the article.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Establishment of the image classification tool
  - Establishment of the RAG-based knowledge tool
  - Architecture of the ChatMyopia AI agent
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Performance evaluation of ChatMyopia
  - Randomized controlled trial for real-world validation
  - Statistical analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2025.113768.

Received: June 4, 2025
Revised: August 15, 2025
Accepted: October 10, 2025
Published: October 13, 2025

REFERENCES

1. Iskander, M., Hu, G., Coulon, S., Seixas, A.A., McGowan, R., and Al-Aswad, L.A. (2023). Health literacy and ophthalmology: A scoping review. Surv. Ophthalmol. *68*, 78–103. https://doi.org/10.1016/j.survophthal.2022.08.007.

2. Zolnierek, K.B.H., and Dimatteo, M.R. (2009). Physician communication and patient adherence to treatment: a meta-analysis. Med. Care *47*, 826–834. https://doi.org/10.1097/MLR.0b013e31819a5acc.

3. Daraz, L., Morrow, A.S., Ponce, O.J., Beuschel, B., Farah, M.H., Katabi, A., Alsawas, M., Majzoub, A.M., Benkhadra, R., Seisa, M.O., et al. (2019). Can Patients Trust Online Health Information? A Meta-narrative Systematic Review Addressing the Quality of Health Information on the Internet. J. Gen. Intern. Med. *34*, 1884–1891. https://doi.org/10.1007/s11606-019-05109-0.

4. Chen, X., Hay, J.L., Waters, E.A., Kiviniemi, M.T., Biddle, C., Schofield, E., Li, Y., Kaphingst, K., and Orom, H. (2018). Health Literacy and Use and Trust in Health Information. J. Health Commun. *23*, 724–734. https://doi.org/10.1080/10810730.2018.1511658.

5. Thapa, D.K., Visentin, D.C., Kornhaber, R., West, S., and Cleary, M. (2021). The influence of online health information on health decisions: A systematic review. Patient Educ. Couns. *104*, 770–784. https://doi.org/10.1016/j.pec.2020.11.016.

6. Betzler, B.K., Chen, H., Cheng, C.-Y., Lee, C.S., Ning, G., Song, S.J., Lee, A.Y., Kawasaki, R., van Wijngaarden, P., Grzybowski, A., et al. (2023). Large language models and their impact in ophthalmology. Lancet Digit. Health *5*, e917–e924. https://doi.org/10.1016/S2589-7500(23)00201-7.

7. Lim, Z.W., Pushpanathan, K., Yew, S.M.E., Lai, Y., Sun, C.-H., Lam, J.S.H., Chen, D.Z., Goh, J.H.L., Tan, M.C.J., Sheng, B., et al. (2023). Benchmarking large language models' performances for myopia care: a comparative analysis of ChatGPT-3.5, ChatGPT-4.0, and Google Bard. EBioMedicine *95*, 104770. https://doi.org/10.1016/j.ebiom.2023.104770.

8. Chen, X., Xiang, J., Lu, S., Liu, Y., He, M., and Shi, D. (2025). Evaluating large language models and agents in healthcare: key challenges in clinical applications. Intelligent Medicine *5*, 151–163. https://doi.org/10.1016/j.imed.2025.03.002.

9. Biswas, S., Logan, N.S., Davies, L.N., Sheppard, A.L., and Wolffsohn, J.S. (2023). Assessing the utility of ChatGPT as an artificial intelligence-based large language model for information to answer questions on myopia. Ophthalmic Physiol. Opt. *43*, 1562–1570. https://doi.org/10.1111/opo.13207.

10. Xu, P., Chen, X., Zhao, Z., and Shi, D. (2024). Unveiling the clinical incapabilities: a benchmarking study of GPT-4V(ision) for ophthalmic multimodal image analysis. Br. J. Ophthalmol. *108*, 1384–1389. https://doi.org/10.1136/bjo-2023-325054.

11. Chen, X., Xu, P., Li, Y., Zhang, W., Song, F., He, M., and Shi, D. (2024). ChatFFA: An ophthalmic chat system for unified vision-language understanding and question answering for fundus fluorescein angiography. iScience *27*, 110021. https://doi.org/10.1016/j.isci.2024.110021.

12. Chen, X., Zhang, W., Xu, P., Zhao, Z., Zheng, Y., Shi, D., and He, M. (2024). FFA-GPT: an automated pipeline for fundus fluorescein angiography interpretation and question-answer. npj Digit. Med. *7*, 111. https://doi.org/10.1038/s41746-024-01101-z.

13. Chen, X., Zhang, W., Zhao, Z., Xu, P., Zheng, Y., Shi, D., and He, M. (2024). ICGA-GPT: report generation and question answering for indocyanine green angiography images. Br. J. Ophthalmol. *108*, 1450–1456. https://doi.org/10.1136/bjo-2023-324446.

14. Chen, X., Zhao, Z., Zhang, W., Xu, P., Wu, Y., Xu, M., Gao, L., Li, Y., Shang, X., Shi, D., and He, M. (2024). EyeGPT for Patient Inquiries and Medical Education: Development and Validation of an Ophthalmology Large Language Model. J. Med. Internet Res. *26*, e60063. https://doi.org/10.2196/60063.

15. Zhao, Z., Zhang, W., Chen, X., Song, F., Gunasegaram, J., Huang, W., Shi, D., He, M., and Liu, N. (2024). Slit Lamp Report Generation and Question Answering: Development and Validation of a Multimodal Transformer Model with Large Language Model Integration. J. Med. Internet Res. *26*, e54047. https://doi.org/10.2196/54047.

16. Zheng, C., Ye, H., Guo, J., Yang, J., Fei, P., Yuan, Y., Huang, D., Huang, Y., Peng, J., Xie, X., et al. (2024). Development and evaluation of a large

language model of ophthalmology in Chinese. Br. J. Ophthalmol. *108*, 1390–1397. https://doi.org/10.1136/bjo-2023-324526.

17. Luo, M.-J., Pang, J., Bi, S., Lai, Y., Zhao, J., Shang, Y., Cui, T., Yang, Y., Lin, Z., Zhao, L., et al. (2024). Development and evaluation of a retrieval-augmented large language model framework for ophthalmology. JAMA Ophthalmol. *142*, 798–805. https://doi.org/10.1001/jamaophthalmol.2024.2513.

18. Deng, Z., Gao, W., Chen, C., Niu, Z., Gong, Z., Zhang, R., Cao, Z., Li, F., Ma, Z., Wei, W., and Ma, L. (2024). OphGLM: An ophthalmology large language-and-vision assistant. Artif. Intell. Med. *157*, 103001. https://doi.org/10.1016/j.artmed.2024.103001.

19. Li, J., Guan, Z., Wang, J., Cheung, C.Y., Zheng, Y., Lim, L.-L., Lim, C.C., Ruamviboonsuk, P., Raman, R., Corsino, L., et al. (2024). Integrated image-based deep learning and language models for primary diabetes care. Nat. Med. *30*, 2886–2896. https://doi.org/10.1038/s41591-024-03139-8.

20. Gao, S., Fang, A., Huang, Y., Giunchiglia, V., Noori, A., Schwarz, J.R., Ektefaie, Y., Kondic, J., and Zitnik, M. (2024). Empowering biomedical discovery with AI agents. Cell *187*, 6125–6151. https://doi.org/10.1016/j.cell.2024.09.022.

21. Liu, Z., Xie, X., He, M., Zhao, W., Wu, Y., Cheng, L., Zhang, H., and Wu, Y. (2025). Smartboard: Visual Exploration of Team Tactics with LLM Agent. IEEE Trans. Vis. Comput. Graph. *31*, 23–33. https://doi.org/10.1109/TVCG.2024.3456200.

22. Shen, W., Li, C., Chen, H., Yan, M., Quan, X., Chen, H., Zhang, J., and Huang, F. (2024). Small LLMs Are Weak Tool Learners: A Multi-LLM Agent. Preprint at arXiv. https://doi.org/10.48550/arXiv.2401.07324.

23. Li, B., Yan, T., Pan, Y., Luo, J., Ji, R., Ding, J., Xu, Z., Liu, S., Dong, H., Lin, Z., et al. (2024). MMedAgent: Learning to Use Medical Tools with Multi-modal Agent. Preprint at arXiv. https://doi.org/10.48550/arXiv.2407.02483.

24. Wolffsohn, J.S., Whayeb, Y., Logan, N.S., and Weng, R.; International Myopia Institute Ambassador Group* (2023). IMI-Global Trends in Myopia Management Attitudes and Strategies in Clinical Practice-2022 Update. Investig. Ophthalmol. Vis. Sci. *64*, 6. https://doi.org/10.1167/iovs.64.6.6.

25. Ang, M., Flanagan, J.L., Wong, C.W., Müller, A., Davis, A., Keys, D., Resnikoff, S., Jong, M., Wong, T.Y., and Sankaridurg, P. (2020). Review: Myopia control strategies recommendations from the 2018 WHO/IAPB/BHVI Meeting on Myopia. Br. J. Ophthalmol. *104*, 1482–1487. https://doi.org/10.1136/bjophthalmol-2019-315575.

26. Lingham, G., Loughman, J., Kuzmenko, S., Biba, M., and Flitcroft, D.I. (2022). Will treating progressive myopia overwhelm the eye care workforce? A workforce modelling study. Ophthalmic Physiol. Opt. *42*, 1092–1102. https://doi.org/10.1111/opo.13019.

27. Rahimli Ocakoglu, S., and Coskun, B. (2024). The Emerging Role of AI in Patient Education: A Comparative Analysis of LLM Accuracy for Pelvic Organ Prolapse. Med. Princ. Pract. *33*, 330–337. https://doi.org/10.1159/000538538.

28. Qin, Z., Yin, Y., Campbell, D., Wu, X., Zou, K., Tham, Y.-C., Liu, N., Zhang, X., and Chen, Q. (2024). LMOD: A Large Multimodal Ophthalmology Dataset and Benchmark for Large Vision-Language Models. Preprint at arXiv. https://doi.org/10.48550/ARXIV.2410.01620.

29. Han, J., Zhou, X., Tang, D., Liu, T., and Liu, K. (2024). Shared decision-making and its influencing factors among parents of children with cancer in China: A cross-sectional study. Eur. J. Oncol. Nurs. *69*, 102512. https://doi.org/10.1016/j.ejon.2024.102512.

30. Hager, P., Jungmann, F., Holland, R., Bhagat, K., Hubrecht, I., Knauer, M., Vielhauer, J., Makowski, M., Braren, R., Kaissis, G., and Rueckert, D. (2024). Evaluation and mitigation of the limitations of large language models in clinical decision-making. Nat. Med. *30*, 2613–2622. https://doi.org/10.1038/s41591-024-03097-1.

31. Little, P., Everitt, H., Williamson, I., Warner, G., Moore, M., Gould, C., Ferrier, K., and Payne, S. (2001). Preferences of patients for patient centred approach to consultation in primary care: observational study. BMJ *322*, 468–472. https://doi.org/10.1136/bmj.322.7284.468.

32. Reaney, M. (2014). The importance of understanding the impact of preference in clinical trials of diabetes interventions. J. Diabetes *6*, 496–498. https://doi.org/10.1111/1753-0407.12187.

33. Wan, P., Huang, Z., Tang, W., Nie, Y., Pei, D., Deng, S., Chen, J., Zhou, Y., Duan, H., Chen, Q., and Long, E. (2024). Outpatient reception via collaboration between nurses and a large language model: a randomized controlled trial. Nat. Med. *30*, 2878–2885. https://doi.org/10.1038/s41591-024-03148-7.

34. Chen, Q., Jin, J., and Yan, X. (2024). Impact of online physician service quality on patients' adoption behavior across different stages: An elaboration likelihood perspective. Decis. Support Syst. *176*, 114048. https://doi.org/10.1016/j.dss.2023.114048.

35. Zeng, D., Qin, Y., Sheng, B., and Wong, T.Y. (2025). DeepSeek's "Low-Cost" Adoption Across China's Hospital Systems: Too Fast, Too Soon? JAMA *333*, 1866–1869. https://doi.org/10.1001/jama.2025.6571.

36. Huo, B., Boyle, A., Marfo, N., Tangamornsuksan, W., Steen, J.P., McKechnie, T., Lee, Y., Mayol, J., Antoniou, S.A., Thirunavukarasu, A.J., et al. (2025). Large Language Models for Chatbot Health Advice Studies: A Systematic Review. JAMA Netw. Open *8*, e2457879. https://doi.org/10.1001/jamanetworkopen.2024.57879.

37. Bedi, S., Liu, Y., Orr-Ewing, L., Dash, D., Koyejo, S., Callahan, A., Fries, J.A., Wornow, M., Swaminathan, A., Lehmann, L.S., et al. (2025). Testing and Evaluation of Health Care Applications of Large Language Models: A Systematic Review. JAMA *333*, 319–328. https://doi.org/10.1001/jama.2024.21700.

38. Qian, B., Sheng, B., Chen, H., Wang, X., Li, T., Jin, Y., Guan, Z., Jiang, Z., Wu, Y., Wang, J., et al. (2024). A Competition for the Diagnosis of Myopic Maculopathy by Artificial Intelligence Algorithms. JAMA Ophthalmol. *142*, 1006–1015. https://doi.org/10.1001/jamaophthalmol.2024.3707.

39. Huang, S., Li, Z., Lin, B., Zhang, S., Yi, Q., and Wang, L. (2023). HPMI: A retinal fundus image dataset for identification of high and pathological myopia based on deep learning. https://doi.org/10.6084/m9.figshare.24800232.v2.

40. Guo, X., Xiao, O., Chen, Y., Wu, H., Chen, L., Morgan, I.G., and He, M. (2017). Three-Dimensional Eye Shape, Myopic Maculopathy, and Visual Acuity: The Zhongshan Ophthalmic Center–Brien Holden Vision Institute High Myopia Cohort Study. Ophthalmology *124*, 679–687. https://doi.org/10.1016/j.ophtha.2017.01.009.

41. Yokoi, T., and Ohno-Matsui, K. (2018). Diagnosis and Treatment of Myopic Maculopathy. Asia. Pac. J. Ophthalmol. *7*, 415–421. https://doi.org/10.22608/APO.2018290.

42. Ohno-Matsui, K., Kawasaki, R., Jonas, J.B., Cheung, C.M.G., Saw, S.-M., Verhoeven, V.J.M., Klaver, C.C.W., Moriyama, M., Shinohara, K., Kawasaki, Y., et al. (2015). International Photographic Classification and Grading System for Myopic Maculopathy. Am. J. Ophthalmol. *159*, 877–883.e7. https://doi.org/10.1016/j.ajo.2015.01.022.

43. Shi, D., Zhang, W., Yang, J., Huang, S., Chen, X., Xu, P., Jin, K., Lin, S., Wei, J., Yusufu, M., et al. (2025). A multimodal visual–language foundation model for computational ophthalmology. npj Digit. Med. *8*, 381. https://doi.org/10.1038/s41746-025-01772-2.

44. Wightman, R., Raw, N., Soare, A., Arora, A., Ha, C., Reich, C., Guan, F., Kaczmarzyk, J., MrT23, Mike, SeeFun, et al. (2023). Rwightman/Pytorch-Image-Models: v0.8.10dev0 Release (Zenodo), Version v0.8.10dev0. https://doi.org/10.5281/ZENODO.4414861.

45. Abrar, M., Sermet, Y., and Demir, I. (2025). An Empirical Evaluation of Large Language Models on Consumer Health Questions. BioMedInformatics *5*, 12. https://doi.org/10.3390/biomedinformatics5010012.

46. Workum, J.D., Volkers, B.W.S., Van De Sande, D., Arora, S., Goeijenbier, M., Gommers, D., and Van Genderen, M.E. (2025). Comparative evaluation and performance of large language models on expert level critical

care questions: a benchmark study. Crit. Care *29*, 72. https://doi.org/10.1186/s13054-025-05302-0.

47. Lam, W.W.T., Fielding, R., Chow, L., Chan, M., Leung, G.M., and Ho, E.Y.Y. (2005). Brief communication: The Chinese medical interview satisfaction scale-revised (C-MISS-R): Development and validation. Qual. Life Res. *14*, 1187–1192. https://doi.org/10.1007/s11136-004-2391-5.

48. Wolf, M.H., Putnam, S.M., James, S.A., and Stiles, W.B. (1978). The medical interview satisfaction scale: Development of a scale to measure patient perceptions of physician behavior. J. Behav. Med. *1*, 391–401. https://doi.org/10.1007/BF00846695.

49. O' Connor, A.M.; Ottawa Hospital Research Institute (1993). User manual – Decision Conflict Scale. https://decisionaid.ohri.ca/docs/develop/User_Manuals/UM_decisional_conflict.pdf.

CellPress
OPEN ACCESS

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
MMAC	Zenodo	https://doi.org/10.5281/zenodo.11025749
HPMI	Figshare	https://doi.org/10.6084/m9.figshare.24800232.v2
Software and algorithms		
Mistral Large: 123b	Mistral AI	https://mistral.ai/news/mistral-large-2407
ChatGPT	OpenAI	https://chatgpt.com/
R (Version 4.3.1)	R software	https://www.r-project.org/

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

The study was approved by the Institutional Review Board of the Hong Kong Polytechnic University (reference number HSEARS20240229009) and was conducted in accordance with the Declaration of Helsinki. The randomized controlled clinical trial was registered at ClinicalTrials.gov (NCT06607822, registration date 2024/09/11). 70 participants (27 male and 43 female) were randomly allocated to the intervention and control groups in a 1:1 ratio using simple random sampling. Sex and age were well-balanced between the two groups at baseline. Demographic details were provided in Table S4. Written informed consent was obtained from all participants.

METHOD DETAILS

Establishment of the image classification tool

For the image classification tool, we utilized two public datasets (Myopic Maculopathy Analysis Challenge (MMAC),[38] High or Pathological Myopia Image (HPMI)[39]), and a retrospective dataset approved by the ethics committee of Zhongshan Ophthalmic Center (2012KYNL002) for model development.[40] The dataset comprises fundus images from participants aged 7-70 years with bilateral high myopia ($\leq$ -6.00 D spherical power in both eyes), excluding those with secondary myopia causes, history of refractive surgeries or intraocular procedures, or any severe systemic conditions. The image classification model was designed to grade myopic maculopathy (MM), a leading cause of blindness and a critical focus of myopia screening.[41] MM classification followed the guidelines established by the META-PM Study Group.[42] MMAC was pre-labeled, while HPMI and our dataset were independently labeled by two ophthalmologists, each with five years of experience. In cases of disagreement, a senior ophthalmologist with 10 years of experience provided a final consensus. The image model was trained and validated on a total of 2,769 fundus images, comprising 1,391 images from MMAC, 789 from HPMI, and 589 from our dataset. Additional details about the image data were provided in Table S5.

The model architecture was based on ViT-large, a Vision Transformer variant with 24 Transformer layers. We initialized the model using pre-trained weights from EyeFound,[43] which is a multimodal foundation model pretrained on millions of multimodal ophthalmic data. Fine-tuning was performed on the composite dataset, with the data split into training, validation, and test sets in an 8:1:1 ratio. We used stratified sampling based on participant ID and class ID to prevent participant overlapping between splits while maintaining class balance across all splits. The model was trained over 50 epochs with a batch size of 64 and an input image size of 224×224 pixels. Auto-augmentation techniques were employed to increase data diversity.[44] The training process included a warmup phase to stabilize optimization and label smoothing to enhance generalization. The checkpoint with the highest Area Under the Receiver Operating Characteristic (AUROC) on the validation set was saved and subsequently used for final evaluation on the test set.

Establishment of the RAG-based knowledge tool

For the RAG-based knowledge tool, we developed a custom-built evidence-based knowledge database, the Myopia Knowledge Database (MKD). The MKD was constructed by integrating data from sources including medical books, peer-reviewed literature, clinical guidelines, and expert consensus. Textbooks on ophthalmology, optometry, neuro-ophthalmology, corneal diseases, glaucoma, lens diseases, and retinal diseases were incorporated into this dataset. Literature was systematically reviewed and selected based on its relevance to myopia management, focusing on pathogenesis, risk factors, clinical presentation, diagnosis, treatment, and management. Clinical practice guidelines from the American Academy of Ophthalmology and the Chinese Health Commission were also included. Expert consensus was gathered through discussions with a panel of Chinese ophthalmologists specializing in myopia. In total, the MKD comprised 12 ophthalmology textbooks and 61 clinical guidelines. Further details about the dataset were provided in Table S6.

The RAG process in ChatMyopia began by encoding input queries into dense vectors using an embedding model (bge-large-zh-v1.5 for Chinese or m3e-large for English). These vectors were matched against a pre-indexed knowledge base stored in FAISS (Facebook AI Similarity Search) using cosine similarity to retrieve the most relevant text chunks. Each chunk was capped at a CHUNK_SIZE of 250 tokens. The retrieved chunks were then embedded into the input prompt and processed by the LLM module to generate the final output.

Architecture of the ChatMyopia AI agent

To address both text-based and image-based inquiries, we developed the patient-centered ChatMyopia framework. Given that Mistral-123B primarily excels in handling complex medical questions,[45,46] the framework leveraged Mistral-123B to interpret questions, decompose tasks, and deliver personalized responses. It integrated two core components mentioned above: an image classification tool for myopic maculopathy grading and a RAG-based knowledge tool for up-to-date professional ophthalmology knowledge. The tool module was designed to be flexible and extendable, allowing for future enhancements by incorporating additional models. Furthermore, to ensure accessibility and ease of use, a simple, user-friendly interface was developed via a web-based front end. To facilitate more interactive and exploratory dialogue, we implemented a "Question Generation" prompt-engineering technique. This approach not only enabled the model to respond to patient inquiries but also suggested follow-up questions that patients may consider after receiving an answer, fostering a more dynamic and patient-centered interaction (Table S7).

QUANTIFICATION AND STATISTICAL ANALYSIS

Performance evaluation of ChatMyopia

ChatMyopia's performance was evaluated across three domains: image classification, single-choice questions (SCQ), and patient-centered free-form question answering.

For image classification, model performance was evaluated on the test set by accuracy, sensitivity, specificity, precision, AUROC, Area Under the Precision-Recall Curve (AUPRC), and F1 score.

For the SCQs, 150 myopia and optometry-related questions were sourced from preparation materials for the National Board Certification Examinations, National Health Professional Qualification Examinations, and National Health Talent Vocational Skills Training Examination in China. Three simulated exams, each containing 50 SCQs (39 knowledge-based and 11 scenario-based), were designed to evaluate the model's ability to handle standardized choice questions. Each question was input into ChatMyopia twice to obtain its scores. To compare ChatMyopia's performance with that of humans, we invited a group of experienced eye care practitioners (ECPs) to complete the same exams. We categorized the ECPs into two groups: five general ECPs (defined as ophthalmologists without a subspecialty focus) and two specialists (comprising ophthalmologists specializing in pediatric care and myopia management, as well as optometrists). Responses from both ChatMyopia and the human participants were scored on a 100-point scale.

For patient-centered question answering, ChatMyopia was tested on 85 open-ended questions gathered from popular online health consultation platforms (e.g., Good Doctor Online) and previously established LLM evaluation question lists.[7,9] These questions spanned topics including pathogenesis, risk factors, clinical presentation, diagnosis, treatment, prevention, and prognosis. The responses were compared with those from the general ECP and GPT-4, the commercial closed-source LLM. Each question was input into ChatMyopia and GPT-4 as a standalone query. The general ECP was required not to use online or external resources and could submit only one final answer per question in a single attempt without revision. All responses were systematically de-identified, reformatted into plain text to remove model-specific cues, and randomly shuffled before presentation to evaluators. Two blinded specialists independently evaluated the responses based on five criteria adapted from our previous study[14] and Luo et al.'s study[17]: accuracy, utility, relevance, safety, and harmlessness (Table S8). Each criterion was rated on a 3-point scale, and disagreements were resolved by consulting a third specialist. Evaluators were not informed in advance that both AI and human answers were included.

Randomized controlled trial for real-world validation

This clinical trial was conducted at the Hong Kong Polytechnic University Optometry Clinic from September 21 to October 26, 2024. The objective was to assess the utility and effectiveness of the ChatMyopia AI agent in improving patients' experience during medical consultations. We hypothesized that ChatMyopia, as a patient education tool, would provide high-quality information, improve disease self-awareness, facilitate positive interactions between patients and ECPs, and ultimately increase patient satisfaction in real-world clinical settings.

Eligible participants were patients aged 18 to 60 years seeking information related to myopia care at the pediatric or high myopia clinics, with no prior experience in digital medicine research, and who provided informed consent. Participants were randomly assigned to either the intervention or control group in a 1:1 ratio using simple random sampling. In the intervention group, participants engaged in a 10-minute interaction with ChatMyopia on a tablet device before meeting their ECPs. ChatMyopia does not provide final diagnosis or treatment decisions, and all clinical decisions remain under the supervision of ECPs. During this interaction, participants could ask questions related to risk factors, symptoms, diagnoses, examinations, treatments, advice, and the interpretation of their fundus photo. In the control group, participants received official leaflets from the Hong Kong Polytechnic University Optometry Clinic and read materials about children's vision care, myopia prevention, and high myopia management for 10 minutes. All participants then proceeded to a standard face-to-face consultation with their ECPs, who were trained to follow standardized communication scripts. During these consultations, the ECPs monitored ChatMyopia's responses and addressed patient questions that required further clarification.

The primary outcome was patient satisfaction, measured using the Chinese version of the Medical Interview Satisfaction Scale-Revised (C-MISS-R), a validated 10-item questionnaire adapted for the Hong Kong population.[47,48] Secondary outcomes included patients' perceptions assessed through a 7-aspect evaluation covering ease of understanding, accuracy in addressing concerns, empathy, improvement in understanding eye conditions, support for future treatments, effectiveness in communication with ECPs, and satisfaction with the provided information. Decision conflict was measured using the 10-item Decision Conflict Scale[49] in patients requiring myopia control treatment decisions (Table S9).

Statistical analysis

The sample size of the clinical trial was estimated based on differences in patient satisfaction between the ChatMyopia AI agent and the standard leaflet from our pilot study (n=10). 64 participants were required to achieve 95% power at a significance level of 0.05 ($\alpha = 0.05$, $\beta = 0.05$). SCQ scores were compared using Repeated Measures Analysis of Variance (RM-ANOVA) with a post hoc least significant difference test. The chi-square test was used to compare the scores between ChatMyopia and individual human in pair-wise ranking evaluation. Friedman test was used to detect the difference in manual evaluation, with pairwise comparisons performed using the Wilcoxon signed-rank test. Mann-Whitney U test was used to compare the C-MISS-R score, patients' perspectives scale, and decision conflict scale between ChatMyopia and leaflet groups. All statistical analyses were conducted using R (Version 4.3.1).