The ROYAL COLLEGE of OPHTHALMOLOGISTS

EDITORIAL

ChatGPT in ophthalmology: the dawn of a new era?

Eye (2024) 38:4–7; https://doi.org/10.1038/s41433-023-02619-4

"Those who can imagine anything, can create the impossible."

-Alan Turing

Imagine living in a world where patients can self-diagnose their medical condition based on self-reported symptoms and/or images; or doctors can manage any medical diseases without considerable prior domain-specific knowledge; or mundane medical administrative work such as drafting hospital discharge letters can be automated.

All above scenarios, once thought to be idealistic aspirations, are now gradually morphing into attainable realities made possible by artificial intelligence (AI), particularly with the recent advent of ChatGPT. In view of the significant hype and controversies, this Editorial aims to demystify the key concepts of ChatGPT and evaluate its roles in medicine, primarily from ophthalmology perspective.

WHAT IS CHATGPT?

ChatGPT (OpenAI, San Francisco, USA) is a generative AI, large language model (LLM)-powered chatbot [1]. It has the ability to process and convert textual and now graphic inputs to text-based outputs, engaging in human-like conversations and creating new human-like contents ("generative"). The model is based on Generative Pretrained Transformer (GPT) architecture and was pre-trained in an unsupervised manner using large-scale text corpora (containing billions of words) derived from books, articles, and various internet-based content, followed by minimal supervised fine-tuning of the model for various downstream tasks (known as "few-shot learning"). The transformer architecture enables processing of the entire input and context wholly and "transforms" it to meaningful human-like text based on statistical prediction of words (guided by previous words and entire context). OpenAI has also created a GPT-3 variant, DALL-E, which can convert textual inputs to graphic outputs. In view of its astounding versatility and functionality, ChatGPT has permeated all fields, encompassing business, entertainment, education, science, legal, and healthcare sectors [2–4] (https://tech.co/news/10-ways-businesses-using-chatgpt).

ChatGPT has undergone several iterations, refining from GPT-1 in 2018, GPT-2 in 2019 (1.5 billion training parameters) and GPT-3 in 2020 (175 billion training parameters), to the more recent ChatGPT-3.5 in November 2022 and ChatGPT-4 in March 2023 (though the latest architecture, training data and parameters remain undisclosed). ChatGPT has quickly accrued over one billion users and a PubMed search on 14/05/2023 returned >400 ChatGPT-related results, highlighting an immense interest in this technology within the medical field.

POTENTIAL PROMISES

ChatGPT has the potential of revolutionising medical field, including ophthalmology, in diverse ways, spanning patients, healthcare professionals/systems, research, and education/training (Fig. 1).

Patients

The plausibility of deploying ChatGPT for medical purposes largely stems from its superior question-answering performance in medical exams where it could attain scores paralleling candidates in the top 10% [1, 5, 6]. However, it is important not to over-extrapolate these findings as question-answering in medical examination does not directly translate to real-world clinical practice where interpersonal skills, clinical reasoning/contextualisation, and holistic patient care are essential. Several studies specifically evaluated ChatGPT's performance in answering ophthalmology questions and demonstrated fair-to-good accuracy (40–80%) [7–9]. This was influenced by question difficulty and domain (e.g. higher accuracy in general medicine than subspecialised domains). Interestingly, the model is receptive to training and refinement [8].

As such, non-medically trained individuals/patients may potentially utilise ChatGPT as a virtual assistant in self-triaging and self-diagnosing his/her ophthalmic conditions, ranging from innocuous to potentially sight-threatening ocular conditions (see examples in Fig. 2A, B). In addition, ChatGPT may be effective in generating patient educational materials, by translating jargon-heavy medical terminology into simplified and/or empathetic language tailored to the level of a layperson (Fig. 2C) [3, 10], or act as a "therapist" for counselling individuals with mental health illness [11]. This will be invaluable as patients with chronic eye diseases or visual impairment are more susceptible to psychological distress/illness [12–14]. ChatGPT's multilingual translation feature can further cater to multi-ethnic patient populations, and potentially be integrated with text-to-speech audio generation (particularly useful for visually impaired patients) and text-to-image or video generation platforms to enhance patient experience.

Healthcare professionals/systems

ChatGPT may empower physicians with curated, synthesised information and guidelines from otherwise diverse and disparate sources, facilitating the process of formulating differential diagnoses, decision-making, and treatment plans (Fig. 2D) [15]. Administrative tasks may potentially be automated and expedited using ChatGPT, including drafting of discharge summaries and medical reports, and answering administrative enquiries like appointment rescheduling and medication refill requests [16, 17]. ChatGPT can acknowledge and rectify its mistakes by prompting,

Received: 14 May 2023 Revised: 22 May 2023 Accepted: 2 June 2023
Published online: 27 June 2023

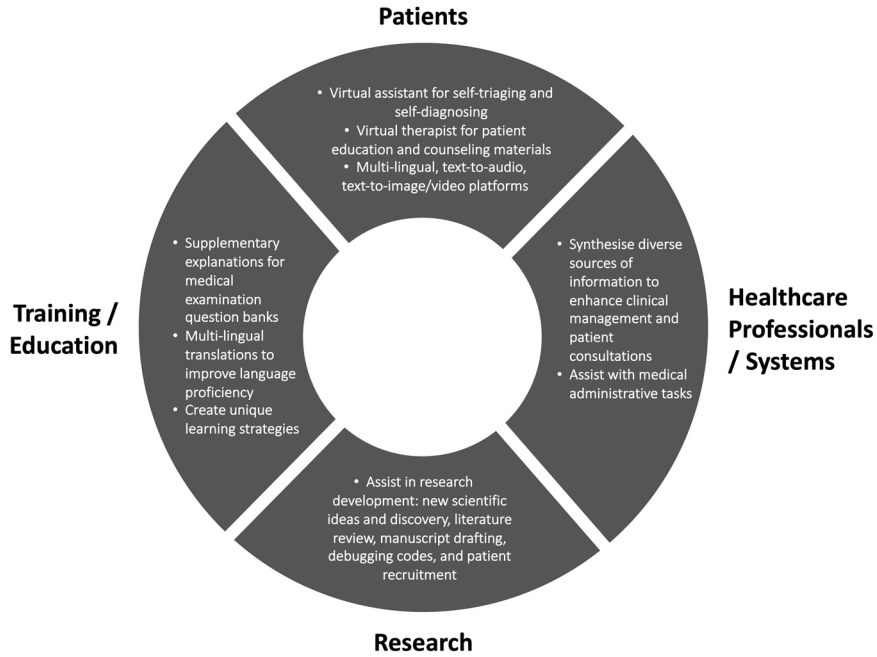

Fig. 1 Potential of ChatGPT in ophthalmology. Potential of ChatGPT in ophthalmology, spanning patients, healthcare professionals/organisations, research, education, and training.

highlighting the importance of human oversight to maximise the benefits of ChatGPT as an assistive tool [18]. It is capable of processing vast structured and unstructured clinical information stored within the electronic health records (EHR), which are increasingly adopted in ophthalmology [19, 20], (https://www.nodaudit.org.uk/) and generate concise summaries of patient record [21]. This allows the physicians to focus their consultation on clinical management and patient communication.

Research

Similarly, ChatGPT's performance in summarisation tasks can accelerate time-consuming research process. ChatGPT can efficiently perform literature review, summarise findings, and draft abstracts/manuscripts, with quality of work being indistinguishable from clinicians/scientists [22, 23]. There are ethical debates on whether ChatGPT should be recognised as an author in research articles. Based on the ICJME criteria, many publishers, including Springer Nature, stated that any LLM, including ChatGPT, should not be included as an author in view of the lack of ability to approve the final manuscript or to take accountability of the research [24, 25]. Nevertheless, ChatGPT's assistance should be acknowledged in publication where appropriate. Further, ChatGPT can also facilitate original research along the model development pipeline, by generating new scientific ideas, cleaning large datasets, debugging codes, assisting patient recruitment, and advancing basic science discovery [18, 26].

Education/training

ChatGPT's encouraging performance in medical examination question-answering tasks and ability to "interact" with human show potential in enhancing medical education and training [4]. In addition to generating answers, ChatGPT can further justify the responses with supplementary logical explanations of high concordance and insight [27–29]. The multilingual function of ChatGPT can be utilised to improve the learning experience of students with language barrier, by using ChatGPT's response as a benchmark and tool to enhance their language proficiency and writing styles [4]. It can also generate ideas, strategies and exam questions or quizzes to stimulate learning (Fig. 2E).

POTENTIAL PITFALLS

As a generative AI, ChatGPT generates outputs based on statistical prediction of the text without human-like reasoning. This can potentially result in plausible-sounding but inaccurate responses, known as "hallucination" or "fabrication" [30]. Lack of recency is another limitation as the model was trained with data up to September 2021. ChatGPT would have missed recent advancements in medicine and ophthalmology (which is rapidly evolving) (https://investors.apellis.com/news-releases/news-release-details/fda-approves-syfovretm-pegcetacoplan-injection-first-and-only). The model was trained using various resources, including unverified internet-based contents, which can exacerbate inaccuracies and lead to patient harm if prematurely deployed to clinical practice without rigorous validation.

Current ChatGPT model lacks transparency and explainability as the responses are not substantiated by references (which renders fact-checking impossible) and details on training dataset/system are undisclosed. Bias may be inadvertently introduced through biased training data, which may lead to biased responses. Other ethical concerns are data privacy and security where personal sensitive data may be inadvertently breached, (https://www.chargedretail.co.uk/2023/05/04/chatgpt-data-breach/) calling for stricter regulatory oversight and cybersecurity. Patients or doctors may upload eye/facial photos for ChatGPT's "medical assistance", which may risk breaching biometric security and/or identity theft [31]. Also, ChatGPT cannot take accountability nor abide by any ethical/moral boundaries. Therefore, doctors and/or patients should only consider ChatGPT as an assistive tool (at best). Potential wide-ranging roles of ChatGPT in research have also raised questions about research work originality (and plagiarism), crediting, and intellectual properties if such technology is used [18].

ChatGPT represents a significant breakthrough in AI and LLM. However, its readiness to be deployed as an automatic decision-making tool in medical field remains to be elucidated [5, 6, 27]. With further model refinement, ChatGPT may reach sufficient maturity, robustness and transparency to be embraced by relevant end users, including patients, healthcare professionals, healthcare systems, researchers, educationalists, regulatory bodies, and

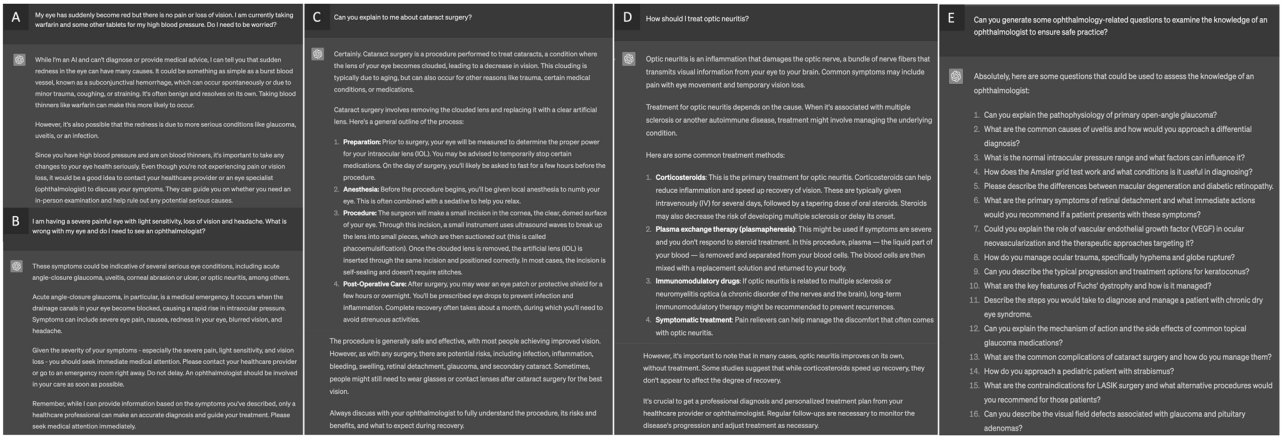

Fig. 2 Several examples of real-life conversation with ChatGPT in relation to the field of ophthalmology (some parts of the generated responses are shortened). A, **B** An example of an individual "consulting" ChatGPT for potential subconjunctival haemorrhage and acute-angle closure glaucoma, respectively. **C** An example of an individual "consulting" ChatGPT to understand about cataract surgery. **D** An example of an individual/ophthalmologist "consulting" ChatGPT for the management of optic neuritis. Note that for any medical advice provided by ChatGPT, it is always concluded with a statement that "consultation with a healthcare professional is crucial". **E** An example of an individual/ophthalmologist using ChatGPT for medical education and training purposes.

governmental stakeholders. When the time comes, this may well herald the dawn of a new era in medicine.

Darren Shu Jeng Ting [1,2,3], Ting Fang Tan[4,5] and Daniel Shu Wei Ting [4,5,6]✉
[1]*Birmingham and Midland Eye Centre, Birmingham, UK.* [2]*Academic Unit of Ophthalmology, Institute of Inflammation and Ageing, University of Birmingham, Birmingham, UK.* [3]*Academic Ophthalmology, School of Medicine, University of Nottingham, Nottingham, UK.* [4]*Artificial Intelligence and Digital Innovation Research Group, Singapore Eye Research Institute, Singapore, Singapore.* [5]*Singapore National Eye Centre, Singapore, Singapore.* [6]*Department of Ophthalmology and Visual Sciences, Duke-National University of Singapore Medical School, Singapore, Singapore.*
✉*email: daniel.ting@duke-nus.edu.sg*

REFERENCES

1. OpenAI. GPT-4 Technical Report. arXiv. 2023:2303.08774.
2. Stokel-Walker C, Van Noorden R. What ChatGPT and generative AI mean for science. Nature. 2023;614:214–6.
3. Will ChatGPT transform healthcare? Nat Med. 2023;29:505–6.
4. Seetharaman R. Revolutionizing medical education: Can ChatGPT boost subjective learning and expression? J Med Syst. 2023;47:61.
5. Gilson A, Safranek CW, Huang T, Socrates V, Chi L, Taylor RA, et al. How does ChatGPT Perform on the United States medical licensing examination? The implications of large language models for medical education and knowledge assessment. JMIR Med Educ. 2023;9:e45312.
6. Thirunavukarasu AJ, Hassan R, Mahmood S, Sanghera R, Barzangi K, El Mukashfi M, et al. Trialling a large language model (ChatGPT) in general practice with the applied knowledge test: observational study demonstrating opportunities and limitations in primary care. JMIR Med Educ. 2023;9:e46599.
7. Mihalache A, Popovic MM, Muni RH. Performance of an artificial intelligence chatbot in ophthalmic knowledge assessment. JAMA Ophthalmol. 2023;e231144.
8. Antaki F, Touma S, Milad D, El-Khoury J, Duval R. Evaluating the Performance of ChatGPT in ophthalmology: an analysis of its successes and shortcomings. Ophthal Sci. 2023; https://doi.org/10.1016/j.xops.2023.100324.
9. Raimondi R, Tzoumas N, Salisbury T, Di Simplicio S, Romano MR. Comparative analysis of large language models in the royal college of ophthalmologists fellowship exams. Eye (Lond). 2023; https://doi.org/10.1038/s41433-023-02563-3.
10. Sharma A, Lin IW, Miner AS, Atkins DC, Althoff T. Human–AI collaboration enables more empathic conversations in text-based peer-to-peer mental health support. Nat Mach Intell. 2023;5:46–57.
11. Graber-Stiehl I. Is the world ready for ChatGPT therapists? Nature. 2023;617:22–4.
12. Demmin DL, Silverstein SM. Visual impairment and mental health: unmet needs and treatment options. Clin Ophthalmol. 2020;14:4229–51.
13. Ting DSJ, Krause S, Said DG, Dua HS. Psychosocial impact of COVID-19 pandemic lockdown on people living with eye diseases in the UK. Eye (Lond). 2021;35:2064–6.
14. Zhou Y, Murrough J, Yu Y, Roy N, Sayegh R, Asbell P, et al. Association between depression and severity of dry eye symptoms, signs, and inflammatory markers in the DREAM study. JAMA Ophthalmol. 2022;140:392–9.
15. Hirosawa T, Harada Y, Yokose M, Sakamoto T, Kawamura R, Shimizu T. Diagnostic accuracy of differential-diagnosis lists generated by generative pretrained transformer 3 chatbot for clinical vignettes with common chief complaints: a pilot study. Int J Environ Res Public Health. 2023;20:3378.
16. Patel SB, Lam K. ChatGPT: the future of discharge summaries? Lancet Digit Health. 2023;5:e107–e8.
17. Singh S, Djalilian A, Ali MJ. ChatGPT and ophthalmology: exploring its potential with discharge summaries and operative notes. Semin Ophthalmol. 2023:1–5.
18. van Dis EAM, Bollen J, Zuidema W, van Rooij R, Bockting CL. ChatGPT: five priorities for research. Nature. 2023;614:224–6.
19. Ting DSJ, Deshmukh R, Ting DSW, Ang M. Big data in corneal diseases and cataract: current applications and future directions. Front Big Data. 2023;6:1017420.
20. Lim MC, Boland MV, McCannel CA, Saini A, Chiang MF, Epley KD, et al. Adoption of electronic health records and perceptions of financial and clinical outcomes among ophthalmologists in the United States. JAMA Ophthalmol. 2018;136:164–70.
21. Yang X, Chen A, PourNejatian N, Shin HC, Smith KE, Parisien C, et al. A large language model for electronic health records. NPJ Digit Med. 2022;5:194.
22. Hutson M. Could AI help you to write your next paper? Nature. 2022;611:192–3.
23. Else H. Abstracts written by ChatGPT fool scientists. Nature. 2023;613:423.
24. Tools such as ChatGPT threaten transparent science; here are our ground rules for their use. Nature. 2023;613:612.
25. Stokel-Walker C. ChatGPT listed as author on research papers: many scientists disapprove. Nature. 2023;613:620–1.
26. Madani A, Krause B, Greene ER, Subramanian S, Mohr BP, Holton JM, et al. Large language models generate functional protein sequences across diverse families. Nat Biotechnol. 2023; https://doi.org/10.1038/s41587-022-01618-2.
27. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. PLOS Digit Health. 2023;2:e0000198.
28. Wei J, Wang X, Schuurmans D, Bosma M, Ichter B, Xia F, et al. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. NeurIPS 2022 Conference. 2022.
29. Yang X, Chen A, PourNejatian N, Shin HC, Smith KE, Parisien C, et al. GatorTron: A Large Clinical Language Model to Unlock Patient Information from Unstructured Electronic Health Records. arXiv. 2022; https://doi.org/10.48550/arXiv.2203.03540.
30. Azamfirei R, Kudchadkar SR, Fackler J. Large language models and the perils of their hallucinations. Crit Care. 2023;27:120.
31. Wang M, Qin Y, Liu J, Li W. Identifying personal physiological data risks to the Internet of everything: the case of facial data breach risks. Humanities and social sciences. Communications. 2023;10:216.

AUTHOR CONTRIBUTIONS

Study conceptualisation and design: DSJT, DSWT. Data collection and interpretation: DSJT, TFT. Drafting of initial manuscript: DSJT, TFT. Critical revision of manuscript: DSWT. Final approval of manuscript: DSJT, TFT, DSWT.

FUNDING

DSJT acknowledges support from the Medical Research Council/Fight for Sight Clinical Research Fellowship (MR/T001674/1). DSWT is supported by the National Medical Research Council, Singapore (NMCR/HSRG/0087/2018; MOH-000655-00; MOH-001014-00), Duke-NUS Medical School (Duke-NUS/RSF/2021/0018; 05/FY2020/EX/15-A58), and Agency for Science, Technology and Research (A20H4g2141; H20C6a0032). These funders were not involved in the conception, execution, or reporting of this review.

COMPETING INTERESTS

DSWT holds a patent on a deep learning system for the detection of retinal diseases. The other authors declare no competing interests.