Review

Opportunities and Challenges of Large Language Models in the Field of Ophthalmology

Shi-Nan Wu[1]*, Xiao-Yu Wang[2]*, Yan-Mei Zeng[2]*, Xiang-Yi Liu[2], Jin-Yu Hu[2], Qian Ling[2], Jie Zou[2], Cheng Chen[2], Liang-Qi He[2], Hong Wei[2], Yi Shao[3#]

[1]Eye Institute of Xiamen University, School of Medicine, Xiamen University, Xiamen, Fujian 361000, China

[2]Department of Ophthalmology, The First Affiliated Hospital of Nanchang University, Nanchang, Jiangxi 330006, China

[3]Department of Ophthalmology, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, National Clinical Research Center for Eye Diseases, Shanghai 200080, China

*Equal contribution

[#]**Correspondence to:** Yi Shao, **email:** freebee99@163.com;

Keywords: Artificial intelligence; Large language models; Ophthalmology; Clinical applications; Challenge.

Received: November 3, 2024 **Accepted:** February 2, 2025 **Published:** February 21, 2025

ABSTRACT

In recent years, ophthalmology, as an image-intensive discipline, has accumulated vast amounts of heterogeneous clinical data, including electronic medical records, imaging data, genomic information, and patient-reported outcomes. These datasets are characterized by multimodality, high dimensionality, and strong structural features, providing fertile ground for the application of large language models (LLMs). LLMs hold great potential in ophthalmology, with applications ranging from automated medical documentation and patient communication support to scientific literature summarization, intelligent diagnostic assistance, decision support systems, and personalized health education. Their capabilities in natural language understanding and cross-modal reasoning may significantly enhance both clinical and research efficiency in the field. However, the implementation of LLMs in ophthalmology faces multiple challenges, such as limited domain-specific corpora, concerns over data privacy, the risk of hallucinations and misinformation, insufficient cross-modal integration, and regulatory and ethical barriers to clinical deployment. Future progress will require interdisciplinary collaboration to optimize model performance, standardize data, and establish robust validation systems to ensure safe, explainable, and effective use in real-world scenarios.

INTRODUCTION

Over the past decade, the field of artificial intelligence has witnessed tremendous progress, particularly in the development of large language models (LLMs). Early natural language processing systems, which relied on rule-based algorithms and statistical methods, gradually gave way to deep learning techniques that could capture semantic nuances and complex linguistic patterns(Huang, Gao, Tao, Du, & Lu, 2022; Manning, 2015). This evolution accelerated with the advent of transformer architectures, which fundamentally changed how models process and generate language by leveraging self-attention mechanisms(Soydaner, 2022). These architectures enabled LLMs to scale up in size and complexity, leading to breakthroughs in machine translation, sentiment analysis, and text generation. With models such as Generative Pre-trained Transformer (GPT), Bidirectional Encoder Representations from Transformers (BERT)(Devlin, Chang, Lee, & Toutanova, 2019; Shreyashree, Sunagar, Rajarajeswari, & Kanavalli, 2022), and their successors, the capacity for understanding context, managing multi-turn conversations, and synthesizing vast amounts of information has grown exponentially(Korngiebel & Mooney, 2021). This rapid evolution not only highlights the innovative strides made in computational linguistics but also underscores the shift towards more generalized AI systems capable of tackling diverse, real-world challenges. As LLMs continue to evolve, they increasingly serve as the backbone for applications across numerous domains, setting the stage for their integration into specialized fields such as medicine and ophthalmology.

In the realm of medicine, LLMs have begun to transform how healthcare professionals interact with data, research, and patient care. Initially deployed for tasks like automated medical transcription and summarization of clinical notes, LLMs have expanded their utility to include diagnostic assistance, treatment recommendations, and even the interpretation of complex genomic data(Thirunavukarasu et al., 2023). Their ability to process and synthesize information from diverse sources—ranging from electronic health records to cutting-edge biomedical literature—allows clinicians to access consolidated, up-to-date insights that can inform patient care decisions(X. Yang et al., 2022). One of the most significant advantages of using LLMs in medicine is their potential to reduce the cognitive burden on clinicians by filtering through vast amounts of unstructured data and presenting actionable information in a concise format. Moreover, these models facilitate enhanced interdisciplinary communication by translating complex medical terminology into accessible language for patients, thereby supporting improved patient education and engagement. Another key benefit lies in their ability to identify hidden patterns and correlations in large datasets, which can lead to earlier detection of diseases and personalized treatment strategies. As these models continue to mature, their integration into medical workflows promises not only to boost efficiency but also to foster innovation in diagnostic methodologies and therapeutic interventions, thereby reshaping the landscape of modern healthcare.

Ophthalmology, with its reliance on high-resolution imaging, detailed clinical records, and nuanced patient narratives, stands to benefit uniquely from the application of LLMs (Betzler et al., 2023). The specialty inherently involves a high degree of data heterogeneity—from retinal scans and optical coherence tomography (OCT) images to structured patient histories and subjective symptom descriptions. LLMs can integrate these diverse data streams by correlating textual information with imaging findings, thus providing a more holistic understanding of ocular health. Furthermore, by automating the extraction and interpretation of data from electronic health records and imaging reports, LLMs can significantly reduce administrative burdens and allow clinicians to devote more time to direct patient care. The models' proficiency in natural language processing also opens new avenues for patient communication, enabling the development of intelligent systems that provide personalized health education and support adherence to treatment plans(Rojas-Carabali et al., 2024; Z. Yang et al., 2024). In addition, LLM-driven decision support systems can assist ophthalmologists in evaluating treatment outcomes and predicting disease progression, thereby improving the overall quality of care(Rong et al., 2024; Sun et al., 2023; S. N. Wu et al., 2024). Ultimately, the integration of LLMs in ophthalmology promises not only to enhance clinical efficiency and diagnostic precision but also to contribute to a deeper, data-driven understanding of ocular diseases, paving the way for more innovative, patient-centered care solutions. This paper summarizes the future applications of LLMs in ophthalmology from several perspectives, including ophthalmic image analysis, diagnosis and treatment prescriptions for eye diseases, medical supply supervision, and surgical treatment of eye conditions (Table 1).

Table 1: Potential Applications of LLMs in the Field of Ophthalmology

Application Areas	Description
Ophthalmic Image Analysis	LLMs can be used for automated analysis of ophthalmic images, identifying lesions in fundus photographs, detecting retinal diseases, and providing preliminary screening and diagnostic support.
Diagnosis and Treatment Prescriptions for Ophthalmic Diseases	Given a patient's medical history and symptoms, LLMs can generate diagnoses and treatment prescriptions for ophthalmologists, providing customized medical recommendations.
Medical Supply Monitoring	LLMs can be used to supervise the production, quality control, and supply chain management of medical supplies, ensuring that pharmaceuticals and instruments in the field of ophthalmology meet standards.
Surgical Treatment of Eye Diseases	LLMs can provide preoperative planning, surgical technique descriptions, and postoperative care recommendations for ophthalmic surgeries, helping doctors improve the precision and safety of the procedures.

APPLICATIONS OF LLMS IN OPHTHALMOLOGY

The Role of LLMs in Ophthalmic Image Analysis

Ophthalmology has a unique advantage in image-based diagnosis due to the rich diagnostic images available in clinical outpatient settings. Compared to other clinical specialties, ophthalmology excels in imaging technologies and image-based diagnostics. Common ophthalmic images include color fundus photographs, OCT images, corneal topography, and more. Currently, AI technology is widely applied in ophthalmology for lesion detection, lesion segmentation, disease classification, and quantitative measurement of ophthalmic parameters. As a powerful AI tool, LLMs can assist in screening eye diseases, supporting diagnosis, and optimizing treatment plans. A team has developed a large model based on fundus fluorescein angiography (FFA) images (FFA-GPT) that can answer visual questions related to the images (Chen, Zhang, et al., 2024). With the continuous development of GPT, GPT-4 includes a visual question-answering system for various ophthalmic imaging modalities (slit-lamp, FFA, ocular ultrasound, etc.), showing good accuracy in examination identification, lesion identification, usability, safety, and diagnostic capacity (Xu, Chen, Zhao, & Shi, 2023). However, as of now, LLMs for ophthalmic image recognition remain limited, with the majority of training data sourced from online datasets, and there is still a lack of specialized training on professional ophthalmic images. There is a deficiency in specific ophthalmic terminology and result descriptions, and it is still far from achieving differential diagnosis of various eye diseases. Several strategies can be employed to address these limitations: First, enhancing the capability of LLMs to refine medical-specific terminology, rather than generating general text, can reduce the occurrence of AI hallucinations (Z. Yang et al., 2024). Second, precise prompt engineering is crucial to producing expected results, including aspects such as prompt length, style, and complexity. In text mining applications in natural language processing, providing a logical framework to obtain specific results, rather than attempting to generate comprehensive answers immediately, will significantly improve model performance (Chen, Zhao, et al., 2024; Z. Zhao et al., 2024; Zheng et al., 2024). Therefore, timely customization of prompts for different medical tasks is necessary. Additionally, deep learning models, including LLMs, often lack interpretability and transparency (X. Wu et al., 2024; H. Zhao et al., 2024). These "black-box" models typically cannot provide a clear decision-making process, raising concerns about trust and accountability in medical environments. Clinicians and patients need to have confidence in the accuracy and reasoning behind AI-based recommendations, which remains a major barrier to widespread adoption. Another challenge is the lack of large-scale, diverse, and high-quality annotated datasets for training LLMs (Omiye, Gui, Rezaei, Zou, & Daneshjou, 2024; X. Yang et al., 2022). Ophthalmic images inherently vary, including differences in imaging equipment, patient populations, and disease manifestations, making it challenging to develop models that generalize across different groups. The scarcity of annotated datasets further complicates the training process. Finally, integrating LLMs into existing medical systems is also a significant challenge(Betzler et al., 2023). Ophthalmic image analysis often requires seamless collaboration between multiple technologies, including image processing algorithms and clinical decision support systems. Ensuring that LLMs can effectively integrate with these systems while maintaining high standards of data privacy and security is crucial for their practical application.

In conclusion, although LLMs have immense potential and significant advantages in ophthalmic image analysis, overcoming challenges such as interpretability, data quality, and system integration is key to achieving their full application in clinical ophthalmology.

The Role of Large Language Models in Ophthalmic Disease Diagnosis, Treatment Prescription, and Medical Supplies Monitoring

The diagnosis and treatment prescriptions for ophthalmic diseases are scientifically, reasonably, and systematically formulated based on the ophthalmic examination data and demographic background data of outpatient patients, with the goal of improving visual quality. Generally, ophthalmic disease treatment prescriptions are tailored to the real-time condition of the patient. Since GPT-4 has been trained with a rich set of online clinical information and relevant literature guidelines, it can provide relatively comprehensive auxiliary diagnostic and treatment advice for general patients. However, it is important to emphasize that GPT-4 typically advises patients to ultimately consult a clinical physician at the end of its responses. This generative language model only provides diagnostic suggestions and cannot make a definitive gold standard diagnosis. Although LLMs have demonstrated broad application prospects in healthcare management and clinical decision support, clinical physicians and patients still need to be cautious in their selection to mitigate the risks of over-relying on biased suggestions generated by these models, which could lead to potential misdiagnoses. Furthermore, for diseases requiring real-time tracking of patient conditions, GPT-4 can serve as an auxiliary tool, offering timely, accurate, and comprehensive assistance to clinical physicians in taking appropriate measures, thus reducing the risk of harm caused by improper medication and maximizing the therapeutic benefits of the treatment. Additionally, some studies have reported that GPT-4 can act as a virtual physician providing real-time guidance for patients, which is useful for those who lack access to professional guidance(Cheng, Guo, et al., 2023). One study evaluated the accuracy of GPT in providing information for common retinal disease patients and found that LLMs are highly accurate in general disease information, but their accuracy in potential treatment recommendations still needs further improvement(Potapenko et al., 2023). Another study reported that ChatGPT can assist in medical education and health knowledge promotion in rural and economically disadvantaged areas, offering background information on trachoma prevention, symptom identification, and treatment guidelines(Pushpanathan et al., 2023). Pushpanathan et al., by comparing GPT-3.5, GPT-4.0, and Google Bard, emphasized the potential of GPT-4.0 in providing accurate and comprehensive responses to inquiries about common ocular symptoms(Masalkhi et al., 2024). Therefore, ChatGPT has broad potential applications in ophthalmic disease diagnosis and treatment as well as in the supervision of pharmaceutical use. By training with large amounts of ophthalmic images, clinical ophthalmic data, and professional content, the effectiveness of LLMs in ophthalmology can be further optimized.

The Role of LLMs in Ophthalmic Surgical Treatment

In the context of ophthalmic surgical treatment, LLMs can enhance both preoperative planning and intraoperative assistance. The integration of LLMs into ophthalmic surgery offers several advantages, although it is also accompanied by specific challenges and limitations that must be addressed to optimize their utility. LLMs can play a significant role in preoperative planning by synthesizing vast amounts of patient data, including medical history, imaging results, and clinical notes(Cheng, Sun, He, Gu, & Wu, 2023). This ability allows LLMs to assist surgeons in identifying potential risks and complications, tailoring treatment plans, and ensuring that all relevant information is considered before proceeding with surgery(Chung et al., 2024). For instance, in complex surgeries such as retinal detachment repair or corneal transplant, LLMs can help identify patterns or correlations in patient data that may not be immediately apparent to clinicians. Intraoperatively, LLMs have the potential to assist in real-time decision-making and surgical navigation. They can analyze ongoing data from various diagnostic devices, such as OCT or intraoperative cameras, and provide suggestions or alerts regarding surgical techniques, instruments, or adjustments. For example, during cataract surgery, LLMs could help in selecting the most appropriate intraocular lens (IOL) based on real-time data analysis, thus improving the precision of the procedure and reducing the likelihood of postoperative complications. Furthermore, LLMs can be used to train surgical teams by providing simulated procedural scenarios and feedback, contributing to continuous learning and skill enhancement. Despite these promising applications, the integration of LLMs into ophthalmic surgery faces several challenges. One of the primary concerns is the reliability and accuracy of LLM-driven recommendations(Miller, 2023). While LLMs can process vast amounts of information, they are not infallible and may generate inaccurate or irrelevant suggestions. This could pose significant risks in high-stakes surgical environments, where precision is critical. Additionally, LLMs are only as good as the data they are trained on, and if the underlying data is incomplete, biased, or unrepresentative, the model's performance may be compromised. Another challenge is the need for seamless integration of LLMs into existing surgical workflows. Ophthalmic surgery requires careful coordination among various specialists, instruments, and technologies, and introducing LLMs into these workflows could disrupt established practices. The learning curve for surgical teams to effectively use LLMs and the potential for human resistance to new technologies could also hinder their widespread adoption. Moreover, the lack of transparency in some LLM decision-making processes presents a barrier to trust and accountability. Surgeons must be able to understand the rationale behind LLM-generated recommendations, particularly when they are involved in critical surgical decisions. Ensuring that LLMs provide explainable and interpretable results is essential for their integration into clinical practice. Lastly, data privacy and security are significant concerns when using LLMs in medical settings. The processing of sensitive patient information must comply with regulations such as the Health Insurance Portability and Accountability Act (HIPAA) in the United States or the General Data Protection Regulation (GDPR) in Europe(Feretzakis et al., 2025; Savage, Wang, & Shieh, 2023). Ensuring that LLMs are used in a way that protects patient confidentiality is paramount.

In conclusion, while LLMs hold substantial promise for advancing ophthalmic surgical treatment, their integration into practice requires overcoming challenges related to accuracy, workflow integration, transparency, and data security. Addressing these issues will be crucial for ensuring that LLMs contribute positively to the field of ophthalmology and improve patient outcomes.

The Role of LLMs in Ophthalmic Medical Education

LLMs, such as ChatGPT and Med-PaLM, are increasingly being explored as tools to enhance ophthalmic medical education. Their ability to generate human-like text, synthesize vast amounts of information, and provide instant responses positions them as promising educational assistants for medical students, residents, and even practicing ophthalmologists(G. Wu et al., 2024). In ophthalmology, LLMs have demonstrated utility in several educational scenarios. First, they serve as interactive tutors capable of explaining complex topics such as ocular embryology, retinal circuitry, or the pathophysiology of glaucoma in a simplified yet accurate manner. Second, LLMs can aid learners in clinical decision-making by simulating patient scenarios, providing differential diagnoses, and discussing the rationale behind treatment strategies, much like "EyeTeacher." (Sevgi, Antaki, & Keane, 2024) This capacity for dynamic, on-demand learning makes them particularly valuable in problem-based learning environments and for self-directed education. Moreover, LLMs have been employed in assisting with board exam preparation, generating multiple-choice questions, and summarizing ophthalmology literature, all of which help streamline and personalize the learning process(Cai et al., 2023). Despite their potential, LLMs face several notable challenges in ophthalmic medical education. One major concern is the accuracy and reliability of generated content. Although LLMs are trained on large-scale biomedical corpora, they can occasionally produce incorrect or outdated information, which may mislead trainees if not properly vetted. This issue is particularly critical in a specialty like ophthalmology, where nuanced clinical distinctions can significantly impact patient outcomes. Another limitation is the models' lack of true clinical understanding and reasoning. While LLMs can mimic expert language, they do not possess real-world clinical experience or visual diagnostic capability—skills that are essential in ophthalmology. The integration of LLMs with image-based AI tools is an ongoing research area, but current standalone LLMs cannot interpret fundus photographs, OCT scans, or slit-lamp images, which limits their role in comprehensive ophthalmic education. Additionally, concerns about bias, data privacy, and the lack of transparency in model training further hinder their broader adoption in medical curricula. Educators also face challenges in integrating LLMs into existing teaching frameworks, as the models require critical oversight and validation by experienced clinicians to ensure educational integrity. In summary, LLMs hold significant promise for enhancing ophthalmic education through interactive learning and knowledge synthesis. However, their implementation must be approached cautiously, with safeguards to ensure accuracy and alignment with clinical best practices. Future advancements that combine LLMs with multimodal inputs and domain-specific fine-tuning may further enhance their educational value in ophthalmology.

LIMITATIONS OF LLMS IN OPHTHALMIC APPLICATIONS

Currently, the application of LLMs in ophthalmology faces several notable limitations. First, LLMs lack interpretability—the underlying analytical mechanisms and learning processes driving their responses are often opaque, making it difficult for clinicians to fully trust or validate their outputs. Second, there is a lack of standardization in training datasets. Many studies assessing LLM performance in ophthalmology use inconsistent criteria, and real-world clinical practice relies heavily on image-based diagnosis, an area where LLMs have yet to demonstrate strong capabilities due to insufficient specialized training and limited access to validated datasets. While datasets for common eye diseases exist, many conditions still lack officially recognized data, and models like ChatGPT, although rich in general internet knowledge, have not undergone focused ophthalmic training. Third, regulatory hurdles remain. To guide clinical decision-making, AI products must comply with medical device regulations in each country, ensuring safety, reliability, and clarity regarding legal responsibility. Fourth, there is a risk of generating incorrect or misleading medical advice. Without proper oversight, users may place undue trust in LLM outputs, potentially leading to harmful health decisions and adverse outcomes. Finally, LLMs may impact medical education by encouraging overreliance and reducing critical thinking. If medical students depend too heavily on AI-generated content, it could hinder the development of essential diagnostic reasoning and practical skills, while fostering an environment that tolerates automation and plagiarism. Addressing these limitations is crucial for the safe and effective integration of LLMs into ophthalmic care.

CONCLUSION

LLMs have made significant progress in the diagnosis, screening, and treatment of ophthalmic diseases and are currently developing at a rapid pace. To date, AI-based screening systems, including those using deep learning, have already been applied in clinical practice and have demonstrated high sensitivity and specificity. However, we are also facing substantial challenges in ophthalmic care due to population growth, aging demographics, the increasing global burden of blindness and visual impairment, and the shortage of ophthalmic medical resources (Li et al., 2024; Zhang, Chen, Liao, & Ding, 2025). To some extent, understanding the potential of emerging tools and seizing the opportunity for enhanced interaction between AI and patient information may be key to reducing the cost of comprehensive eye care. Moreover, given China's vast repository of medical data, the standardization of LLM-based screening and diagnostic systems, such as those built on ChatGPT, could significantly enhance their application prospects in ophthalmology. The emergence and popularization of generative language models have brought great convenience to clinical practice, medical education, and scientific research. However, it is important to acknowledge that such technologies also have limitations. AI cannot fully replace clinicians in the diagnostic and treatment process. Skilled physicians are still essential in clinical settings, and a qualified doctor cannot rely solely on AI systems for patient care. Looking ahead, it is hoped that collaboration between AI systems like GPT and clinical practitioners will help deliver earlier diagnosis and treatment to more patients with ocular diseases, ultimately improving their quality of life. At the same time, healthcare professionals should take on the responsibility of promoting and disseminating these technologies, ensuring that patients in all regions can benefit from them.

REFERENCE

1. Huang D, Gao F, Tao X, Du Q, Lu J: Toward semantic communications: Deep learning-based image semantic coding. IEEE Journal on Selected Areas in Communications 2022, 41(1):55-71.
2. Manning CD: Computational linguistics and deep learning. Computational Linguistics 2015, 41(4):701-707.
3. Soydaner D: Attention mechanism in neural networks: where it comes and where it goes. Neural Computing and Applications 2022, 34(16):13371-13385.
4. Devlin J, Chang M-W, Lee K, Toutanova K: Bert: Pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers): 2019; 2019: 4171-4186.
5. Shreyashree S, Sunagar P, Rajarajeswari S, Kanavalli A: A literature review on bidirectional encoder representations from transformers. Inventive Computation and Information Technologies: Proceedings of ICICIT 2021 2022:305-320.
6. Korngiebel DM, Mooney SD: Considering the possibilities and pitfalls of Generative Pre-trained Transformer 3 (GPT-3) in healthcare delivery. NPJ Digital Medicine 2021, 4(1):93.
7. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW: Large language models in medicine. Nature medicine 2023, 29(8):1930-1940.
8. Yang X, Chen A, PourNejatian N, Shin HC, Smith KE, Parisien C, Compas C, Martin C, Costa AB, Flores MG: A large language model for electronic health records. NPJ digital medicine 2022, 5(1):194.
9. Betzler BK, Chen H, Cheng C-Y, Lee CS, Ning G, Song SJ, Lee AY, Kawasaki R, van Wijngaarden P, Grzybowski A: Large language models and their impact in ophthalmology. The Lancet Digital Health 2023, 5(12):e917-e924.
10. Yang Z, Wang D, Zhou F, Song D, Zhang Y, Jiang J, Kong K, Liu X, Qiao Y, Chang RT: Understanding natural language: Potential application of large language models to ophthalmology. Asia-Pacific Journal of Ophthalmology 2024:100085.
11. Rojas-Carabali W, Agrawal R, Gutierrez-Sinisterra L, Baxter SL, Cifuentes-González C, Wei YC, Arputhan AJ, Kannapiran P, Wong S, Lee B: Natural Language Processing in Medicine and Ophthalmology: A Review for the 21st-century clinician. Asia-Pacific Journal of Ophthalmology 2024:100084.
12. Wu SN, Qin DY, Zhu L, Guo SJ, Li X, Huang CH, Hu J, Liu Z: Uveal melanoma distant metastasis prediction system: A retrospective observational study based on machine learning. Cancer Science 2024, 115(9):3107-3126.
13. Rong R-Y, Shen Y-K, Wu S-N, Xu S-H, Hu J-Y, Zou J, He L, Chen C, Kang M, Ying P: Prediction model for ocular metastasis of breast cancer: machine

learning model development and interpretation study. BMC cancer 2024, 24(1):1472.
14. Sun JQ, Wu SN, Mou ZL, Wen JY, Wei H, Zou J, Li QJ, Liu ZL, Xu SH, Kang M: Prediction model of ocular metastasis from primary liver cancer: Machine learning-based development and interpretation study. Cancer Medicine 2023, 12(20):20482-20496.
15. Chen X, Zhang W, Zhao Z, Xu P, Zheng Y, Shi D, He M: ICGA-GPT: report generation and question answering for indocyanine green angiography images. British Journal of Ophthalmology 2024, 108(10):1450-1456.
16. Xu P, Chen X, Zhao Z, Shi D: Evaluation of a digital ophthalmologist app built by gpt4-v (ision). medRxiv 2023:2023.2011. 2027.23299056.
17. Zheng C, Ye H, Guo J, Yang J, Fei P, Yuan Y, Huang D, Huang Y, Peng J, Xie X: Development and evaluation of a large language model of ophthalmology in Chinese. British Journal of Ophthalmology 2024, 108(10):1390-1397.
18. Zhao Z, Zhang W, Chen X, Song F, Gunasegaram J, Huang W, Shi D, He M, Liu N: Slit Lamp Report Generation and Question Answering: Development and Validation of a Multimodal Transformer Model with Large Language Model Integration. Journal of medical Internet research 2024, 26:e54047.
19. Chen X, Zhao Z, Zhang W, Xu P, Wu Y, Xu M, Gao L, Li Y, Shang X, Shi D: EyeGPT for Patient Inquiries and Medical Education: Development and Validation of an Ophthalmology Large Language Model. Journal of medical Internet research 2024, 26:e60063.
20. Wu X, Zhao H, Zhu Y, Shi Y, Yang F, Liu T, Zhai X, Yao W, Li J, Du M: Usable XAI: 10 strategies towards exploiting explainability in the LLM era. arXiv preprint arXiv:240308946 2024.
21. Zhao H, Chen H, Yang F, Liu N, Deng H, Cai H, Wang S, Yin D, Du M: Explainability for large language models: A survey. ACM Transactions on Intelligent Systems and Technology 2024, 15(2):1-38.
22. Omiye JA, Gui H, Rezaei SJ, Zou J, Daneshjou R: Large language models in medicine: the potentials and pitfalls: a narrative review. Annals of internal medicine 2024, 177(2):210-220.
23. Cheng K, Guo Q, He Y, Lu Y, Xie R, Li C, Wu H: Artificial intelligence in sports medicine: could GPT-4 make human doctors obsolete? Annals of Biomedical Engineering 2023, 51(8):1658-1662.
24. Potapenko I, Boberg-Ans LC, Stormly Hansen M, Klefter ON, van Dijk EH, Subhi Y: Artificial intelligence-based chatbot patient information on common retinal diseases using ChatGPT. Acta Ophthalmologica (1755375X) 2023, 101(7).
25. Pushpanathan K, Lim ZW, Yew SME, Chen DZ, Lin HAHE, Goh JHL, Wong WM, Wang X, Tan MCJ, Koh VTC: Popular large language model chatbots' accuracy, comprehensiveness, and self-awareness in answering ocular symptom queries. Iscience 2023, 26(11).

26. Masalkhi M, Ong J, Waisberg E, Zaman N, Sarker P, Lee AG, Tavakkoli A: ChatGPT to document ocular infectious diseases. Eye 2024, 38(5):826-828.
27. Cheng K, Sun Z, He Y, Gu S, Wu H: The potential impact of ChatGPT/GPT-4 on surgery: will it topple the profession of surgeons? International Journal of Surgery 2023, 109(5):1545-1547.
28. Chung P, Fong CT, Walters AM, Aghaeepour N, Yetisgen M, O'Reilly-Shah VN: Large language model capabilities in perioperative risk prediction and prognostication. JAMA surgery 2024, 159(8):928-937.
29. Miller R: A Surgical Perspective on Large Language Models. Annals of Surgery 2023, 278(2):e211-e213.
30. Savage T, Wang J, Shieh L: A large language model screening tool to target patients for best practice alerts: development and validation. JMIR Medical Informatics 2023, 11:e49886.
31. Feretzakis G, Vagena E, Kalodanis K, Peristera P, Kalles D, Anastasiou A: GDPR and Large Language Models: Technical and Legal Obstacles. Future Internet 2025, 17(4):151.
32. Wu G, Del Buono M, Wong A, Zhao W, Nguyen M, Satheesh S, Lee DA: Can AI Large Language Models and AI Assistants help educate our Retinitis Pigmentosa (RP) patients? Investigative Ophthalmology & Visual Science 2024, 65(7):2327-2327.
33. Sevgi M, Antaki F, Keane PA: Medical education with large language models in ophthalmology: custom instructions and enhanced retrieval capabilities. British Journal of Ophthalmology 2024, 108(10):1354-1361.
34. Cai LZ, Shaheen A, Jin A, Fukui R, Yi JS, Yannuzzi N, Alabiad C: Performance of generative large language models on ophthalmology board–style questions. American journal of ophthalmology 2023, 254:141-149.
35. Zhang Y, Chen J, Liao L, Ding X: Global Trends and Projections for Avoidable Visual Impairment among Working-Age Individuals: A Population-Based Study. Am J Ophthalmol 2025, 271:304-315.
36. Li Y, Wang H, Guan Z, Guo C, Guo P, Du Y, Yin S, Chen B, Jiang J, Ma Y et al: Persistence of severe global inequalities in the burden of blindness and vision loss from 1990 to 2019: findings from the Global Burden of Disease Study 2019. Br J Ophthalmol 2024, 108(2):301-309.