



doi: 10.12419/es24120902

View this article at: <https://dx.doi.org/10.12419/es24120902>

• Original Article •

Large language models in ophthalmology: a bibliometric analysis

Ruyue Shen^{1,2}, Eunice See Heng Lee¹, Xiaoyan Hu¹, Clement C. Tham^{1,2,3,4}, Carol Y. Cheung^{1,2}

1. Department of Ophthalmology and Visual Sciences, The Chinese University of Hong Kong, Hong Kong SAR 999077, China.
2. Lam Kin Chung, Jet King-Shing Ho Glaucoma Treatment and Research Centre, Department of Ophthalmology and Visual Sciences, The Chinese University of Hong Kong, Hong Kong SAR 999077, China.
3. Hong Kong Eye Hospital, Hong Kong SAR 999077, China.
4. Department of Ophthalmology and Visual Sciences, The Prince of Wales Hospital, Hong Kong SAR 999077, China.

HIGHLIGHTS

- What is already known on this topic: The current literature on large language models (LLMs) and their applications in healthcare has seen a significant increase in publication output since 2021, reflecting the growing adoption of LLM technologies. This trend underscores the necessity for ongoing research to understand the implications of LLMs integration in clinical settings.
- What this study adds: This study revealed that while research was predominantly concentrated in high-income countries, notable exceptions exist with China leading in specific LLMs applications. These anomalies highlighted the diverse global landscape of AI research and its varying adoption rates across different economies.
- How this study might impact research, practice, or policy: The current focus on the clinical applications of LLMs, particularly in diagnosis and patient education, suggested a paradigm shift in healthcare delivery. This study emphasizes the potential for LLMs to enhance clinical decision-making and patient engagement, paving the way for future research and policy initiatives to optimize LLMs use in diverse healthcare contexts.

Received date: 2025-05-27; Revised date: 2025-06-25; Accepted date: 2025-07-17; Published online: 2025-09-26

Corresponding author: Ruyue Shen, E-mail: ruyueshen@cuhk.edu.hk.

Full Text

Abstract: **Backgrounds:** With the rapid development of artificial intelligence (AI), large language models (LLMs) have emerged as a potent tool for invigorating ophthalmology across clinical, educational, and research fields. Their accuracy and reliability have undergone tested. This bibliometric analysis aims to provide an overview of research on LLMs in ophthalmology from both thematic and geographical perspectives. **Methods:** All existing and highly cited LLM-related ophthalmology research papers published in English up to 24th April 2025 were sourced from Scopus, PubMed, and Web of Science. The characteristics of these publications, including publication output, authors, journals, countries, institutions, citations, and research domains, were analyzed using Biblioshiny and VOSviewer software. **Results:** A total of 277 articles from 1,459 authors and 89 journals were included in this study. Although relevant publications began to appear in 2019, there was a significant increase starting from 2023. He M and Shi D are the most prolific authors, while *Investigative Ophthalmology & Visual Science* stands out as the most prominent journal. Most of the top-publishing countries are high-income economies, with the USA taking the lead, and the University of California is the leading institution. VOSviewer identified 5 clusters in the keyword co-occurrence analysis, indicating that current research focuses on the clinical applications of LLMs, particularly in diagnosis and patient education. **Conclusions:** While LLMs have demonstrated effectiveness in retaining knowledge, their accuracy in image-based diagnosis remains limited. Therefore, future research should investigate fine-tuning strategies and domain-specific adaptations to close this gap. Although research on the applications of LLMs in ophthalmology is still in its early stages, it holds significant potential for advancing the field.

Keywords: artificial intelligence; large language models; ophthalmology; bibliometric analysis

Cite this article as: Shen RY, Lee ESH, Xiaoyan Hu XY, Tham CC, Cheung CY. Large language models in ophthalmology: a bibliometric analysis. *Eye Science*, 2025, 2(3): 222-237. doi: 10.12419/es24120902.

INTRODUCTION

Large language models (LLMs) are deep learning neural networks trained on large amounts of textual data. This training enables them to be proficient in natural language processing (NLP) tasks, such as comprehension, generation, and manipulation.^[1] Generative artificial intelligence (AI) chatbots, which are developed using LLMs, can engage in conversations with users. These interactive and human-like experiences stand in stark contrast to other foundation models. These other models are typically trained to perform specific tasks, such as detection, classification, segmentation, and text-to-image retrieval, and do not necessitate human-AI interaction. ChatGPT, a well-known LLM, was fine-tuned by OpenAI from the LLM Generative Pre-training Transformer (GPT)-3.5 and was launched in

November 2022. Its latest version, GPT-4o was released in May 2024. This new version boasts enhanced image processing capabilities and higher accuracy.^[2]

LLMs have been put to the test for a variety of clinical tasks, including disease detection, medicine prescription, the creation of operational and discharge notes, and responding to patients' inquiries. As a relatively novel topic, there are limited bibliometric analyses on LLM-related medical research. Moreover, most researchers tend to adopt a broad scope rather than focusing on specific medical specialties.^[3] There have been bibliometric analyses on AI in ophthalmology^[4] or AI in specific ophthalmology fields, such as cataracts and myopia.^[5-6] However, at present, there are no bibliometric analyses that specifically highlight LLMs in ophthalmology.

In this study, we utilized bibliometric methods to

analyze publication characteristics, including output, authors, journals, countries, and institutions, as well as citation patterns and research domains. Our aim is to evaluate the current status and future directions of LLMs in ophthalmology.

METHODS

Data source and search strategy

We retrieved publications related to LLMs in ophthalmology from three databases, Scopus, PubMed, and Web of Science Core Collection, on 24th April 2025. Articles published after this date were excluded. Our search strategy was formulated following a preliminary literature review. To enhance the effectiveness of the search, we combined specific LLM-related keywords, such as “large language model*”, with ophthalmic terms including “ophthalm*” and “eye disease*”. In addition, common ophthalmic diseases, like “cataract”, “glaucoma”, and “diabetic retinopathy”, as well as LLM programs such as “ChatGPT” and “Google Gemini”, were separately listed as free-text queries to ensure comprehensive coverage. The specific retrieval search strings are presented in Supplementary Table 1.

Screening strategy

Two researchers independently conducted the screening process to guarantee the reproducibility and accuracy of the study. All publications were evaluated based on the following inclusion criteria: 1) published in English; 2) articles or conference papers; and 3) relevant to LLMs in ophthalmology. Editorials, Comments, Letters, and book chapters were excluded from this bibliometric analysis. The citation management software (EndNote Desktop Version 20.0) was used to screen eligible publications. Figure 1 illustrates the search and screening process.

Statistical analysis

This bibliometric analysis focused on publication characteristics, citations, and research domains. The Biblioshiny (version 2.0), a web interface under an R statistical programming package Bibliometrix, was used to analyse publication features.^[7] VOSviewer software (version 1.6.18; Centre for Science and Technology Studies, Leiden University) was used to analyze bibliometrics data, including citations, authorship, and journal indices, so as to quantify research trends in this specific field.^[8] The publication output, top publishing

Table 1 Top 8 authors with the most publications in LLMs in ophthalmology

Rank based on output	authors	Articles	Articles Fractionalized	Rank based on citations	Total citations	G-index based on citations
1	He Mingguang	9	1.046	1	74	8
2	Shi Danli	9	1.256	7	32	5
3	Chen Xiaolan	8	1.131	7	32	5
4	Tham Yih Chung	8	1.567	1	66	8
5	Zhang Weiyi	8	1.006	6	24	4
6	Delsoz, Mohammad	7	0.880	3	88	7
7	Madadi Yeganeh	7	0.880	3	88	7
8	Yousefi Siamak	7	0.880	3	88	7

authors, journals, countries, and institutions, were assessed. Impact factors of journals were obtained from Clarivate’s Journal Citation Reports.^[9] The top-cited articles were identified with Biblioshiny. A co-citation map generated by VOSviewer visualized the relationships between articles, thereby revealing the intellectual structures. We compute the number of links for each node and the total link strength. Biblioshiny calculated the betweenness centrality, which measures the number of times a node lies on the shortest path between two other nodes, as an indicator of its influence.^[10] Finally, VOSviewer produced keyword co-occurrence maps to highlight major research themes.

RESULTS

A total of 2,082 articles were retrieved for initial screening, with 475 sourced from PubMed, 865 from Scopus, and 742 articles from Web of Science. After screening, 1,441 articles were excluded for the following reasons: 1) 304 were not peer-reviewed journal papers or conference papers; 2) 596 were not published in English; 3) 365 were duplicated publications; and 4) 176 were not relevant to LLMs in ophthalmology. Finally, 277 papers were included in the bibliometric analysis (Figure 1).

Characteristics of publications

Publications output

Since the inception of research in this field in 2019, the publication output has generally shown an upward trend. From 2019 to 2022, only 6 articles were published. In 2023, the number increased to 50 articles, and in 2024, there was a 3-fold increase to 166 articles. As of the current date, 55 works have been published or are projected to be published in 2025.

Figure 2 illustrates the monthly fluctuations in article publication. There was a notable peak in June 2024, with 44 articles accounting for 15.9% (44/277) of all papers. Smaller peaks were also observed in December 2023 and October 2024.

Authors

A total of 1,459 authors were included in this analysis. The top authors in this field were defined as those who have published the largest number of relevant articles (Table 1). He M and Shi D had the highest output, with 9 articles each. These 2 authors have collaborated on several articles, including “Slit Lamp Report Generation and Question Answering: Development and Validation of a Multimodal Transformer Model with Large Language Model Integration”^[11] and “ChatFFA: An ophthalmic chat system for unified vision-language understanding

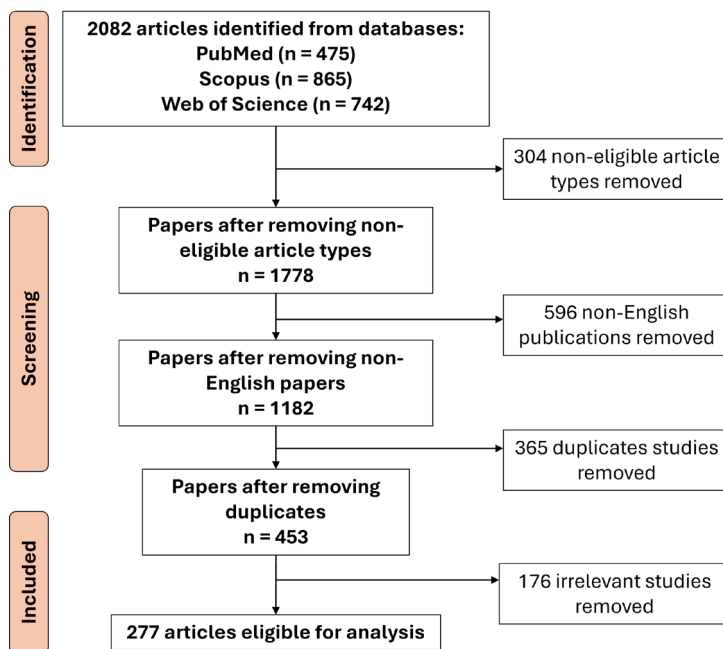


Figure 1 PRISMA flow diagram detailing the searching and screening process

and question answering for fundus fluorescein angiography”^[11]. The G-index,^[11] an author-level metric that assigns greater weight to highly cited articles by identifying the largest number of top publications whose combined citations are at least the square of that number, reflects an author's overall impact. He Mingguang and Tham Yih Chung had the highest G-index scores of 8.

Journals

89 journals were included in this analysis. The top 5 journals with the most publications in this field are *Investigative Ophthalmology & Visual Science*, *Eye*, *British Journal of Ophthalmology*, *Cureus Journal of Medical Science and Ophthalmology Science*, contributing 14.4% (40/277), 6.86% (19/277), 5.05% (14/277), 4.69% (13/277) and 3.97% (11/277) of the total articles respectively. These 5 journals can be considered dominant, as the next-ranking journals have only 8 or fewer articles. They also ranked among the top 6 based on the total number of citations their articles received, except for *Investigative Ophthalmology & Visual Science*, which has a lower ranking.

The top 10 journals (Figure 3) started publishing articles in this field in 2023, except for the *British Journal of Ophthalmology*, which had its first relevant publication

in 2022. The number of publications from these journals has increased from 2023 to 2025, showing a positive trend. In particular, the *British Journal of Ophthalmology* experienced the largest percentage increase (1,300%).

The ranks based on impact factors differ from those based on the number of articles or local citations, with *JAMA Ophthalmology*, *Journal of Medical Internet Research*, and *Investigative Ophthalmology & Visual Science* occupying the top ranks.

Countries & institutions

The articles in this analysis were published in 46 countries. Figure 4 shows an uneven geographical distribution of publications, with a high concentration in North America and East Asia.

Frequent international collaborations are evident in this field, with a total of 191 inter-country links (Figure 4). The most prominent collaborations include those between USA and Singapore (12 articles), China and Singapore (12 articles), USA and the United Kingdom (11 articles), USA and the United Kingdom and Singapore (10 articles) and USA and China (10 articles).

The articles in this analysis originated from 431 institutions. The top 10 contributing institutions (Figure 5) are from USA, Singapore, Canada, the United Kingdom

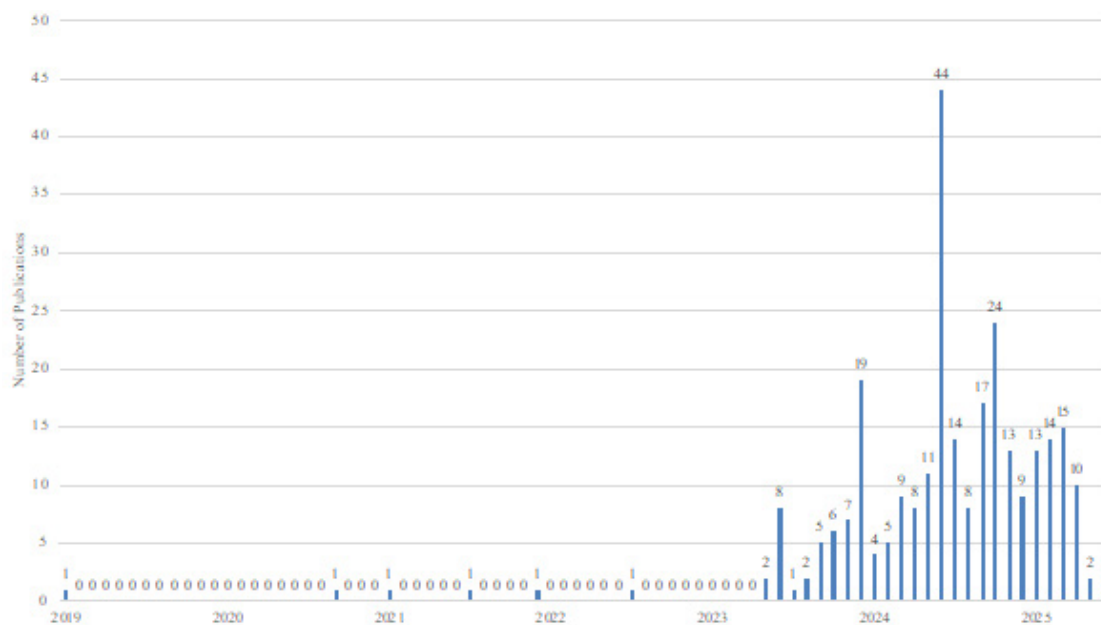


Figure 2 Graph showing monthly publications output from 2021 to 2025

Monthly intervals are used to allow for a more meaningful analysis of a temporal scale, as LLM is a relatively novel topic in the field of ophthalmology.

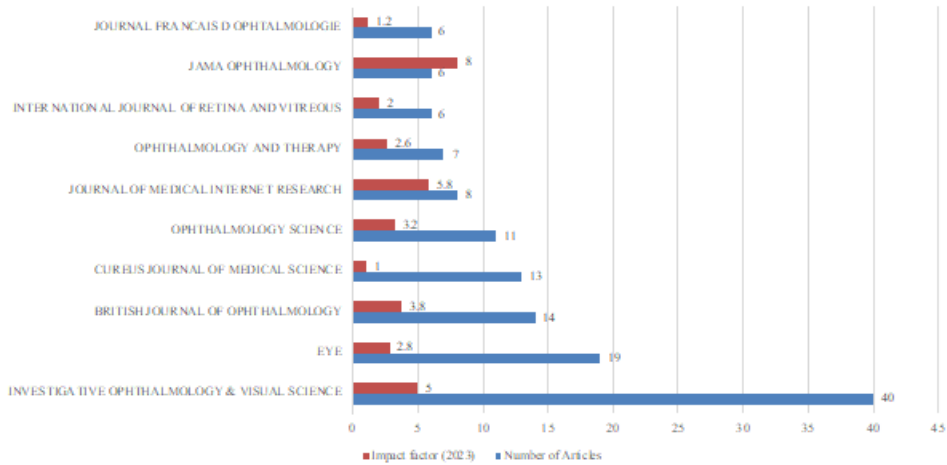


Figure 3 Graph of the top 10 journals with the most publications in LLMs in ophthalmology

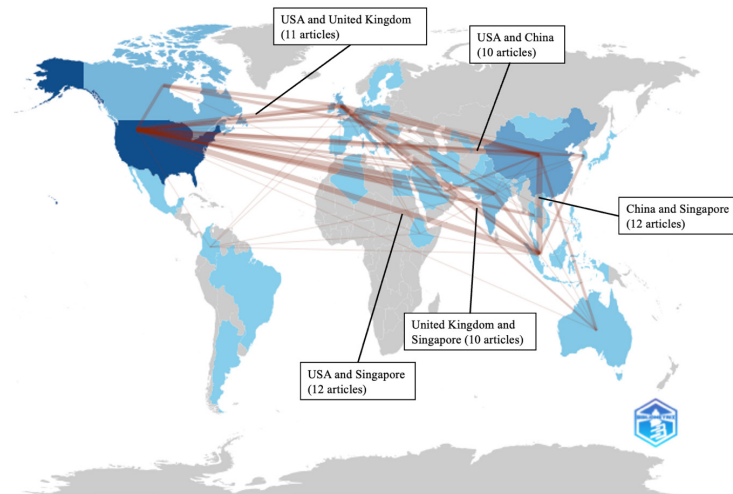


Figure 4 Choropleth map showing the inter-country collaborations of publications in LLMs in ophthalmology

A darker shade of blue indicates higher degree of collaboration. Grey colour indicates no publications. The thickness of the connecting lines is proportional to the number of collaborations between the two connected countries. The most prominent collaborations are labelled.

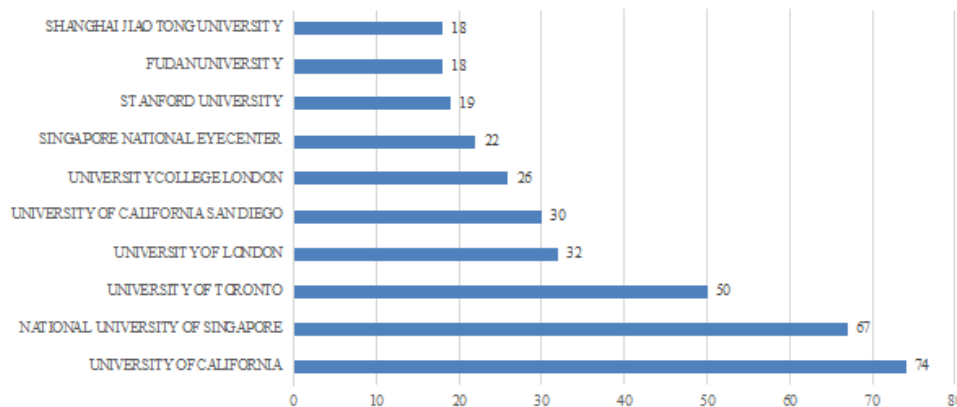


Figure 5 Graph of the top 10 affiliations with most publications in LLMs in ophthalmology

and China, which corresponds to the top 5 contributing countries in this field. The University of California, National University of Singapore and University of Toronto ranked highest with 74, 67, and 50 articles respectively.

Citations

Table 2 shows the 10 articles with the highest number of citations, while Supplementary Figure 1 presents a co-citation network with 63 nodes representing the most co-cited articles in our dataset. The articles with the highest total link strength (Supplementary Table 2) were often cited together with other articles in this field, indicating their high influence and significance. This is likely due to their monumental discoveries, such as Mihalache et al.'s study,^[12] or detailed exploration of

current LLM applications, such as Antaki et al.'s study.^[13] 7 out of the 10 top-cited articles also rank high in terms of total link strength.

Research domains

Most of the top-occurring keywords (Supplementary Table 3) revolve around LLMs and AI, with ChatGPT being a highly researched LLM. The applications of LLM by ophthalmologists and in patient education are a hot topic, while glaucoma is the most studied disease.

The keyword co-occurrence network (Supplementary Figure 2) has 5 clusters (Table 3). Based on these clusters, the major themes in this field are 1) clinical practice (clusters 1, 3); 2) education of patients and medical personnel (cluster 4); 3) management and quality-control of LLMs (cluster 2).

Table 2 Top 10 cited articles in LLMs in ophthalmology

Article	Corresponding author	Journal	Publication year	Number of citations
Evaluating the performance of ChatGPT in ophthalmology	R. Duval	<i>Ophthalmology Science</i>	2023	224
Performance of an artificial intelligence chatbot in ophthalmic knowledge assessment	R. H. Muni	<i>Jama Ophthalmology</i>	2023	163
Benchmarking large language models' performances for myopia care: a comparative analysis of ChatGPT-3.5, ChatGPT-4.0, and Google Bard	Y.C. Tham	<i>eBioMedicine</i>	2023	140
Appropriateness and readability of ChatGPT-4-Generated responses for surgical treatment of retinal diseases	A.E. Kuriyan	<i>Ophthalmology Retina</i>	2023	95
ChatGPT and ophthalmology: exploring its potential with discharge summaries and operative notes	M.J. Ali	<i>Seminars in Ophthalmology</i>	2023	90
Comparison of ophthalmologist and large language model chatbot responses to online patient eye care questions	S.Y. Wang	<i>JAMA Network Open</i>	2023	86
Performance of generative large language models on ophthalmology board-style questions	L.Z. Cai	<i>American Journal of Ophthalmology</i>	2023	71
Artificial Intelligence in ophthalmology: a comparative analysis of GPT-3.5, GPT-4, and human expertise in answering StatPearls questions	M. Moshirfar	<i>Cureus Journal of Medical Science</i>	2023	64
New meaning for NLP: the trials and tribulations of natural language processing with GPT-3 in ophthalmology	P.A. Keane	<i>British Journal of Ophthalmology</i>	2022	58
The use of ChatGPT to assist in diagnosing glaucoma based on clinical case reports	S. Yousefi	<i>Ophthalmology and Therapy</i>	2023	56

Table 3 6 major clusters in keyword co-occurrence network

Cluster no.	Cluster category	Colour	Number of keywords	Examples of keywords
1	clinical practice	red	15	diagnosis, imaging, management
2	management of LLMs	green	12	accuracy, quality, reliability
3	health and ophthalmic conditions	blue	11	keratoconus, myopia, retina
4	patient education	yellow	7	health literacy, education materials, patient education
5	large fundamental ideas	purple	6	artificial intelligence, large language model, telemedicine

DISCUSSION

This study used bibliometrics data from 277 articles from Scopus, PubMed, and Web of Science to analyze the publication trends, current characteristics, and research hotspots of LLMs in the field of ophthalmology.

Characteristics of publications

LLM research initiated in the late 1990s and gradually extended into the medical field in 2011. It began with applications in general practice, then advanced to mental health and oncology, before reaching more complex specialties in 2021. Although articles concerning LLMs in ophthalmology first appeared in 2019,^[14] the first significant surge in publications occurred in June 2023. The relative scarcity of research in this field prior to 2023 can be attributed to the inherent complexity of ophthalmology, where imaging plays a pivotal role in diagnosis. Early LLMs were limited to integrating only textual information, and there was a dearth of structured, labelled ophthalmology data available for training and testing these models. However, the successive releases of GPT-3 in 2020, ChatGPT in 2022, and multimodal GPT-4 in 2023 have created new opportunities for LLM applications in ophthalmology.

The earliest publications in this area focused on two main aspects: 1) knowledge assessment of LLMs, demonstrating that ChatGPT could correctly answer approximately half of the multiple-choice ophthalmology exam questions,^[12,15] and 2) the potential and associated concerns of applying LLM in ophthalmic clinical,

education, and research settings.^[11] Over the years, researchers have continued to evaluate the performance of LLMs in ophthalmology professional examinations. As the accuracy of LLMs has improved, the research focus has gradually shifted towards their application in clinical practice, ranging from relative simple knowledge-retainment tasks, such as answering patients' inquiries, to more challenging tasks like disease detection and diagnosis.^[16-17] Knowledge assessments have also evolved from multiple-choice questions to open-ended questions that required LLMs to generate comprehensive and specific responses.^[12] Since 2025, there has been further exploration into the ability of LLMs to integrate visual and textual data.^[12] Given that these preliminary breakthroughs have not yet yielded promising results, it is anticipated that such research will continue to expand, with more clinical trials being conducted as LLMs strive to transform ophthalmological clinical practice in the foreseeable future.

The most productive researchers in this field are He Mingguang and Shi Danli. As members of the School of Optometry in Hong Kong Polytechnic University, their collaborative efforts in refining LLMs for the interpretation of various ophthalmic imaging modalities and question-answering, exemplified by programs like EyeGPT and FFA-GPT,^[12] indicate a new trend from May 2024 towards engaging LLMs in multimodal tasks. Their contributions may have played a significant role in a 7-fold increase in cumulative publications in China from 2023 to 2025, enabling China to become the country with 2nd highest number of relevant publications

in this field.

The geographic distribution of publications is highly uneven, with high-income economies, as classified by the World Bank,^[18] dominating the field. LLM research demands advanced network infrastructure and well-organized medical datasets, which low-income countries may lack due to limited access to advanced technology and well-structured healthcare databases. However, China and India are exceptions, as their large populations provide substantial patient sample sizes for conducting LLM-related research.

The ranking for top-producing and top-cited journals does not align with their impact factors, as shown in Figure 3. One possible reason is that the number of published articles and citations for the journals in our dataset ($n=277$) is calculated within this specific set, whereas impact factors take into account all articles published by each journal. Journals with a longer history of establishment and a more rigorous peer review process, such as *JAMA Ophthalmology*, may enjoy higher prestige and reliability compared to newer journals like *Eye*, and thus have higher impact factors.

A comparison between the publication characteristics of ophthalmology-related LLM research and the overall trends in LLM research reveals several notable insights. The broader field of LLM research experienced a sharp increase in publications between 2019 and 2020, followed by steady growth through 2023.^[19] This surge coincides with rapid advancements in model architectures, training strategies, and the increasing availability of large datasets and computational resources. In contrast, ophthalmology-related LLM publications first appeared around 2019 but only showed a remarkable acceleration from 2023 to 2024. This lag likely reflects the translational gap between foundational technical innovations and their clinical adoption. The application of LLMs and natural language processing (NLP) techniques in healthcare settings requires additional time for validation, regulatory approval, and integration into clinical workflows, which explains the delayed uptake. The top collaborating countries in LLM research are USA, England, India, and Canada, which is similar to the situation in ophthalmology, except for India. This discrepancy may be due to India's unique research focus, which often emphasizes socially and humanitarian-

centered applications of LLMs, such as language accessibility and public health, rather than specialized clinical fields like ophthalmology. Moreover, differences in resource allocation and healthcare infrastructure across countries may influence the extent of clinical LLM research in ophthalmology. These observations underscore the complex ecosystem of LLM research, where technological breakthroughs precede clinical translation, and geographical research emphases reflect broader socio-economic and healthcare priorities. As LLM capabilities continue to evolve, as evidenced by recent advances in reasoning models and multimodal integration, future ophthalmology research is expected to accelerate, potentially narrowing this translational gap.

Citations

The most cited and co-cited papers are related to two research areas. The first area is Knowledge Assessment. The paper with the highest citation count, which also had the strongest total link strength in the co-citation network, reported that when ChatGPT Plus answered multiple-choice questions from two sources, it achieved accuracies of 59.4% and 49.2%.^[13] The second-most cited paper also obtained similar results.^[12] Studies on knowledge assessment were frequently referenced because they demonstrated the reliability of LLMs, which is crucial in discussions about whether LLMs should be implemented in healthcare. The other research area is clinical application. In the fifth-most cited paper, it was found that ChatGPT could rapidly construct discharge summaries and operative notes, but the accuracy was lacking. Interestingly, ChatGPT acknowledged its inaccuracy, self-corrected, and avoided similar mistakes when answering subsequent prompts.^[20] On the other hand, the fifth most co-cited paper reported that when ChatGPT diagnosed glaucoma based on textual case descriptions, it achieved an accuracy of 72.7%, surpassing the average accuracy of three ophthalmology residents.^[20] Although LLMs could not complete tasks with absolute accuracy, further training and fine-tuning may pave the way for the future integration of these efficient tools into healthcare.

Research domains

Our study provided a thorough overview of the primary research domains: 1) Incorporating LLMs into

clinical practice; 2) Using LLMs to educate patients and medical personnel; and 3) The management and quality-control of LLMs. Our analysis of keyword clusters reveals a clear evolution in LLM research within ophthalmology. Early work focused on foundational NLP techniques and model development, while recent studies emphasize clinical applications, specific eye diseases, and integration into clinical practice. Emerging themes include patient-centered applications and ethical considerations, reflecting a shift from technical innovation toward translational and implementation science. This progression highlights the need for interdisciplinary collaboration to address clinical validation, safety, and real-world integration. Understanding these trends can guide future research to ensure LLMs effectively enhance patient care and healthcare delivery.

Incorporating LLMs into clinical practice

Many studies have demonstrated the potential of LLMs to assist ophthalmologists efficiently and objectively. ChatGPT's ability to diagnose and triage patients based on textual case descriptions has been shown to match or surpass that of ophthalmologists in several studies.^[21]

Applications of generative AI in ocular imaging are an emerging hotspot for ophthalmology, but only 28 articles have tested the ability of LLMs to handle visual data inputs, such as interpreting retinal photographs and making subsequent diagnoses. Multimodal LLMs are still in their infancy and have generated mediocre outcomes, such as a 50.7% accuracy in the diagnosis of retinal cases.^[22] Since ophthalmology relies heavily on the analysis of ocular imaging, including optical coherence tomography (OCT) and fundus photographs, future research should concentrate on making task-specific refinements to multimodal LLMs, such as GPT-4V, to improve their diagnostic abilities. Other applications of image interpretation, such as predictions of disease progression or treatment response, should also be explored. Although LLMs are not designed for image analysis, unlike computer vision models, they can engage in a variety of clinical applications (not task-specific) and are conveniently available online, thus can be adopted into healthcare more efficiently. Aside from improving their accuracy, researchers have sought to expand LLM

applications to non-English-speaking countries, such as ChatFFA.^[22]

LLMs may also assist ophthalmologists in implementing treatment plans for patients. Since human knowledge and memory may be limited, and human judgment may be skewed or not sufficiently thorough, suggestions from LLMs may enhance clinical outcomes. Studies have shown how ChatGPT-4 selected appropriate surgical procedures and choices of intraocular tamponade for patients with retinal detachment,^[23] and outperformed ophthalmologists in generating an assessment and management plan for glaucoma and retinal cases.^[24]

Educating patients and medical personnel

The high frequency of the keyword "patient education" underscores the importance of educating patients about their diagnosis and treatment. Unfortunately, public hospitals are crowded and fast-paced, so physicians have little time to address patients' concerns, often leading to patient confusion, dissatisfaction, and poorer clinical outcomes as patients are less able to comply with their prescribed treatments. With the ability to score highly in ophthalmology exams,^[25] construct human-like responses, and even show empathy when programmed to do so, chatbots may help generate patient-targeted educational materials or answer patients' inquiries. A study in 2022 observed low readability in online ophthalmic materials, where the reading level surpassed that of average Americans.^[26] ChatGPT-4 could generate patient-targeted educational materials about uveitis with higher readability.^[27] Since LLMs are accessible and cost-effective, their utilization in patient education could improve patient satisfaction, optimize the use of healthcare resources, and reduce ophthalmologists' workload.

Furthermore, medical education is important for the cultivation of medical students and ophthalmologists alike. Although textbooks and websites offer a plethora of information, it may be difficult to obtain the most relevant source, especially for ophthalmologists, as patients have specific needs and anatomical variations. Seygi et al.^[28] developed 3 GPTs for ophthalmic education: 'EyeTeacher', which explains concepts using a Q&A method; 'EyeAssistant', which answers ophthalmologists' inquiries regarding patient management; and "The

GPT for GA”, a clinical assistant specific to geographic atrophy. The release of custom GPTs enables users to modify GPTs with natural language processing and create novel GPTs for specific educational purposes. Hence, LLMs can facilitate patient care by compensating for the knowledge gaps of medical personnel.

Management and quality-control of LLMs

Although LLMs have the potential to improve patient care and education in ophthalmology, barriers still exist regarding their implementation. Studies show that LLMs are subject to hallucinations, misinterpretation, bias, and inconsistency. In one study, 24.4% of ChatGPT’s responses to cataract patients’ post-surgery inquiries had the potential to cause harm, and 9.5% did not align with the scientific consensus.^[29] Therefore, the accuracy and consistency of LLMs must be improved. Ophthalmologists must also remain critical and only incorporate LLMs’ suggestions if they are scientifically proven to be beneficial, non-maleficent, and align with the ophthalmologist’s knowledge.

Furthermore, the lack of encryption in ChatGPT conversations poses cybersecurity threats. Prompt injection attacks have been shown to bypass security filters,^[30] which endangers patients’ confidentiality and infringes on medical ethics.

Limitations

This study has several limitations. First, we excluded non-English and non-peer-reviewed articles. This exclusion may have resulted in the omission of relevant studies, particularly regional literature and research published in other languages. Consequently, it could limit the global comprehensiveness and diversity of perspectives captured in our bibliometric mapping. Second, our search was restricted to Scopus, PubMed, and Web of Science. Databases such as Embase may contain additional pertinent publications that were not included in our study, which could potentially affect the completeness of our dataset. Third, our analysis focused on publication characteristics, citations, and research domains. Future studies should employ more advanced methods, such as exploratory factor analysis and Latent Dirichlet Allocation, to gain deeper insights. Additionally, some high-frequency keywords, like

“artificial intelligence,” were too broad to analyse effectively. Moreover, burst keyword analysis was not feasible because the topic only emerged in 2019. A follow-up study conducted in 5 to 10 years would be valuable for more accurately tracking and interpreting evolving research trends.

In conclusion, LLMs hold great promise for streamlining administrative tasks and supporting ophthalmologists in diagnosis and decision-making. However, issues such as hallucinations and cybersecurity must be addressed, and rigorous testing is required before their clinical use. Future research should focus on bridging the gap between research and clinical application, especially in the field of ocular images.

Correction Notice

None

Acknowledgement

This study was supported in part by the Endowment Fund for Lam Kin Chung. Jet King-Shing Ho Glaucoma Treatment and Research Centre, Hong Kong.

Author Contributions

- (I) Conception and design: Ruyue Shen
- (II) Administrative support: Clement C. Tham; Carol Y. Cheung
- (III) Provision of study materials or patients: Ruyue Shen
- (IV) Collection and assembly of data: Eunice See Heng Lee
- (V) Data analysis and interpretation: Ruyue Shen; Eunice See Heng Lee; Xiaoyan Hu
- (VI) Manuscript writing: All authors
- (VII) Final approval of manuscript: All authors

Funding

This study was supported by Health and Medical Research Fund, Hong Kong (11220386, 12230246).

Conflict of Interests

None of the authors has any conflicts of interest to disclose. All authors have declared in the completed the ICMJE uniform disclosure form.

Patient Consent for Publication

None

Ethical Statement

None

Provenance and Peer Review

This article was a standard submission to our journal. The article has undergone peer review with our anonymous review system.

Data Sharing Statement

None

Open Access Statement

This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license).

References

- Bharathi Mohan G, Prasanna Kumar R, Vishal Krishh P, et al. An analysis of large language models: their impact and potential applications. *Knowl Inf Syst*, 2024, 66(9): 5047-5070. DOI:10.1007/s10115-024-02120-8.
- OpenAI. Accessed August 28, 2024. <https://openai.com/index/hello-gpt-4o/>.
- Wu J, Ma Y, Wang J, et al. The application of ChatGPT in medicine: a scoping review and bibliometric analysis. *J Multidiscip Healthc*, 2024, 17: 1681-1692. DOI: 10.2147/JMDH.S463128.
- Jiang X, Xie M, Ma L, et al. International publication trends in the application of artificial intelligence in ophthalmology research: an updated bibliometric analysis. *Ann Transl Med*, 2023, 11(5): 219. DOI: 10.21037/atm-22-3773.
- Chen S, Huang L, Li X, et al. Hotspots and trends of artificial intelligence in the field of cataracts: a bibliometric analysis. *Int Ophthalmol*, 2024, 44(1): 258. DOI: 10.1007/s10792-024-03207-5.
- Wu XY, Fang HH, Xu YW, et al. Bibliometric analysis of hotspots and trends of global myopia research. *Int J Ophthalmol*, 2024, 17(5): 940-950. DOI: 10.18240/ijjo.2024.05.20.
- Aria M, Cuccurullo C. Bibliometrix: an R-tool for comprehensive science mapping analysis. *J Informetr*, 2017, 11(4): 959-975. DOI: 10.1016/j.joi.2017.08.007.
- Guo Y, Hao Z, Zhao S, et al. Artificial intelligence in health care: bibliometric analysis. *J Med Internet Res*, 2020, 22(7): e18228. DOI: 10.2196/18228.
- Gurav R. Data from: Journal Citation Reports (JCR): Impact Factor 2024. . 2024. Deposited June 2024. DOI: 10.13140/RG.2.2.34797.60640
- Peng S, Zhou Y, Cao L, et al. Influence analysis in social networks: a survey. *J Netw Comput Appl*, 2018, 106: 17-32. DOI: 10.1016/j.jnca.2018.01.005.
- Dossantos J, An J, Javan R. Eyes on AI: ChatGPT's transformative potential impact on ophthalmology. *Cureus*, 2023, 15(6): e40765. DOI: 10.7759/cureus.40765.
- Mihalache A, Popovic MM, Muni RH. Performance of an artificial intelligence chatbot in ophthalmic knowledge assessment. *JAMA Ophthalmol*, 2023, 141(6): 589-597. DOI: 10.1001/jamaophthalmol.2023.1144.
- Antaki F, Touma S, Milad D, et al. Evaluating the performance of ChatGPT in ophthalmology: an analysis of its successes and shortcomings. *Ophthalmol Sci*, 2023, 3(4): 100324. DOI:10.1016/j.xops.2023.100324.
- Lee JH, Jeong MS, Cho JU, et al. Developing a ophthalmic chatbot system. 2021 15th Int Conf Ubiquitous Inf Manag Commun IMCOM. 2021.
- Teebagy S, Colwell L, Wood E, Faustina M. ChatGPT in Ophthalmology: A Pilot Study. Meeting Abstract. *Invest Ophth Vis Sci*, 2023, 64(8).
- Olis M, Dyjak P, Weppelmann TA. Performance of three artificial intelligence chatbots on Ophthalmic Knowledge Assessment Program materials. *Can J Ophthalmol*, 2024, 59(4): e380-e381. DOI: 10.1016/j.jcjo.2024.01.011.
- Haddad F, Saade JS. Performance of ChatGPT on ophthalmology-related questions across various examination levels: observational study. *JMIR Med Educ*, 2024, 10: e50842. DOI: 10.2196/50842.
- World Bank Country and Lending Groups – World Bank Data Help Desk. <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups>.
- Fan L, Li L, Ma Z, et al. A bibliometric review of large language models research from 2017 to 2023. *ACM Trans Intell Syst Technol*, 2024, 15(5): 1-25. DOI: 10.1145/3664930.
- Singh S, Djalilian A, Ali MJ. ChatGPT and ophthalmology: exploring its potential with discharge summaries and operative notes. *Semin Ophthalmol*, 2023, 38(5): 503-507.

- DOI: 10.1080/08820538.2023.2209166.
21. Delsoz M, Raja H, Madadi Y, et al. The use of ChatGPT to assist in diagnosing glaucoma based on clinical case reports. *Ophthalmol Ther*, 2023, 12(6): 3121-3132. DOI: 10.1007/s40123-023-00805-x.
 22. Mihalache A, Huang RS, Mikhail D, et al. Interpretation of clinical retinal images using an artificial intelligence chatbot. *Ophthalmol Sci*, 2024, 4(6): 100556. DOI: 10.1016/j.xops.2024.100556.
 23. Carlà MM, Gambini G, Baldascino A, et al. Exploring AI-chatbots' capability to suggest surgical planning in ophthalmology: ChatGPT versus Google Gemini analysis of retinal detachment cases. *Br J Ophthalmol*, 2024, 108(10): 1457-1469. DOI: 10.1136/bjo-2023-325143.
 24. Huang AS, Hirabayashi K, Barna L, et al. Assessment of a large language model's responses to questions and cases about glaucoma and retina management. *JAMA Ophthalmol*, 2024, 142(4): e236917. DOI: 10.1001/jamaophthalmol.2023.6917.
 25. Fowler T, Pullen S, Birkett L. Performance of ChatGPT and Bard on the official part 1 FRCOphth practice questions. *Br J Ophthalmol*, 2024, 108(10): 1379-1383. DOI: 10.1136/bjo-2023-324091.
 26. Cohen SA, Pershing S. Readability and accountability of online patient education materials for common retinal diseases. *Ophthalmol Retina*, 2022, 6(7): 641-643. DOI: 10.1016/j.oret.2022.03.015.
 27. Kianian R, Sun D, Crowell EL, et al. The use of large language models to generate education materials about uveitis. *Ophthalmol Retina*, 2024, 8(2): 195-201. DOI: 10.1016/j.oret.2023.09.008.
 28. Sevgi M, Antaki F, Keane PA. Medical education with large language models in ophthalmology: custom instructions and enhanced retrieval capabilities. *Br J Ophthalmol*, 2024, 108(10): e325046. DOI: 10.1136/bjo-2023-325046.
 29. Chowdhury M, Lim E, Higham A, et al. Can large language models safely address patient questions following cataract surgery? [C]//Proceedings of the 5th Clinical Natural Language Processing Workshop. Toronto, Canada. Stroudsburg, PA, USA: ACL, 2023: 131-137. DOI: 10.18653/v1/2023.clinicalnlp-1.17.
 30. Gupta M, Akiri C, Aryal K, et al. From ChatGPT to ThreatGPT: impact of generative AI in cybersecurity and privacy. *IEEE Access*, 2023, 11: 80218-80245. DOI:10.1109/ACCESS.2023.3300381.

SUPPLEMENTARY MATERIALS

Supplementary Table 1 The search strategy for Scopus, PubMed, and Web of Science (As of 24th April 2025)

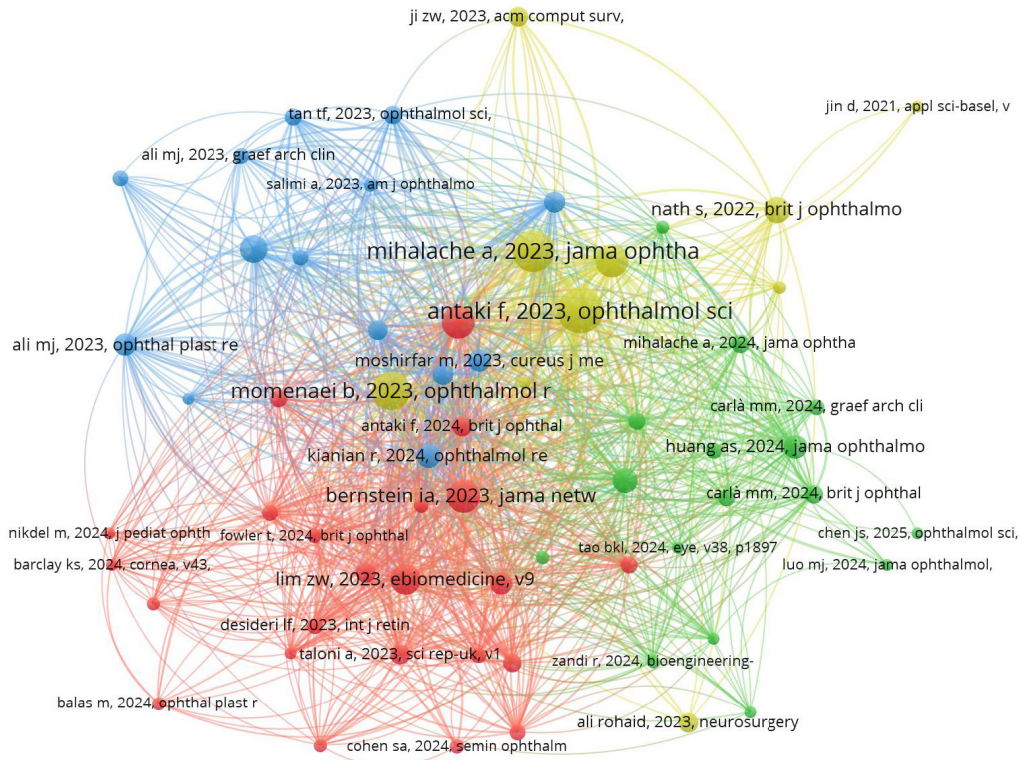
Database	Search string	Result
Scopus	TITLE-ABS-KEY ("ophthalm*" OR "ocular" OR "optic" OR "eye disease*" OR "orbital disease*" OR "retina" OR "Graves Ophthalmopathy" OR "Retinal disease*" OR "Acute Retinal necrosis" OR "Diabetic Retinopathy" OR "Orbital Lymphoma" OR "Optic Nerve" OR "Dry Eye" OR "Asthenopia" OR "Conjunctival disease*" OR "optic neuropathy" OR "Uveal Disease*" OR "Eye Neoplasm" OR "Eyelid Disease*" OR "Eye Hemorrhage" OR "Scleral Disease*" OR "thyroid-associated ophthalmopathy" OR "acute optic neuritis" OR "Age-Related Macular Degeneration" OR "Retinal Necrosis Syndrome" OR "glaucoma" OR "cataract" OR "myopia" OR "strabismus" OR "retinoblastoma" OR "amblyopia") AND TITLE-ABS-KEY ("large language model*" OR "natural language processing" OR "ChatGPT" OR "chatbot*" OR "GPT" OR "OpenAI" OR ("Google" AND "BERT") OR ("Anthropic" AND "Claude") OR ("Meta" AND "LLaMA") OR ("Stanford" AND "Alpaca"))	865
PubMed	("ophthalm*[Title/Abstract] OR "ocular"[Title/Abstract] OR "optic"[Title/Abstract] OR "eye disease*[Title/Abstract] OR "orbital disease*[Title/Abstract] OR "retina"[Title/Abstract] OR "Graves Ophthalmopathy"[Title/Abstract] OR "Retinal disease*[Title/Abstract] OR "Acute Retinal necrosis"[Title/Abstract] OR "Diabetic Retinopathy"[Title/Abstract] OR "Orbital Lymphoma"[Title/Abstract] OR "Optic Nerve"[Title/Abstract] OR "Dry Eye"[Title/Abstract] OR "Asthenopia"[Title/Abstract] OR "Conjunctival disease*[Title/Abstract] OR "optic neuropathy"[Title/Abstract] OR "Uveal Disease*[Title/Abstract] OR "Eye Neoplasm"[Title/Abstract] OR "Eyelid Disease*[Title/Abstract] OR "Eye Hemorrhage"[Title/Abstract] OR "Scleral Disease*[Title/Abstract] OR "thyroid-associated ophthalmopathy"[Title/Abstract] OR "acute optic neuritis"[Title/Abstract] OR "Age-Related Macular Degeneration"[Title/Abstract] OR "Retinal Necrosis Syndrome"[Title/Abstract] OR "glaucoma"[Title/Abstract] OR "cataract"[Title/Abstract] OR "myopia"[Title/Abstract] OR "strabismus"[Title/Abstract] OR "retinoblastoma"[Title/Abstract] OR "amblyopia"[Title/Abstract]) AND ("large language model*[Title/Abstract] OR "natural language processing"[Title/Abstract] OR "ChatGPT"[Title/Abstract] OR "chatbot*[Title/Abstract] OR "GPT"[Title/Abstract] OR "OpenAI"[Title/Abstract] OR ("Google"[Title/Abstract] AND "BERT"[Title/Abstract]) OR ("Anthropic"[Title/Abstract] AND "Claude"[Title/Abstract]) OR ("Meta"[Title/Abstract] AND "LLaMA"[Title/Abstract]) OR ("Stanford"[Title/Abstract] AND "Alpaca"[Title/Abstract]))	475
Web of Science	(ALL=("ophthalm*" OR "ocular" OR "optic" OR "eye disease*" OR "orbital disease*" OR "retina" OR "Graves Ophthalmopathy" OR "Retinal disease*" OR "Acute Retinal necrosis" OR "Diabetic Retinopathy" OR "Orbital Lymphoma" OR "Optic Nerve" OR "Dry Eye" OR "Asthenopia" OR "Conjunctival disease*" OR "optic neuropathy" OR "Uveal Disease*" OR "Eye Neoplasm" OR "Eyelid Disease*" OR "Eye Hemorrhage" OR "Scleral Disease*" OR "thyroid-associated ophthalmopathy" OR "acute optic neuritis" OR "Age-Related Macular Degeneration" OR "Retinal Necrosis Syndrome" OR "glaucoma" OR "cataract" OR "myopia" OR "strabismus" OR "retinoblastoma" OR "amblyopia")) AND ALL=("large language model*" OR "natural language processing" OR "ChatGPT" OR "chatbot*" OR "GPT" OR "OpenAI" OR ("Google" AND "BERT") OR ("Anthropic" AND "Claude") OR ("Meta" AND "LLaMA") OR ("Stanford" AND "Alpaca"))	742

Supplemental Table 2 Top 9 articles in LLMs in ophthalmology based on total link strength values in co-citation network

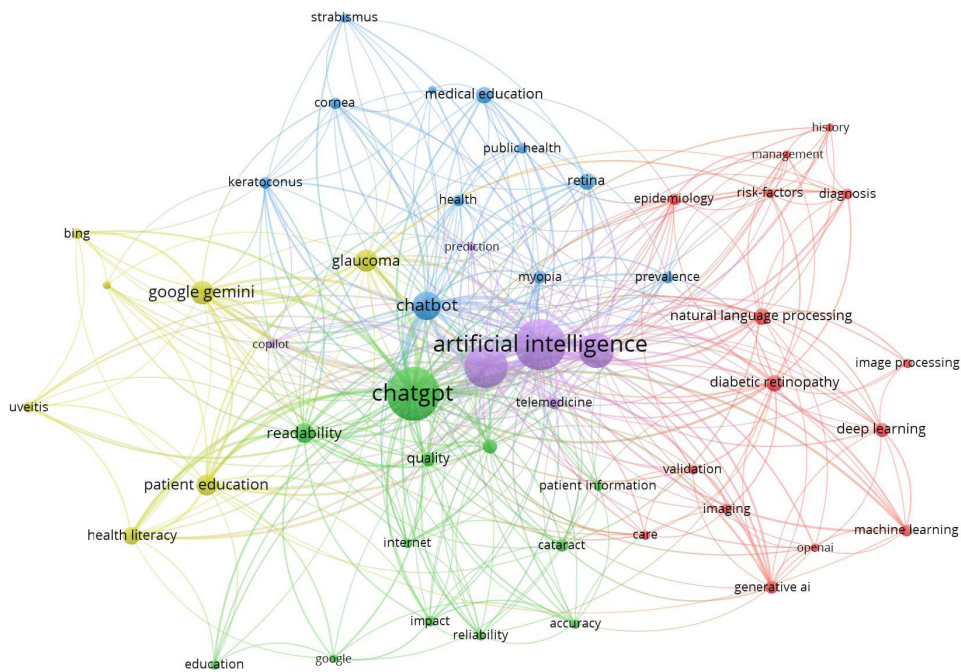
Article	Total link strength
Evaluating the performance of ChatGPT in ophthalmology: an analysis of its successes and shortcomings	432
Performance of an artificial intelligence chatbot in ophthalmic knowledge assessment	410
Appropriateness and readability of ChatGPT-4-Generated responses for surgical treatment of retinal diseases	381
Performance of generative large language models on ophthalmology board-style questions	338
The use of ChatGPT to assist in diagnosing glaucoma based on clinical case reports	322
Comparison of ophthalmologist and large language model chatbot responses to online patient eye care questions	294
Benchmarking large language models' performances for myopia care: a comparative analysis of ChatGPT-3.5, ChatGPT-4.0, and Google Bard	267
Artificial intelligence-based ChatGPT chatbot responses for patient and parent questions on vernal keratoconjunctivitis	250
Capabilities of GPT-4 in ophthalmology: an analysis of model entropy and progress towards human-level medical question answering	215

Supplemental Table 3 Top 10 occurring keywords in LLM in ophthalmology

Keyword	Occurrence	Total link strength
ChatGPT	114	375
artificial intelligence	108	359
large language model	75	247
ophthalmology	50	186
chatbot	35	150
google gemini	22	86
glaucoma	21	75
patient education	18	77
readability	17	73



Supplemental Figure 1 Co-citation network of publications in LLMs in ophthalmology
 Threshold: frequency ≥ 5; Number of nodes: 63; Number of edges: 1,537; Total link strength: 4,615.



Supplemental Figure 2 Keyword co-occurrence network on LLMs in ophthalmology
 Threshold: frequency ≥ 3; Number of nodes: 51; Number of edges: 408; Total link strength: 1,247.