Content available at: https://www.ipinnovative.com/open-access-journals

IP International Journal of Ocular Oncology and Oculoplasty

Journal homepage: https://ijooo.org/

Original Research Article

Assessing AI chatbots efficacy in ophthalmic triage and referrals: A comparative study

Sarita Panda[1], Piyush Jain[1]*, Ankita Mishra[1], Deepika Priyadarshini[1], Radhakanta Bhoi[1]

[1]*Dept. of Ophthalmology, M.K.C.G Medical College, Berhampur, Odisha, India*

ARTICLE INFO

Article history:
Received 29-05-2024
Accepted 29-08-2024
Available online 02-11-2024

Keywords:
Artificial Intelligence
Ophthalmology
AI Chatbots
Diagnostic Accuracy
Triage
Referrals

ABSTRACT

Aims: To evaluate the efficacy of AI chatbots (OpenAI ChatGPT GPT-3.5, Google Bard, and WebMD) compared to human ophthalmology trainees in triage and referrals for common ophthalmic conditions.
Materials and Methods: A single-center study was conducted at MKCG Medical College, Berhampur, Odisha, involving six ophthalmology trainees. The performance of AI chatbots was assessed based on diagnostic accuracy and triage categorization. Key performance indicators included the accuracy of the top three suggested diagnoses and concordance in recommendations for investigations and referrals.
Results: Physician respondents identified the correct diagnosis among the top three suggestions in 95% of cases. Google Bard achieved 90% accuracy, ChatGPT 85%, and WebMD 20%. High concordance was observed between physician and AI recommendations for investigations and referrals.
Conclusion: AI chatbots demonstrate promising potential in supporting triage and referral decisions for ophthalmic conditions. While human expertise remains crucial, AI tools can augment diagnostic accuracy, improve efficiency, and enhance patient care. Future research should focus on refining AI algorithms, integrating clinical data, and exploring real-world implementation strategies.

1. Introduction

Ophthalmology, the realm of vision, is embracing the technological wave with AI chatbots emerging as potential game-changers. These virtual assistants hold immense promise for enhancing patient care, improving accessibility, and augmenting medical expertise.[1]

1. Patient education and triage: Chatbots can offer 24/7 access to reliable eye health information, empowering patients to understand symptoms, manage appointments, and even perform preliminary self-triage.[2]
2. Glaucoma management: Chatbots can remind patients about medication adherence, answer questions about drops, and provide emotional support, potentially leading to better glaucoma control.[3]
3. Physician support: Chatbots can handle routine inquiries, freeing up ophthalmologists' time for complex cases and consultations[4–6] They can also analyze vast amounts of data to aid in diagnosis and treatment planning.

Imagine a world where patients with blurry vision, eye pain, or sudden flashes of light could instantly access a trusted advisor, capable of analyzing their symptoms, suggesting potential diagnoses, and even guiding them towards the appropriate specialist.[7]

This is no longer science fiction, but the potential of AI chatbots in ophthalmic care

* *Corresponding author*.
E-mail address: drpiyushjain721@gmail.com (P. Jain).

https://doi.org/10.18231/j.ijooo.2024.026

Globally, nearly 2.2 billion people have vision impairment, with preventable blindness affecting millions. Precise triage and referral are crucial to saving sight and preventing complication[8,9]

Every year, millions of people visit ophthalmologists with eye problems, yet accurate diagnosis and timely referrals remain critical challenges. Could AI chatbots be the answer?

This cross-sectional study aimed to:

1. Evaluate the accuracy of AI Chatbot's in identifying and prioritizing diagnoses from ophthalmic clinical vignettes.
2. Compare the performance of AI Chatbot's with human ophthalmology trainees in triage and referral recommendations.
3. Assess the potential of AI Chatbot's to support clinical decision-making in ophthalmology.

2. Materials and Methods

One hundred clinical vignettes encompassing a spectrum of common ophthalmic conditions encountered in a tertiary care outpatient department were meticulously crafted. These vignettes included diverse presentations, varying degrees of severity, and pertinent historical details to mimic real-world patient encounters. Conditions covered included cataracts, glaucoma, diabetic retinopathy, age-related macular degeneration, corneal abnormalities, uveitis, and strabismus. Each vignette adhered to a standardized format, presenting patient demographics, chief complaint, presenting symptoms, past medical history, ophthalmic history, family history, and relevant social history.

2.1. Participant selection

2.1.1. Four groups of participants were recruited for this study

2.1.1.1. Ophthalmology trainees. A group of six ophthalmology trainees at various were recruited to represent the perspective of junior physicians encountering various ophthalmic conditions in the OPD setting

2.1.1.2. Open AI Chat GPT (GPT 3.5). This state-of-the-art generative pre-trained transformer model was employed, recognized for its proficiency in understanding and responding to natural language queries.

2.1.1.3. Gemini. Google AI's advanced generative model, Bard, was included due to its demonstrated capabilities in healthcare applications and clinical reasoning.

2.1.1.4. WebMD. As a widely used online platform for medical information, WebMD was incorporated to represent the readily available resources.

2.1.2. Standardized prompt

To ensure consistency and comparability, all participants were presented with the same standardized prompt for each vignette. This prompt included the patient's demographics, presenting complaint, current symptoms, past medical and ophthalmic history, medications, and any available diagnostic test results. Participants were instructed to:

1. Generate a list of possible diagnoses: They were asked to identify the most likely diagnosis based on the provided information and rank them in order of probability.
2. Recommend a management plan: This included suggestions for further investigations, referrals to specialists if necessary, and initial treatment recommendations.

2.2. Data collection and analysis

Responses from all participants were collected and anonymized. Diagnoses listed by each participant were compared against a pre-defined "gold standard" diagnosis established by a panel of senior ophthalmologists. Accuracy was measured by calculating the percentage of cases where the true diagnosis was listed among the top three suggestions. Management plans were assessed for alignment with established referral guidelines and best practices in ophthalmic care. Inter-rater reliability among physician respondents was evaluated using Cohen's kappa coefficient.

2.3. Endpoint and scoring system

2.3.1. Diagnosis

1. Top 3 Diagnosis
   (a) For each vignette, participants will provide their top 3 most likely diagnoses.
   (b) Scoring: Each diagnosis will be compared to a pre-defined list of established, correct diagnoses for the vignette.
   (c) Exact Match: 3 points
   (d) Partially Correct: 2 points (e.g., if the participant lists a specific subtype of the correct diagnosis)
   (e) Incorrect: 0 points
2. Triage
   (a) Triage Category: Participants will categorize the urgency of the case based on the presented symptoms and potential diagnosis.
   (b) Urgent: Requires immediate specialist consultation or intervention.
   (c) Semi-urgent: Requires referral within a specific timeframe (e.g., within 24 hours).
   (d) Non-urgent: Can be managed by a primary care physician or scheduled for a follow-up appointment.

3. Referral
   (a) Score: Each participant received 1 point for each vignette where the referral recommendation (specialist referral, follow-up with primary care, no referral) corresponded to accepted practices for the given diagnosis.

2.3.2. *Overall scoring*

The total score for each participant will be calculated by summing the points earned across all diagnoses, triage categories, and referral recommendations for all 60 vignettes. This score will provide a comprehensive comparison of performance across the different participant groups.

3. Results

1. Diagnosis accuracy - Physician respondents listed the appropriate diagnosis among the top three suggestions in 95% of cases. Google Gemini correctly identified the diagnosis in 90% of cases, followed by ChatGPT at 85% and WebMD at 20%. High concordance was observed between physician and AI recommendations for investigations and referrals.
2. Traige accuracy
   (a) Physician respondents correctly categorized the urgency level (urgent, semi-urgent, non-urgent) in 97 (97%) of the cases.
   (b) Gemini demonstrated high accuracy in triage, correctly classifying urgency in 93 (93%) of the cases.
3. ChatGPT's triage accuracy was slightly lower, with correct categorization in 87 (87% of the cases
4. WebMD's performance in triage was significantly weaker, with accurate urgency classification in only 35 (35%) of the cases.

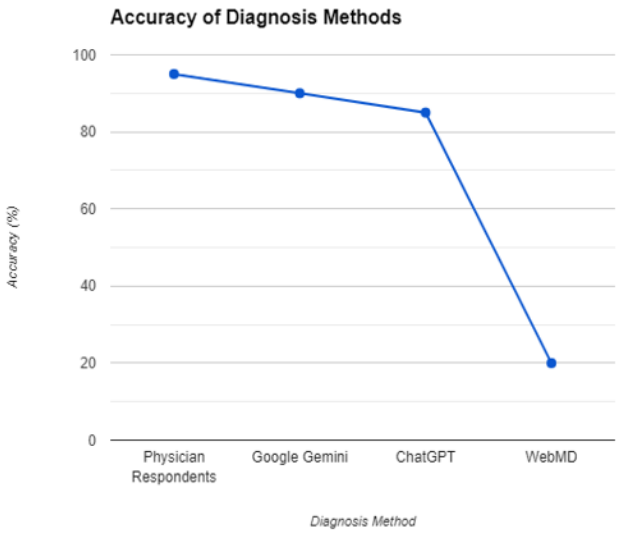

Figure 1: Accuracy of diagnosis methods

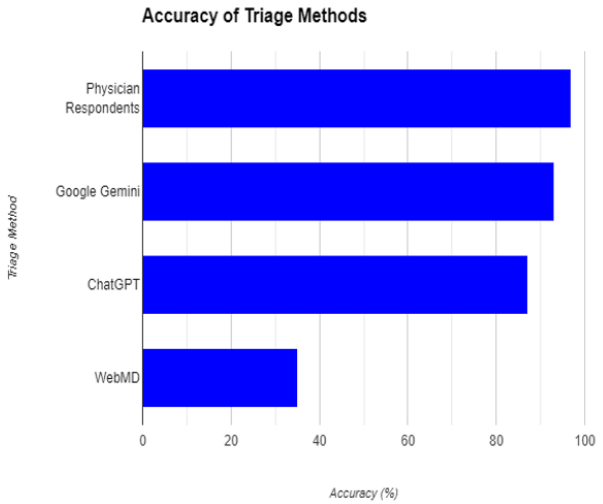

Figure 2: Accuracy of triage methods

4. Discussion

The present study evaluated the triage ability of GPT technology using ChatGPT and Google Gemini across a wide range of ophthalmic conditions. The results demonstrated high diagnostic and triage accuracy for Google Gemini, comparable to that of physicians. Both ChatGPT and Gemini outperformed the existing online medical triage service, WebMD Symptom Tracker. These findings suggest that AI chatbots can serve as valuable adjuncts to human expertise in ophthalmology.[10]

These results highlight the potential of AI chatbots to:

4.1. *Augment clinician decision-making*

By providing accurate preliminary diagnoses and suggesting appropriate referrals, AI can streamline the triage process and improve efficiency.

Early and accurate diagnosis allows for timely intervention and potentially better patient outcomes. AI chatbots can also empower patients with information and guide them towards seeking appropriate medical attention.[11]

AI-powered tools can provide basic triage and information in underserved areas or after clinic hours, potentially reducing healthcare disparities.[12–15]

5. Conclusion

This study demonstrates the promising potential of AI chatbots in supporting triage and referral decisions for ophthalmic conditions. While human expertise remains

Table 1: Clinical vignettes

	Chat GPT	
1.I am a 28year old male . Today work I splashed cement mixture in my eyes . I am developing blurring of vision and pain.	I'm not a medical professional, but it's crucial to treat eye injuries seriously. Splashing cement mixture in your eyes can be dangerous. Please seek immediate medical attention. Rinse your eyes with clean, lukewarm water for at least 15 minutes while waiting for medical help. Do not rub your eyes, as this may cause further damage. If possible, let your coworkers or supervisor know about the incident so they can assist you in seeking medical help. Eye injuries require prompt attention to minimize potential damage	I understand you're in a worrisome situation and experiencing concerning symptoms after getting cement mixture in your eyes. It's crucial to seek medical attention immediately. Cement contains highly alkaline substances that can cause severe and permanent damage to your eyes if not treated promptly. Do not: Rub your eyes. Apply any medication or ointment without consulting a medical professional. Try to remove any remaining cement particles yourself. Instead: Call emergency services or proceed to the nearest emergency department immediately. Rinse your eyes with clean, lukewarm water continuously for at least 15 minutes. Avoid putting your head down, as this can cause the cement to settle further into your eye. Cover both eyes loosely with a clean, damp cloth to protect them from further irritation
2. I am 32years old female. Today I developed sudden vision loss with no pain no trauma.	I'm not a medical professional, but sudden vision loss is a serious symptom that requires immediate medical attention. It could be caused by various conditions, including issues with blood flow to the eye or the retina. Please seek emergency medical care as soon as possible. Contact your local emergency services or go to the nearest emergency room for a thorough evaluation by a healthcare professional. Delay in seeking medical attention for sudden vision loss can be critical, so it's essential to act quickly.	I understand that you're experiencing sudden vision loss without any pain or trauma, and that's very concerning. While I cannot provide medical advice or diagnose your condition, I strongly urge you to seek immediate medical attention. Sudden vision loss, even in the absence of pain, can be a sign of a serious underlying medical condition that requires prompt evaluation and treatment. Call emergency services or proceed to the nearest emergency department. Contact your ophthalmologist or primary care physician as soon as possible. They can assess your situation and recommend further evaluation or treatment. Seek urgent care at a nearby walk-in clinic.
3.Mujhe glaucoma ho gaya hai		Here's the thing: I can't diagnose medical conditions. But I can tell you some symptoms of glaucoma to look for: Vision loss: This can be gradual or sudden, and usually affects peripheral vision first. Blind spots: You might have areas where you can't see anything, even though your eye is open. Tunnel vision: Your field of vision may narrow, like looking through a tunnel. Eye pain: This is more common with acute angle-closure glaucoma, a less frequent type. Redness: Your eye may appear red and irritated. Halos around lights: You might see rainbow-colored rings around lights.

paramount, AI tools can serve as valuable adjuncts, enhancing diagnostic accuracy, efficiency, and patient care. Future research should focus on refining AI algorithms, integrating clinical data, and exploring real-world implementation strategies.

6. Limitations

Single-center study: The findings may not be generalizable to other healthcare settings.

Clinical vignettes: Real-world patient consultations can be more complex, potentially impacting AI accuracy.

Limited scope: The study focused on common conditions. Performance with rarer cases needs further evaluation.

7. Source of Funding

None.

8. Conflict of Interest

None.

References

1. Topol EJ. High-performance medicine: The convergence of human and artificial intelligence. *Nat Med*. 2019;25:44–56.

doi:10.1038/s41591-018-0300-7.
2. Lee P, Bubeck S, Petro J. Benefits, Limits, and Risks of GPT-4 as an AI Chatbot for Medicine. *N Engl J Med*. 2023;388:1233–9. doi:10.1056/NEJMsr2214184.
3. Michels TC, Ivan O. Glaucoma: Diagnosis and Management. *Am Fam Physician*. 2023;107(3):253–62.
4. Muehlematter UJ, Daniore P, Vokinger KN. Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe (2015–20): a comparative analysis. *Lancet Digit Health*. 2021;3(3):195–203.
5. Wiedemann P. Artificial intelligence in ophthalmology. *Int J Ophthalmol*. 2023;16(9):1357–60.
6. Li Z, Wang L, Wu X, Jiang J, Qiang W, Xie H, et al. Artificial intelligence in ophthalmology: The path to the real-world clinic. *Cell Rep Med*. 2023;4(7):101095. doi:10.1016/j.xcrm.2023.101095.
7. Nagendran M, Chen Y, Lovejoy CA, Gordon AC, Komorowski M, Topol EJ, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. *BMJ*. 2021;368:m689. doi:10.1136/bmj.m689.
8. National Eye Institute. (2020, December). Facts About Glaucoma. National Institutes of Health (.gov) . Available from: https://www.nei.nih.gov/learn-about-eye-health/outreach-resources/outreach-materials/facts-about-glaucoma.
9. Sonmez SC, Sevgi M, Antaki F, Huemer J, Keane P. Generative artificial intelligence in ophthalmology: current innovations, future applications and challenges. *Br J Ophthalmol*. 2024;108(10):1335–40.
10. American Academy of Ophthalmology. (2023, September). What is Glaucoma? American Academy of Ophthalmology . Available from: https://www.aao.org/glaucoma.
11. Keenan TDL, Loewenstein A. Artificial intelligence for home monitoring devices. *Curr Opin Ophthalmol*. 2023;34(5):441–8.
12. Bajwa J, Munir U, Nori A, Williams B. Artificial intelligence in healthcare: transforming the practice of medicine. *Future Healthc J*. 2021;8(2):188–92.
13. Sun G, Zhou YH. AI in healthcare: navigating opportunities and challenges in digital communication. *Front Digit Health*. 2023;5:1291132. doi:10.3389/fdgth.2023.1291132.
14. Altamimi I, Altamimi A, Alhumimidi AS, Altamimi A, Temsah MH. What is Glaucoma? *Cureus*. 2023;15(6):e40922. doi:10.7759/cureus.40922..
15. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction. *BMC Med Res Methodol*. 2019;19(1):64. doi:10.1186/s12874-019-0681-4.

Author's biography

Sarita Panda, HOD

Piyush Jain, Junior Resident https://orcid.org/0009-0005-5582-348X

Ankita Mishra, Junior Resident https://orcid.org/0009-0001-5003-284X

Deepika Priyadarshini, Assistant Professor https://orcid.org/0009-0002-2270-6074

Radhakanta Bhoi, Assistant Professor

Cite this article: Panda S, Jain P, Mishra A, Priyadarshini D, Bhoi R. Assessing AI chatbots efficacy in ophthalmic triage and referrals: A comparative study. *IP Int J Ocul Oncol Oculoplasty* 2024;10(3):135-139.