INNOVAPATH™
The Premier Journal Bridging Discovery & Innovation

Letter to the Editor
https://doi.org/10.63501/dbw8bs33

Advancing Glaucoma Care: Comparing Vision Language Models and Ophthalmology Foundation Models

Erum Habib[1]

[1]Dow University of Health Sciences

Corresponding Author Email: erumhabib131@gmail.com

Dear Editor,

Glaucoma is a progressive optic neuropathy characterized by retinal ganglion cell loss, optic nerve head cupping, and irreversible visual field defects. Pathophysiologically, elevated intraocular pressure (IOP) and impaired aqueous humor outflow contribute to axonal damage and retinal nerve fiber layer thinning. Clinically, patients present with peripheral vision loss, visual field scotomas, and, in advanced stages, central vision impairment. Despite pharmacological therapies (topical prostaglandin analogues and beta-blockers), laser trabeculoplasty, and surgical interventions such as trabeculectomy and minimally invasive glaucoma surgery (MIGS), delayed diagnosis remains a significant challenge, with glaucoma contributing significantly to global blindness (1).

Recent advances in artificial intelligence (AI) have introduced Vision Language Models (VLMs) and Ophthalmology Foundation Models (OFMs) as diagnostic tools. VLMs integrate multimodal data, including retinal images and clinical text, using transformer-based architectures, enabling contextual interpretation of fundus and OCT findings. Their working principle relies on large-scale pretraining across diverse datasets to enhance generalizability and accuracy. OFMs, in contrast, are domain-specific models trained exclusively on ophthalmic datasets, optimized for tasks such as glaucoma detection, diabetic retinopathy grading, and age-related macular degeneration (AMD) screening. (2) Both approaches demonstrate high efficacy: VLMs achieve robust cross-disease diagnostic accuracy, while OFMs excel in specialized ophthalmic tasks. Beyond glaucoma, these models have been applied to diabetic retinopathy, AMD, and keratoconus, improving prognosis through earlier detection and personalized treatment planning. (3)

Comparative literature highlights their complementary strengths. In one study, *Chia et al.* reported that OFMs trained on Moorfields Eye Hospital datasets achieved higher sensitivity for glaucoma detection than general-purpose VLMs, underscoring the value of domain-specific training. (2) Conversely, *Li et al.* introduced VisionUnite, a VLM enhanced with clinical knowledge, which demonstrated strong performance across multiple ophthalmic conditions, including glaucoma and diabetic retinopathy, with improved interpretability. (4) Together, these findings suggest that while OFMs provide precision in narrowly defined tasks, VLMs offer scalability and adaptability across diverse clinical scenarios.

Despite the promise, several limitations persist. VLMs require massive multimodal datasets, raising concerns about data privacy and standardization. OFMs, while highly accurate, risk overfitting and limited generalizability beyond ophthalmology. Both approaches face challenges in clinical integration, including regulatory approval, clinician trust, and infrastructure requirements. Moreover, disparities in access to

https://doi.org/10.63501/dbw8bs33

advanced imaging technologies may hinder equitable implementation, particularly in low-resource settings. (2,5)

In conclusion, Vision Language Models and Ophthalmology Foundation Models represent transformative tools for glaucoma care, offering complementary pathways to enhance diagnostic accuracy and patient outcomes. Their earliest adoption is likely to occur at critical junctures in the glaucoma care pathway: triage in community-based screening programs, risk stratification in tertiary clinics, and longitudinal monitoring of visual fields in established patients. To ensure safe and effective translation, concrete next steps should include multicenter prospective validation studies, integration into OCT imaging platforms, and pilot implementation within teleophthalmology programs. These strategies will not only establish clinical trust and regulatory confidence but also reduce costs and expand accessibility. By embedding these tools into real-world workflows, glaucoma care can move toward earlier detection, personalized risk management, and improved long-term outcomes.

Author contribution statement

E.H: Conceptualization, Methodology, Software, Data curation, Writing Original draft preparation, Reviewing and Editing.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

The authors received no funding.

Clinical trial registration

Not applicable.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Consent for publication is not applicable as this study involves publicly available data.

Ethics, Consent to Participate, and Consent to Publish declarations:

Not applicable.

Acknowledgements

Not applicable.

Provenance and peer review

Not commissioned, externally peer-reviewed.

REFERENCES

1. Tham YC, Li X, Wong TY, Quigley HA, Aung T, Cheng CY. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. Ophthalmology. 2014 Nov;121(11):2081–90.
2. Chia MA, Antaki F, Zhou Y, Turner AW, Lee AY, Keane PA. Foundation models in ophthalmology. Br J Ophthalmol. 2024 Sept 20;108(10):1341–8.
3. Popescu Patoni SI, Muşat AAM, Patoni C, Popescu MN, Munteanu M, Costache IB, et al. Artificial intelligence in ophthalmology. Romanian J Ophthalmol. 2023;67(3):207–13.
4. Li Z, Song D, Yang Z, Wang D, Li F, Zhang X, et al. VisionUnite: A Vision-Language Foundation Model for Ophthalmology Enhanced with Clinical Knowledge [Internet]. arXiv; 2025 [cited 2025 Dec 6]. Available from: http://arxiv.org/abs/2408.02865
5. Kazemzadeh K. Artificial intelligence in ophthalmology: opportunities, challenges, and ethical considerations. Med Hypothesis Discov Innov Ophthalmol. 2025 May 10;14(1):255–72.