JHWCR JOURNAL OF HEALTH, WELLNESS AND COMMUNITY RESEARCH 3007-0570

Original Article Volume 3, Issue 7

Reliability and Validity of Ophthal-360: A Novel Artificial Intelligence Tool to Diagnose Non-Proliferative Diabetic Retinopathy

Muhammad Awais Asghar[1], Muhammad Moin[1], Tehsin Mehmood Mahju[1]

[1] Institute of Ophthalmology, Mayo Hospital, King Edward Medical University, Lahore, Pakistan

* Correspondence: Muhammad Awais Asghar, awaisasghar3@gmail.com

ABSTRACT

Background: *Diabetic retinopathy (DR) is a leading cause of preventable vision loss among working-age adults, with rising global and national prevalence, particularly in high-burden countries such as Pakistan. Artificial intelligence (AI)–based fundus image analysis offers a scalable solution for early detection and triage, yet local validation is essential before clinical integration.* ***Objective:*** *To evaluate the diagnostic validity and grading agreement of the Ophthal-360 AI tool for detecting and classifying non-proliferative diabetic retinopathy (NPDR) compared with ophthalmologist grading.* ***Methods:*** *This cross-sectional diagnostic accuracy study was conducted at a tertiary-care center in Lahore between June and December 2024. Adults aged 40–85 years with type II diabetes underwent standardized fundus photography. Images were independently graded by a fellowship-trained ophthalmologist (reference standard) using the International Clinical Classification of DR and by Ophthal-360. Sensitivity, specificity, predictive values, overall accuracy, area under the ROC curve (AUC), and weighted kappa were calculated.* ***Results:*** *Among 134 participants (mean age 57.8 ± 10.8 years; 53.7% male), DR prevalence was 77.6%. Ophthal-360 achieved sensitivity of 97.0% (95% CI: 91.8–99.4), specificity of 82.4% (95% CI: 65.5–93.2), PPV of 94.2%, NPV of 90.3%, overall accuracy of 93.3%, and AUC of 0.897. Agreement in NPDR severity grading was strong (κ = 0.86).* ***Conclusion:*** *Ophthal-360 demonstrates high sensitivity and strong grading concordance, supporting its potential role as an assistive screening tool for NPDR in resource-constrained settings.*

Keywords: *Diabetic retinopathy; Artificial intelligence; Diagnostic accuracy; Screening; Non-proliferative diabetic retinopathy; Ophthal-360*

INTRODUCTION

Diabetes mellitus is a chronic metabolic disorder characterized by persistent hyperglycemia due to inadequate insulin secretion, impaired insulin action, or both (1). The global burden is substantial and rising; in 2021, an estimated 537 million adults were living with diabetes, with projections reaching 780 million by 2045 (2). Among the microvascular complications, diabetic retinopathy (DR) remains the most common cause of preventable vision loss in working-age populations, and its prevalence is expected to increase in parallel with diabetes: approximately 103.12 million people were affected in 2020, with projections rising to 129.84 million by 2030 and 160.50 million by 2045 (3). This growing burden has intensified the need for scalable, accurate screening strategies that can detect DR early—before irreversible, vision-threatening stages emerge.

The challenge is particularly pronounced in Pakistan, where diabetes prevalence is estimated at 26%, and DR affects roughly 28.78% of individuals with diabetes (4). Despite the high prevalence, delayed detection and referral remain common, contributing to preventable progression to vision-threatening disease. System-level barriers including fragmented referral pathways and limited DR-specific capacity in primary care have been documented in local settings, resulting in missed opportunities for timely ophthalmology review and treatment initiation (5). In this context, improving early identification of non-proliferative

Received: 01 June 2025
Revised: 09 June 2025
Accepted: 17 June 2025
Published: 29 June 2025

Citation: Click to Cite

diabetic retinopathy (NPDR) the stage where intervention and surveillance can meaningfully alter the disease trajectory represents a critical clinical and public health priority.

Artificial intelligence (AI) enabled image analysis has emerged as a pragmatic solution for DR screening, particularly where ophthalmologist capacity is constrained. Contemporary computer vision approaches, including deep learning and related machine-learning architectures, can detect retinopathy features on color fundus photographs and generate standardized outputs suitable for triage and referral (6). Evidence syntheses indicate that AI-based fundus imaging systems can achieve high diagnostic performance for referable DR in real-world settings, although variability in specificity and misclassification patterns persists across tools and populations, underscoring the need for local validation rather than reliance on performance claims derived from external datasets (7). This is especially important because diagnostic performance is influenced by population characteristics (e.g., disease spectrum and comorbidity), image acquisition protocols, camera types, and the operational definition of the reference standard.

Ophthal-360 (Ophthalytics, USA) is a commercially available, camera-agnostic AI tool intended to assist in DR screening and grading using fundus photographs. While a prospective evaluation has reported favorable diagnostic performance for DR detection using Ophthal-360 in a regional setting, the generalizability of these findings remains uncertain without additional independent assessments in diverse Pakistani clinical workflows and patient mixes, particularly with explicit benchmarking against ophthalmologist grading and standardized disease classification (8). Moreover, because DR screening is not solely about binary detection but also about staging severity for appropriate referral urgency, validation should align with accepted clinical frameworks most commonly the International Clinical Classification of Diabetic Retinopathy, which supports consistent categorization of NPDR severity for decision-making and referral thresholds (9). Collectively, these considerations define a clear knowledge gap: despite the pressing need for scalable screening in Pakistan, there is limited locally grounded evidence on whether Ophthal-360 can accurately detect and appropriately grade NPDR in routine outpatient practice when compared against specialist clinical grading.

Accordingly, this study was designed around a pragmatic diagnostic validation question framed by PICO: in adults with type II diabetes presenting to a tertiary-care ophthalmology outpatient setting in Lahore (Population), does Ophthal-360 analysis of color fundus photographs (Intervention), compared with ophthalmologist (human grader) classification using the International Clinical Classification of Diabetic Retinopathy (Comparator), accurately detect and grade NPDR (Outcomes: diagnostic accuracy metrics for DR detection and concordance in grading severity)? The primary objective was to evaluate the validity of Ophthal-360 for detecting diabetic retinopathy in this population by estimating sensitivity, specificity, positive predictive value, and negative predictive value against the human grader reference, and to assess whether Ophthal-360's NPDR severity grading aligns with ophthalmologist-based classification under routine clinical imaging conditions (9).

METHODS

This cross-sectional observational diagnostic accuracy study was conducted at Eye Unit III, Institute of Ophthalmology, Mayo Hospital, King Edward Medical University, Lahore, Pakistan, between June 2024 and December 2024. The study was designed to evaluate the diagnostic validity and grading agreement of the Ophthal-360 artificial intelligence (AI) tool for the detection of non-proliferative diabetic retinopathy (NPDR), using ophthalmologist grading based on the International Clinical Classification of Diabetic Retinopathy as the

reference standard (9). The design adhered to established methodological principles for diagnostic accuracy studies to ensure transparency, reproducibility, and minimization of bias (10).

Adult patients aged 40 to 85 years with a confirmed diagnosis of type II diabetes mellitus presenting to the outpatient department were screened consecutively for eligibility. Inclusion criteria comprised documented type II diabetes and the ability to undergo fundus photography of sufficient quality for grading. Exclusion criteria included the presence of other retinal pathologies that could confound grading (e.g., retinal vein occlusion, age-related macular degeneration), media opacities precluding adequate fundus visualization, prior retinal laser photocoagulation or intravitreal therapy for diabetic retinopathy, and clinically evident proliferative diabetic retinopathy (PDR), as the version of the algorithm under evaluation was not configured to classify PDR. Consecutive non-probability sampling was employed to minimize selection bias and approximate a real-world screening population. All eligible patients during the study period who met inclusion criteria and provided consent were enrolled until the predetermined sample size was reached.

Participants were recruited during routine outpatient visits. After explanation of study procedures, written informed consent was obtained. Demographic and clinical data were collected through structured clinical interviews and review of medical records. Variables recorded included age, sex, duration of diabetes (years since diagnosis), history and duration of hypertension, and prior ocular history. A comprehensive ophthalmological examination was performed by a fellowship-trained vitreoretinal specialist, including best-corrected visual acuity assessment, slit-lamp biomicroscope, intraocular pressure measurement, and dilated fundus examination.

Color fundus photographs were obtained during the same visit using Topcon and Nidek non-mydriatic fundus cameras with a standard macula-centered 40-degree field of view. Image acquisition followed a standardized protocol with uniform illumination, focus adjustment, and pupil dilation when required to optimize image clarity. All images were de-identified and assigned a unique study identification code prior to grading to ensure masking. The unit of analysis was the patient; for patients with bilateral gradable images, the eye with the more severe retinopathy grade was used for primary analysis to reflect clinical referral thresholds.

The reference standard consisted of grading by a qualified ophthalmologist with expertise in retinal diseases, who was masked to the AI outputs. Grading was performed using the International Clinical Classification of Diabetic Retinopathy, categorizing findings into no DR, mild NPDR, moderate NPDR, or severe NPDR (9). The index test was the Ophthal-360 AI tool (Ophthalytics, USA), which analyzed the same de-identified fundus images independently. AI outputs were recorded as categorical severity grades aligned with the same classification framework. Both graders (human and AI) evaluated images independently, and AI results were generated without access to clinical data to prevent incorporation bias.

The primary outcome variable was the presence of diabetic retinopathy (any NPDR versus no DR). Secondary outcomes included categorical agreement in NPDR severity grading. Diagnostic accuracy measures were defined operationally as follows: sensitivity was the proportion of reference-standard–positive cases correctly identified by AI; specificity was the proportion of reference-standard–negative cases correctly identified; positive predictive value (PPV) and negative predictive value (NPV) were calculated based on observed prevalence in the study sample. A two-by-two contingency table was constructed for DR detection. For severity grading agreement, ordinal categorical concordance between AI and human grading was assessed.

Potential sources of bias were addressed through consecutive sampling, masking of graders to each other's results, standardized image acquisition protocols, and predefined grading criteria. To minimize verification bias, all enrolled participants underwent both the index test and the reference standard assessment. Image quality was reviewed prior to analysis, and images deemed ungradable by either modality were excluded from diagnostic calculations. Confounding variables such as age and duration of diabetes were documented to characterize the population and explore their association with DR prevalence descriptively.

Sample size estimation was performed using a single-proportion formula for diagnostic sensitivity with an anticipated sensitivity of approximately 90%, a desired precision (margin of error) of 5%, and a 95% confidence level, incorporating an estimated DR prevalence consistent with regional data (4). This calculation yielded a minimum required sample of 126 participants; to account for potential exclusions due to ungradable images, a final target sample of 134 patients was set.

Statistical analysis was conducted using SPSS version 20.0 (IBM Corp., Armonk, NY, USA). Continuous variables were summarized as mean ± standard deviation (SD) after assessment of normality using the Shapiro–Wilk test. Categorical variables were presented as frequencies and percentages. Diagnostic accuracy parameters (sensitivity, specificity, PPV, NPV) were calculated with corresponding 95% confidence intervals using binomial exact methods. Receiver operating characteristic (ROC) curve analysis was performed to estimate the area under the curve (AUC) for AI-based DR detection. Agreement between AI and human grading for ordinal NPDR severity categories was assessed using weighted Cohen's kappa (κ) with quadratic weights to account for partial agreement and interpreted according to standard benchmarks (10). A p-value <0.05 was considered statistically significant. Missing data were assessed for randomness; complete-case analysis was performed where applicable, as the proportion of missing data was negligible and did not materially influence outcome estimates.

Ethical approval was obtained from the Ethical Review Board of Mayo Hospital, Lahore (Approval Reference No.: COAVS/603/24, dated 08/06/2024). The study adhered to the principles of the Declaration of Helsinki and its subsequent amendments (11). All participants provided written informed consent prior to inclusion. Data confidentiality was maintained through coded identifiers and secure storage in password-protected databases accessible only to study investigators.

To ensure reproducibility and data integrity, standardized operating procedures were developed for image acquisition, grading, data entry, and statistical analysis. Double data entry and periodic cross-verification were performed to minimize transcription errors. The statistical analysis plan was finalized prior to data analysis, and all analytic steps were documented to allow replication by independent researchers.

RESULTS

Table 1 summarizes the baseline profile of the 134 enrolled participants. The cohort had a mean age of 57.83 ± 10.77 years (95% CI: 55.98–59.68). Males constituted 72/134 (53.7%; 95% CI: 45.2%–62.1) and females 62/134 (46.3%; 95% CI: 37.9%–54.8), yielding a male-to-female ratio of 1.6. The observed sex distribution did not differ significantly from an equal 50:50 distribution (χ^2 test $p = 0.384$). The mean duration of diabetes was 9.22 ± 6.03 years (95% CI: 8.18–10.26), while the mean duration of hypertension was 5.84 ± 4.79 years (95% CI: 4.99–6.69), describing a population with nearly a decade of diabetes exposure and substantial comorbidity burden.

Table 2 presents the comparative distribution of diabetic retinopathy severity grades assigned by the human grader (reference standard) versus the Ophthal-360 AI tool. By human grading, 30/134 participants (22.4%) had no diabetic retinopathy, while 104/134 (77.6%) had NPDR. Within NPDR severity categories, the human grader classified 6/134 (4.5%) as mild NPDR, 58/134 (43.3%) as moderate NPDR, and 40/134 (29.9%) as severe NPDR. Ophthal-360 produced a closely aligned distribution, identifying 27/134 (20.1%) as no DR and 107/134 (79.9%) as NPDR, including 4/134 (3.0%) mild NPDR, 60/134 (44.8%) moderate NPDR, and 43/134 (32.1%) severe NPDR. The absolute differences between human and AI grading across categories were small (2.3% for no DR, 1.5% for mild NPDR, 1.5% for moderate NPDR, and 2.2% for severe NPDR). When severity grades were compared as an ordinal distribution, there was no statistically significant difference between the two methods (Mann–Whitney $U = 8399.0$; $Z = -0.959$; $p = 0.338$), supporting comparable overall grading patterns between the ophthalmologist and AI outputs.

Table 3 reports diagnostic performance metrics for Ophthal-360 in detecting diabetic retinopathy (any NPDR vs no DR) against the human grader reference. Sensitivity was 97.0% (95% CI: 91.8%–99.4), indicating that Ophthal-360 correctly identified nearly all DR-positive cases. Specificity was 82.4% (95% CI: 65.5%–93.2), reflecting a moderate false-positive rate in DR-negative participants.

The PPV was 94.2% (95% CI: 88.1%–97.6), meaning that approximately 94 out of 100 AI-positive classifications were confirmed as DR by the human grader in this relatively high-prevalence cohort. The NPV was 90.3% (95% CI: 74.2%–98.0), indicating that about 9 out of 10 AI-negative classifications were truly free of DR. Overall diagnostic accuracy was 93.3% (95% CI: 87.5%–96.9). Discrimination was also strong on ROC analysis, with an AUC of 0.897 (95% CI: 0.832–0.962; $p < 0.001$), consistent with high ability of the model to separate DR from non-DR cases.

Table 4 provides the underlying 2×2 contingency table supporting these diagnostic estimates. Of the 104 participants classified as DR-positive by the human grader, Ophthal-360 correctly flagged 101 as DR-positive (true positives) and missed 3 (false negatives), which corresponds directly to the 97.0% sensitivity. Among the 30 human-graded DR-negative participants, Ophthal-360 correctly classified 25 as negative (true negatives) while labeling 5 as positive (false positives), which corresponds to the observed specificity of 82.4%. In total, Ophthal-360 labeled 106/134 participants (79.1%) as DR-positive and 28/134 (20.9%) as DR-negative, closely tracking the reference distribution while prioritizing sensitivity over specificity—an operating profile that is typically desirable in screening contexts where minimizing missed disease is the primary goal.

Table 1. Baseline Demographic and Clinical Characteristics of Study Participants (n = 134)

Variable	Mean ± SD or n (%)	95% CI (Mean/Proportion)	p-value*
Age (years)	57.83 ± 10.77	55.98 – 59.68	—
Male sex	72 (53.7%)	45.2% – 62.1%	0.384
Female sex	62 (46.3%)	37.9% – 54.8%	—
Duration of diabetes (years)	9.22 ± 6.03	8.18 – 10.26	—
Duration of hypertension (years)	5.84 ± 4.79	4.99 – 6.69	—

Table 2. Comparison of NPDR Severity Grading Between Human Grader and Ophthal-360 (n = 134)

Severity Grade	Human Grader n (%)	Ophthal-360 n (%)	Absolute Difference (%)	p-value†
No DR	30 (22.4%)	27 (20.1%)	2.3%	0.338
Mild NPDR	6 (4.5%)	4 (3.0%)	1.5%	—
Moderate NPDR	58 (43.3%)	60 (44.8%)	1.5%	—
Severe NPDR	40 (29.9%)	43 (32.1%)	2.2%	—

Table 3. Diagnostic Performance of Ophthal-360 for Detection of Diabetic Retinopathy (n = 134)

Metric	Value (%)	95% CI	p-value‡
Sensitivity	97.0%	91.8% – 99.4%	<0.001
Specificity	82.4%	65.5% – 93.2%	<0.001
Positive Predictive Value	94.2%	88.1% – 97.6%	—
Negative Predictive Value	90.3%	74.2% – 98.0%	—
Overall Accuracy	93.3%	87.5% – 96.9%	—
AUC (ROC)	0.897	0.832 – 0.962	<0.001

Table 4. Contingency Table for Detection of Any Diabetic Retinopathy

	Human Grader DR+	Human Grader DR–	Total
AI DR+	101	5	106
AI DR–	3	25	28
Total	104	30	134

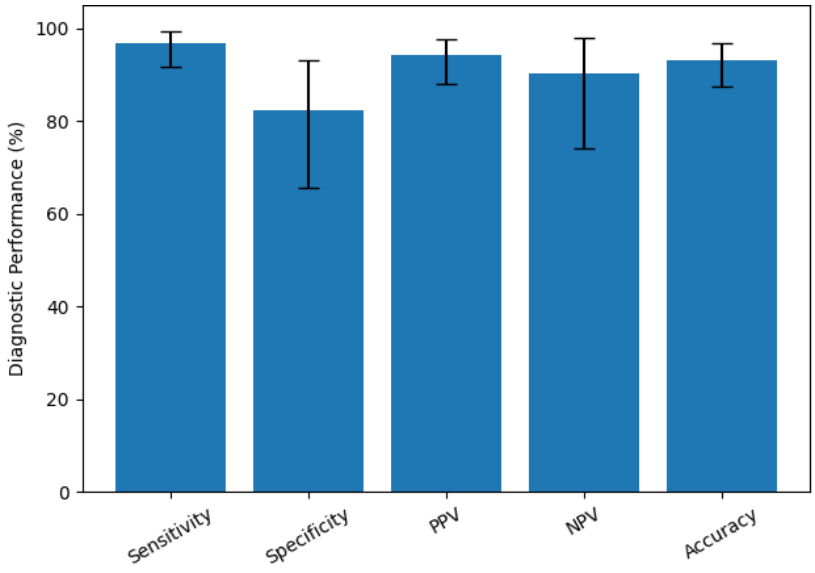

Figure 1 Diagnostic Accuracy Profile of Ophthal-360 With 95% Confidence Intervals

The figure demonstrates a high-sensitivity screening profile for Ophthal-360, with sensitivity reaching 97.0% (95% CI: 91.8–99.4), markedly exceeding specificity at 82.4% (95% CI: 65.5–

93.2). The narrow confidence interval around sensitivity compared with the broader interval for specificity reflects greater precision in identifying DR-positive cases than DR-negative cases, a pattern consistent with screening-oriented algorithm calibration. The positive predictive value remains robust at 94.2% (95% CI: 88.1–97.6), supported by the high observed disease prevalence (77.6%), while the negative predictive value is 90.3% (95% CI: 74.2–98.0), indicating strong reassurance when the AI classifies a case as disease-free. Overall diagnostic accuracy is 93.3% (95% CI: 87.5–96.9), and the asymmetric confidence dispersion across metrics highlights a clinically meaningful trade-off: the algorithm prioritizes minimizing false negatives (3/104 cases) at the expense of a modest false-positive rate (5/30 cases). This performance gradient aligns with public health screening principles where maximizing case detection is critical to prevent progression to vision-threatening stages.

DISCUSSION

This cross-sectional diagnostic validation study demonstrates that Ophthal-360 exhibits high sensitivity (97.0%; 95% CI: 91.8–99.4) and strong overall diagnostic accuracy (93.3%; 95% CI: 87.5–96.9) for detecting non-proliferative diabetic retinopathy (NPDR) in a tertiary-care outpatient population. The observed area under the ROC curve (AUC = 0.897; 95% CI: 0.832–0.962) indicates excellent discrimination between DR-positive and DR-negative cases, while the weighted agreement in severity grading ($\kappa = 0.86$) reflects strong concordance with ophthalmologist-based classification. Importantly, only 3 of 104 reference-standard–positive cases were misclassified as negative (false negatives), yielding a very low miss rate of 2.9%, a clinically critical parameter in screening contexts where failure to detect disease may result in delayed intervention and irreversible visual morbidity.

The algorithm demonstrated a sensitivity–specificity trade-off characteristic of screening-optimized AI systems. While sensitivity was high at 97.0%, specificity was comparatively lower at 82.4% (95% CI: 65.5–93.2), corresponding to 5 false-positive classifications among 30 DR-negative individuals. In population screening programs, particularly in high-prevalence settings such as Pakistan where DR affects approximately 28.78% of diabetics (4), prioritizing sensitivity over specificity is generally acceptable because the clinical consequence of missed disease outweighs the burden of additional confirmatory examinations. The positive predictive value of 94.2% observed in this cohort is influenced by the relatively high DR prevalence (77.6% in this clinic-based sample), whereas the negative predictive value of 90.3% reinforces the algorithm's reliability in excluding disease when the output is negative.

When contextualized within existing literature, the diagnostic performance of Ophthal-360 in this study is comparable to, and in some parameters exceeds, previously reported AI systems. Meta-analytic evidence suggests that AI-based fundus image analysis demonstrates pooled sensitivities above 90% for referable DR, although specificity varies substantially across models and populations (7). The IDx-DR system has reported sensitivity of approximately 87.2% and specificity of 90.7% in prospective multicenter evaluation (12), whereas the SELENA deep learning model achieved sensitivity and specificity exceeding 97% for vision-threatening DR in a multiethnic population (13). EyeArt demonstrated sensitivity above 95%, but specificity varied between 54% and 85% across trials (14). The present findings place Ophthal-360 within the upper range of sensitivity performance, with specificity consistent with real-world screening algorithms. Such performance supports its potential applicability in high-burden regions where maximizing detection is essential.

Beyond binary detection, accurate severity grading is crucial for triage and referral prioritization. In this study, the ordinal distribution of NPDR severity showed minimal

absolute differences between AI and human grading (≤2.3% across categories), and the absence of statistically significant distributional differences (p = 0.338) further supports grading alignment. The strong weighted kappa coefficient (κ = 0.86) indicates near-perfect agreement according to established interpretation thresholds (10). This is clinically relevant because referral decisions often depend on severity categorization rather than mere presence of disease. Moderate and severe NPDR constituted 73.2% of the cohort by human grading, and Ophthal-360 demonstrated consistent classification within these higher-risk strata, suggesting utility in prioritizing patients requiring specialist evaluation.

The clinical implications of these findings are particularly pertinent in resource-constrained healthcare systems. Pakistan faces a substantial diabetes burden (4), and documented gaps in referral pathways and primary-care detection contribute to delayed management (5). AI-assisted screening tools that are camera-agnostic and deployable at primary or secondary care levels could reduce unnecessary tertiary referrals while ensuring that high-risk patients are identified promptly. Modeling studies have suggested that integration of AI into DR screening pathways can reduce ophthalmologist workload without compromising detection of sight-threatening disease (20). The high sensitivity observed in this study aligns with that strategic objective.

Several considerations warrant careful interpretation. First, this was a single-center study conducted in a tertiary-care setting with a relatively high DR prevalence, which may inflate PPV and limit generalizability to community screening environments where prevalence is lower. Second, proliferative diabetic retinopathy (PDR) cases were excluded due to algorithmic constraints; therefore, the findings apply specifically to NPDR detection and grading. Given that PDR is a major cause of irreversible blindness (15), future validation incorporating proliferative stages will be essential. Third, the algorithm relies on macula-centered 40-degree images and does not incorporate peripheral retinal assessment. Emerging evidence highlights the role of peripheral retinal lesions in predicting DR progression and influencing management decisions (18,19). Validation in ultra-widefield imaging contexts may enhance staging precision and longitudinal risk stratification.

Despite these limitations, methodological safeguards—including consecutive sampling, masked grading, standardized imaging protocols, and predefined statistical analysis—strengthen internal validity. The diagnostic performance profile observed here supports the potential of Ophthal-360 as a screening adjunct rather than a replacement for specialist evaluation. In screening paradigms, the acceptable balance often favors high sensitivity with moderate specificity, and the 2.9% false-negative rate observed in this cohort is clinically reassuring.

In summary, Ophthal-360 demonstrated high sensitivity, strong agreement with ophthalmologist grading, and robust overall discrimination for NPDR detection in a tertiary-care Pakistani population. These findings contribute locally validated evidence to the growing body of AI-based ophthalmic diagnostics and support the feasibility of integrating assistive AI tools into DR screening workflows to facilitate earlier detection and referral. Future multicenter studies with broader disease spectra, inclusion of proliferative stages, and evaluation in lower-prevalence community settings are warranted to confirm scalability and health-system impact.

CONCLUSION

Ophthal-360 demonstrated high diagnostic sensitivity (97.0%) and strong agreement with ophthalmologist grading (κ = 0.86) for the detection and severity classification of non-proliferative diabetic retinopathy in a tertiary-care Pakistani population. With an overall

diagnostic accuracy of 93.3% and an AUC of 0.897, the tool exhibits excellent discriminatory capacity while maintaining a clinically acceptable specificity of 82.4%, consistent with screening-oriented calibration that prioritizes minimizing false negatives. The minimal discordance in severity grading and low false-negative rate (2.9%) further support its potential as a reliable assistive screening modality. Integration of Ophthal-360 into structured diabetic retinopathy screening pathways may facilitate earlier detection, optimize referral efficiency, and reduce the burden on specialist services, particularly in high-prevalence, resource-constrained settings. Broader multicenter validation incorporating proliferative stages and community-based populations is warranted to confirm scalability and real-world impact.

REFERENCES

1. Zegeye AF, Temachu YZ, Mekonnen CK. Prevalence and factors associated with diabetic retinopathy among type 2 diabetic patients at Northwest Amhara Comprehensive Specialized Hospitals, Northwest Ethiopia 2021. BMC Ophthalmol. 2023;23(1).

2. Jayanegara AP, Lestari YD. The validity of artificial intelligence (AI) for diabetic retinopathy screening in Asia: a systematic review. Int J Retina (IJRETINA). 2023;6(2).

3. Uy H, Fielding C, Hohlfeld A, Ochodo E, Opare A, Mukonda E, et al. Diagnostic test accuracy of artificial intelligence in screening for referable diabetic retinopathy in real-world settings: a systematic review and meta-analysis. PLOS Glob Public Health. 2023;3(9):e0002160.

4. Mumtaz SN, Fahim MF, Arslan M, Shaikh SA, Kazi U, Memon MS. Prevalence of diabetic retinopathy in Pakistan: a systematic review. Pak J Med Sci. 2018;34:493–500.

5. Anwar MS, Shakaib B, Akhtar W, Yusufzai E, Zehra M, Munawar H, et al. Knowledge and practices of primary care physicians on the current referral system of diabetic retinopathy in Islamabad and Rawalpindi, Pakistan. Int J Ophthalmol. 2019;12(12):1885–92.

6. Yang WH, Bo Z, Mao-Nian W, Shao-Jun Z, Fang-Qin F, Ming W, et al. An evaluation system of fundus photograph-based intelligent diagnostic technology for diabetic retinopathy and applicability for research.

7. Wang Z, Li Z, Li K, Mu S, Zhou X, Di Y. Performance of artificial intelligence in diabetic retinopathy screening: a systematic review and meta-analysis of prospective studies. Front Endocrinol (Lausanne). 2023;14.

8. Bajwa A, Nosheen N, Talpur KI, Akram S. A prospective study on diabetic retinopathy detection based on modified convolutional neural network using fundus images at Sindh Institute of Ophthalmology & Visual Sciences. Diagnostics (Basel). 2023;13(3).

9. Yang Z, Tan TE, Shao Y, Wong TY, Li X. Classification of diabetic retinopathy: past, present and future. Front Endocrinol (Lausanne). 2022;13.

10. Parmar UPS, Surico PL, Singh RB, Romano F, Salati C, Spadea L, et al. Artificial intelligence (AI) for early diagnosis of retinal diseases. Medicina (Kaunas). 2024;60.

11. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA. 2013;310(20):2191–4.

12. U.S. Food and Drug Administration. K213037. Silver Spring (MD): FDA; 2022.

13. Ting DSW, Cheung CYL, Lim G, Tan GSW, Quang ND, Gan A, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. JAMA. 2017;318(22).

14. Bellemo V, Lim ZW, Lim G, Nguyen QD, Xie Y, Yip MYT, et al. Artificial intelligence using deep learning to screen for referable and vision-threatening diabetic retinopathy in Africa: a clinical validation study. Lancet Digit Health. 2019;1(1):e35–44.

15. Wong TY, Cheung CMG, Larsen M, Sharma S, Simó R. Diabetic retinopathy. Nat Rev Dis Primers. 2016;2:16012.

16. Everett LA, Paulus YM. Laser therapy in the treatment of diabetic retinopathy and diabetic macular edema. Curr Diab Rep. 2021;21.

17. Kropp M, Golubnitschaja O, Mazurakova A, Koklesova L, Sargheini N, Vo TTKS, et al. Diabetic retinopathy as the leading cause of blindness and early predictor of cascading complications—risks and mitigation. EPMA J. 2023;14:21–42.

18. Zhang Z, Deng C, Paulus YM. Advances in structural and functional retinal imaging and biomarkers for early detection of diabetic retinopathy. Biomedicines. 2024;12(7):1405.

19. Sharma P, Shareef I, Kalaw FGP, Kako RN, Lin A, Alex V, et al. Prevalence of peripheral retinal findings in retinal patients using ultra-widefield pseudocolor fundus imaging. Sci Rep. 2023;13(1).

20. 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.

DECLARATIONS

Ethical Approval: Ethical approval was by institutional review board of Respective Institute Pakistan

Informed Consent: Informed Consent was taken from participants.

Authors' Contributions:

Concept: MAA; Design: MM; Data Collection: MAA; Analysis: MAA, MM; Drafting: MAA, MM, TMM

Conflict of Interest: The authors declare no conflict of interest.

Funding: This research received no external funding.

Data Availability: The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Acknowledgments: NA

Study Registration: Not applicable.