Clinical Validation of an Artificial Intelligence–Driven Diabetic Retinopathy Screening System: Evidence from a Large Multicenter Study in India

S. logeshwaran[1], S.Jeeva Lakshmi[2], P. Pandia Vadivu[3]
[1]*Executive Director & Chief Duty Medical Officer, The Horizon Medical Clinics, British Columbia, Canada,*
[2]*MDS student, CSI Dental college, Madurai, Tamil Nadu*
[3]*Associate Professor, Tamil Nadu Open University*

***Abstract*—Diabetic retinopathy (DR) remains one of the foremost causes of preventable visual impairment and blindness globally, with a disproportionately high burden in low- and middle-income countries such as India. Despite strong evidence that early detection and timely intervention can substantially reduce vision loss, large-scale DR screening is hindered by shortages of trained ophthalmologists, infrastructural constraints, and limited access to specialist eye care, particularly in rural and underserved regions. To overcome these barriers, an AI-Driven Diabetic Retinopathy Screening System (AIDRSS) was developed using advanced deep-learning algorithms for automated detection and grading of DR from retinal fundus images. This article presents findings from a large multicenter validation study conducted across diverse Indian healthcare settings, evaluating the diagnostic accuracy, clinical utility, and scalability of AIDRSS. The system demonstrated high diagnostic performance, achieving 92% sensitivity and 88% specificity, with excellent detection of referable diabetic retinopathy. The results underscore the potential of AI-enabled screening tools to facilitate early diagnosis, optimize referral pathways, and support population-level eye care programs in resource-constrained environments. The study highlights the role of artificial intelligence as a transformative approach for strengthening diabetic eye care delivery and reducing preventable blindness in India.**

***Index Terms*—Artificial intelligence; Diabetic retinopathy; Automated screening; Deep learning; Retinal fundus imaging; Multicenter validation; Resource-limited settings; Public health ophthalmology; Digital health; India**

I. INTRODUCTION

Diabetic retinopathy (DR) is one of the most common microvascular complications of diabetes mellitus and remains a leading cause of preventable blindness among working-age adults worldwide. According to the International Diabetes Federation, over 537 million adults globally are living with diabetes, a number projected to rise to 643 million by 2030, substantially increasing the global burden of diabetic eye disease. Recent meta-analyses estimate that approximately one in three individuals with diabetes develops some form of diabetic retinopathy, and nearly one in ten experiences vision-threatening DR, underscoring the urgency of scalable screening strategies (Yau et al., 2012; Teo et al., 2021).

In India, the public health challenge is particularly pronounced. Epidemiological studies indicate that diabetic retinopathy affects approximately 13–18% of the general population, with prevalence rising to 35–40% among individuals with poor glycemic control or long-standing diabetes. Given India's rapidly expanding diabetic population and marked urban–rural disparities in healthcare access, DR represents a critical yet underdiagnosed cause of visual impairment. Although early identification and timely intervention such as laser photocoagulation, intravitreal therapy, and metabolic control can reduce the risk of severe vision loss by up to 90%, population-level screening remains inadequate due to shortages of trained ophthalmologists, uneven distribution of retinal services, and dependence on manual image grading.

Conventional DR screening programs rely on fundus photography followed by expert interpretation, a process that is labor-intensive, time-consuming, and difficult to scale in low-resource settings. The World Health Organization has repeatedly emphasized the need for task-shifting and technology-enabled screening approaches to strengthen eye care delivery within primary healthcare systems, particularly in low- and middle-income countries. In this context, artificial intelligence (AI) particularly deep learning has emerged as a transformative tool capable of automating retinal image analysis with accuracy comparable to that of human experts.

Globally, several landmark studies have demonstrated the feasibility and effectiveness of AI-based DR screening. Gulshan et al. (2016) and Ting et al. (2017) independently validated deep learning algorithms achieving high sensitivity and specificity for DR detection across multiethnic populations. Subsequent real-world implementations, including FDA-approved systems such as IDx-DR and national screening pilots in the United Kingdom and Singapore, have provided further evidence that AI can safely augment clinical decision-making and expand screening coverage without increasing specialist workload.

Building on these international advancements, the AI-Driven Diabetic Retinopathy Screening System (AIDRSS) was developed to address the specific epidemiological, infrastructural, and demographic challenges of the Indian healthcare landscape. The system employs a deep learning architecture trained on a large and diverse dataset of retinal fundus images and is designed to automatically detect and grade diabetic retinopathy according to the International Clinical Diabetic Retinopathy (ICDR) severity scale. Unlike earlier single-center or retrospective evaluations, AIDRSS underwent validation through a large multicenter study, enabling robust assessment across heterogeneous imaging conditions, patient profiles, and healthcare settings.

By systematically evaluating diagnostic performance, clinical applicability, and scalability, this study positions AIDRSS within the growing global movement toward AI-enabled population eye health, while providing contextually relevant evidence for its integration into national diabetic care and blindness-prevention programs in India and comparable resource-constrained settings.

II. RESEARCH METHODS

Study Design and Setting

A multicenter, cross-sectional diagnostic accuracy study was conducted in Kolkata, India, to evaluate the performance of an Artificial Intelligence–Driven Diabetic Retinopathy Screening System (AIDRSS). The study was designed in accordance with the Standards for Reporting Diagnostic Accuracy Studies (STARD) and adhered to internationally accepted principles for AI evaluation in healthcare. Multiple healthcare facilities representing varied clinical environments and patient demographics were included to ensure broad generalizability and robustness of findings.

Study Population

A total of 5,029 adult participants with diagnosed diabetes mellitus were enrolled through consecutive sampling. Inclusion criteria comprised individuals aged ≥18 years with confirmed diabetes who underwent routine retinal screening. Exclusion criteria included poor-quality images unsuitable for grading, prior retinal surgery, or coexisting ocular pathologies that could confound diabetic retinopathy (DR) assessment. Each participant contributed images from both eyes, resulting in a dataset of 10,058 macula-centric retinal fundus images.

Image Acquisition and Quality Control

Retinal fundus images were acquired using standardized non-mydriatic fundus cameras operated by trained technicians. Images were captured under consistent illumination and focus protocols to minimize variability. A centralized quality assurance process was implemented to assess image clarity, field definition, and artifact presence. Images failing predefined quality thresholds were excluded from analysis to ensure reliability of both human and algorithmic grading.

AI System Architecture and Image Preprocessing

The AIDRSS employs a deep learning architecture based on convolutional neural networks (CNNs), optimized for retinal image analysis. Prior to inference, images underwent automated preprocessing to enhance diagnostic features. This included Contrast Limited Adaptive Histogram Equalization (CLAHE) to improve local contrast, noise reduction algorithms to suppress imaging artifacts, and normalization techniques to standardize color and illumination variability across centers.

The algorithm was trained on a large, diverse retinal image dataset independent of the validation cohort, ensuring unbiased performance assessment during this study. The system outputs both binary (DR present/absent) and multi-class severity predictions.

Reference Standard and Grading Protocol

Diabetic retinopathy grading followed the International Clinical Diabetic Retinopathy (ICDR) Severity Scale, categorizing images as:

- DR0: No diabetic retinopathy
- DR1: Mild non-proliferative DR
- DR2: Moderate non-proliferative DR
- DR3: Severe non-proliferative DR
- DR4: Proliferative DR

Each image was independently evaluated by experienced retina specialists who were masked to AI outputs. Discrepancies were resolved through consensus discussion. These expert assessments served as the gold standard reference against which AIDRSS predictions were compared.

Outcome Measures

The primary outcomes were diagnostic sensitivity (true positive rate) and specificity (true negative rate) for detection of any diabetic retinopathy. Secondary analyses included performance for referable diabetic retinopathy (DR3–DR4), which requires urgent ophthalmic intervention. Additional metrics such as positive predictive value, negative predictive value, and confusion matrices were examined to provide a comprehensive assessment of diagnostic performance.

Statistical Analysis

Statistical analyses were conducted at the image level using standard diagnostic accuracy frameworks. Sensitivity and specificity were calculated with corresponding 95% confidence intervals. Agreement between AIDRSS and expert graders was assessed using inter-rater reliability measures where applicable. Subgroup analyses examined algorithm performance across varying image qualities and demographic characteristics to evaluate consistency and robustness.

Ethical Considerations

The study protocol received approval from the appropriate institutional ethics committees. All procedures complied with the Declaration of Helsinki and national ethical guidelines for biomedical research. Written informed consent was obtained from all participants prior to enrollment. Data were anonymized before analysis to ensure patient confidentiality and data security.

Reproducibility and Transparency

To support transparency and reproducibility, system architecture details, evaluation metrics, and validation protocols were documented in accordance with emerging best practices for AI research reporting. The methodology aligns with recommendations from international regulatory and academic bodies for clinical deployment of AI-based diagnostic systems.

III. RESULTS

Diagnostic Performance of AIDRSS

The AI-Driven Diabetic Retinopathy Screening System (AIDRSS) demonstrated robust and clinically meaningful diagnostic performance when evaluated against expert retina specialist grading across the multicenter dataset. For the detection of any diabetic retinopathy, the system achieved an overall sensitivity of 92% and an overall specificity of 88%, indicating a high capacity to correctly identify individuals with disease while maintaining a low false-positive rate.

These findings reflect the algorithm's strong discriminative ability across a heterogeneous population and varied imaging conditions, reinforcing its suitability for deployment in real-world screening environments. The high specificity is particularly relevant for minimizing unnecessary referrals and reducing the burden on already constrained ophthalmic services.

Detection of Referable Diabetic Retinopathy

Notably, AIDRSS achieved 100% sensitivity for referable diabetic retinopathy (DR3–DR4), ensuring that all cases requiring urgent ophthalmic evaluation were correctly identified. This level of performance is of critical clinical importance, as missed diagnoses at this stage are strongly associated with irreversible vision loss. The system's flawless detection of sight-threatening disease underscores its potential role as a reliable front-line screening tool capable of prioritizing high-risk patients for timely specialist care.

Population-Level Implications

The algorithm's consistent performance across multiple centers highlights its scalability and robustness in diverse clinical settings, including resource-limited environments where access to ophthalmologists is limited. By accurately distinguishing non-referable from referable disease, AIDRSS supports efficient triage, optimizes referral pathways, and enables the redistribution of specialist resources toward patients with advanced disease.

In addition to validating the AI system, the study reaffirmed the substantial prevalence of diabetic retinopathy among individuals with elevated blood glucose levels, consistent with existing epidemiological evidence. This finding further emphasizes the necessity of integrating systematic retinal screening into routine diabetes management programs to mitigate long-term visual morbidity.

Summary of Key Findings

Overall, the results demonstrate that AIDRSS meets and, in critical aspects, exceeds performance thresholds recommended for population-based diabetic retinopathy screening. Its high sensitivity, particularly for referable disease, combined with strong specificity, positions the system as a safe, effective, and scalable solution for early detection of diabetic retinopathy, with significant implications for reducing preventable blindness in India and comparable low- and middle-income settings.

IV. DISCUSSION

This multicenter validation study provides compelling evidence for the clinical utility and public health relevance of the AI-Driven Diabetic Retinopathy Screening System (AIDRSS) in addressing one of the most pressing challenges in diabetes-related eye care early detection of diabetic retinopathy (DR) at scale. The findings underscore the transformative potential of artificial intelligence to strengthen screening pathways, optimize resource utilization, and reduce preventable vision loss, particularly in resource-constrained healthcare settings.

Diagnostic Accuracy and Scalability

The high diagnostic accuracy demonstrated by AIDRSS, with 92% sensitivity and 88% specificity, indicates that the system performs at a level comparable to expert human graders for initial DR screening. Most notably, the system achieved 100% sensitivity for referable diabetic retinopathy, ensuring that all sight-threatening cases were correctly identified. This is a critical safety benchmark for screening tools, as false negatives at advanced disease stages carry significant clinical and ethical implications.

The consistency of performance across a large, multicenter cohort highlights the robustness and generalizability of the algorithm across diverse patient populations, imaging conditions, and healthcare environments. These attributes are essential for real-world implementation and support the feasibility of deploying AIDRSS as a front-line screening solution capable of handling large screening volumes without compromising diagnostic reliability.

Relevance to Resource-Limited Settings

A major strength of AIDRSS lies in its applicability to low- and middle-income countries (LMICs), where the burden of diabetes is rapidly increasing while access to ophthalmic specialists remains limited. In India, ophthalmologist-to-population ratios are particularly unfavorable in rural and peri-urban regions, resulting in delayed diagnosis and advanced disease presentation.

By enabling automated interpretation of retinal images, AIDRSS allows trained primary healthcare providers, nurses, and community health workers to conduct initial DR screening without requiring specialist expertise. This task-shifting approach aligns with World Health Organization recommendations for strengthening primary eye care and has the potential to substantially expand screening coverage, reduce delays in referral, and prevent progression to irreversible vision loss.

Integration into Public Health and Diabetes Care Programs

The study's findings support the integration of AI-enabled DR screening into national diabetes control and blindness prevention programs. Given the strong association between poor glycemic control and DR prevalence reaffirmed in this study, embedding AIDRSS within routine diabetes care pathways could facilitate systematic, opportunistic screening and early risk stratification.

Furthermore, integration with digital health platforms and mobile screening units could enhance outreach to remote and underserved populations. Linking AI-based screening outputs with electronic health records and referral networks may further streamline care coordination, improve follow-up adherence, and support data-driven public health planning.

Alignment with Emerging AI Innovations in Ophthalmology
The performance of AIDRSS is consistent with and complementary to other emerging AI-based DR screening solutions developed and validated both globally and within India. Initiatives such as MadhuNETrAI, an AI-powered mobile screening application developed by AIIMS in collaboration with Wadhwani AI, have reported high detection accuracies and are undergoing real-world deployment. Collectively, these innovations signal a rapid maturation of AI technologies in ophthalmology and a growing evidence base supporting their safety, effectiveness, and scalability.
The convergence of multiple validated AI systems strengthens confidence in AI-enabled screening as a sustainable component of future eye care delivery models, particularly in countries facing disproportionate disease burden and workforce constraints.

Implications and Future Directions
While the present study establishes strong diagnostic validity, future research should focus on longitudinal outcomes, cost-effectiveness analyses, and implementation science evaluations to assess long-term impact on visual morbidity and health system efficiency. Regulatory frameworks, clinician acceptance, patient trust, and ethical governance will also play critical roles in determining successful large-scale adoption.

V. CONCLUSION OF DISCUSSION

Overall, the findings demonstrate that AIDRSS represents a meaningful advancement in diabetic retinopathy screening by combining high diagnostic accuracy with scalability and real-world feasibility. Its ability to reliably identify sight-threatening disease highlights its potential to transform population-level diabetic eye care, address critical gaps in access to specialist services, and support global initiatives aimed at preventing avoidable blindness through early detection and timely intervention.

Conclusion
The AI-Driven Diabetic Retinopathy Screening System (AIDRSS) demonstrates strong clinical validity and scalability, offering an effective solution for early detection of diabetic retinopathy in resource-limited settings. With high sensitivity and specificity, and complete detection of referable disease, AIDRSS can support large-scale screening, improve timely referrals, and contribute to reducing preventable vision loss. Its integration into routine diabetes care and public health programs holds significant promise for strengthening eye care delivery and improving population health outcomes. AIDRSS offers a reliable and scalable solution for diabetic retinopathy screening, demonstrating high diagnostic accuracy and complete detection of referable disease. Its implementation can enhance early diagnosis, improve access to eye care, and support efforts to reduce preventable blindness in resource-limited settings.

REFERENCES

[1] Dey, A. K., Walia, P., Somvanshi, G., et al. (2025). AI-driven diabetic retinopathy screening: Multicentric validation of AIDRSS in India. arXiv. https://arxiv.org/
[2] Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., et al. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA, 316(22), 2402–2410. https://doi.org/10.1001/jama.2016.17216
[3] International Diabetes Federation. (2023). IDF diabetes atlas (10th ed.). Brussels, Belgium: Author.
[4] Li, Z., Keel, S., Liu, C., He, Y., Meng, W., Scheetz, J., et al. (2018). An automated grading system for detection of vision-threatening diabetic retinopathy on retinal images. JAMA Ophthalmology, 136(4), 351–358. https://doi.org/10.1001/jamaophthalmol.2017.6197
[5] Rajaraman, S., Antani, S. K., Poostchi, M., Silamut, K., Hossain, M. A., Maude, R. J., &

Thoma, G. R. (2019). Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images. PeerJ, 7, e6290.

[6] Teo, Z. L., Tham, Y. C., Yu, M., Chee, M. L., Rim, T. H., Cheung, N., et al. (2021). Global prevalence of diabetic retinopathy and projection of burden through 2045: Systematic review and meta-analysis. Ophthalmology, 128(11), 1580–1591. https://doi.org/10.1016/j.ophtha.2021.04.027

[7] Ting, D. S. W., Cheung, C. Y., Lim, G., Tan, G. S. W., Quang, N. D., Gan, A., et al. (2017). Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations. JAMA, 318(22), 2211–2223. https://doi.org/10.1001/jama.2017.18152

[8] World Health Organization. (2019). World report on vision. Geneva: WHO Press.

[9] World Health Organization. (2022). WHO guideline on the use of digital technologies to support eye care. Geneva: World Health Organization.

[10] Wadhwani AI & All India Institute of Medical Sciences. (2025). Evaluation of MadhuNETrAI: An AI-based screening tool for diabetic retinopathy [Press release]. The Economic Times.

[11] AI for Sight Initiative. (2025). Public health implications of artificial intelligence–based diabetic retinopathy screening in India. Devdiscourse.

[12] Yau, J. W. Y., Rogers, S. L., Kawasaki, R., Lamoureux, E. L., Kowalski, J. W., Bek, T., et al. (2012). Global prevalence and major risk factors of diabetic retinopathy. Diabetes Care, 35(3), 556–564. https://doi.org/10.2337/dc11-1909