https://doi.org/10.12662/interagir.v21.6139

Artigo de Revisão

Artificial intelligence in the diagnosis of diabetic retinopathy: a systematic review of diagnostic accuracy and clinical impact (2015–2025)

ABSTRACT

Objectives: This systematic review evaluated the diagnostic accuracy and clinical impact of artificial intelligence (AI) systems used for the early detection of diabetic retinopathy (DR) from 2015 to 2025, emphasizing performance, applicability, and implications for large-scale screening. **Methods:** Following the PRISMA 2020 guidelines, searches were conducted in PubMed, Scopus, Web of Science, CINAHL, and Google Scholar. Eligible studies included observational designs, clinical trials, and external validation studies using deep learning algorithms to identify DR in adults with diabetes. Exclusion criteria covered case reports, narrative reviews, studies without accuracy metrics, and those with overlapping datasets. Data were organized in Excel, analyzed with RevMan 5.4, and evaluated using QUADAS-2 and GRADE. **Results:** Of the 2,000 records initially selected, 14 studies met the inclusion criteria, involving more than 50,000 participants. Most employed convolutional neural networks, reporting sensitivities of 86% to 99% and specificities of 89% to 99%. Standalone systems, such as ARDA and IDx-DR, demonstrated specialist-level accuracy, often achieving area under the curve values above 0.95. Multicenter studies confirmed feasibility in real-world primary care and teleophthalmology settings. **Conclusion:** AI systems exhibit high accuracy and operational scalability for DR screening, supporting broader access and earlier diagnosis. However, methodological heterogeneity highlights the need for standardized reporting, diverse population datasets, and ongoing algorithmic monitoring. These findings position AI as a promising component of global strategies to reduce preventable vision loss.

Keywords: diabetic retinopathy; telediagnosis; artificial intelligence; deep learning.

Gabriela Hilgemberg
MD, School of Medicine, Pontifical Catholic University of Paraná (PUC-PR).
ORCID: https://orcid.org/0009-0002-0684-4522.
Email: hilgemberggabriela@gmail.com

Mac Gayver da Silva Castro, PhD
Doctor of Medical Surgical-Sciences, School of Medicine, Federal University of Ceará (FAMED-UFC). Professor, Christus University Center (Unichristus).
ORCID: https://orcid.org/0000-0002-6251-4694.
Email: mac.castro@unichristus.edu.br

Autor correspondente:
Mac Gayver da Silva Castro
E-mail: mac.castro@unichristus.edu.br

Submetido em: 05/11/2025
Aprovado em: 08/01/2026
Publicado em: 24/11/2026

Como citar este artigo: HILGEMBERG; Gabriela; CASTRO, Mac Gayver da Silva. Artificial intelligence in the diagnosis of diabetic retinopathy: a systematic review of diagnostic accuracy and clinical impact (2015–2025). **Revista Interagir**, Fortaleza, v. 21, p. 1-4, 2026. DOI: https://doi.org/10.12662/interagir.v21.6139

Licença: Este artigo é publicado sob a licença Creative Commons Atribuição 4.0 Internacional (CC BY 4.0) – https://creativecommons.org/licenses/by/4.0/deed.pt_BR. © Os autores, 2026.

1 INTRODUCTION

Diabetic retinopathy (DR) is the leading cause of preventable blindness in working-age adults. Traditional screening – based on manual interpretation of fundus images – is expensive, subjective, and limited by a shortage of specialists (Ting *et al.*, 2017). Artificial intelligence (AI) achieves sensitivities and specificities above 90%, rivaling specialists (Bellemo *et al.*, 2019).

2 METHODS

This review followed the PRISMA 2020 and PRISMA-DTA guidelines and was conducted in PubMed, Scopus, Web of Science, CINAHL, and Google Scholar for studies published between January 2015 and September 2025. The search combined three conceptual blocks: diabetic retinopathy, artificial intelligence, and diagnostic processes, using AND between conceptual groups and OR within each group.

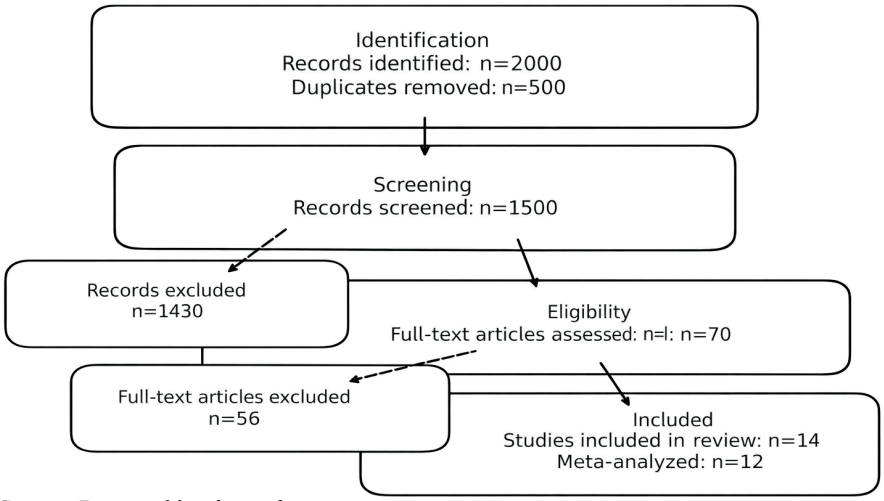
We included original studies that applied AI or deep learning to diagnose DR in adults using fundus photography, OCT, or OCTA that reported at least one diagnostic accuracy metric. We excluded non-original reports, studies without extractable accuracy data, duplicate datasets, non-human experiments, and articles without full text from 2015 to 2025.

Two reviewers independently selected titles/abstracts and full texts, extracted data using a standardized form, and assessed quality with QUADAS-2; disagreements were resolved through discussion or by a third reviewer. Of the 14 eligible studies, 12 underwent random-effects meta-analysis in RevMan 5.4, while the others were synthesized narratively.

3 RESULTS

The search identified 2,000 records; 500 duplicates were removed. After screening 1,500 titles/abstracts and 70 full texts, 14 studies were included.

Figure 1 - PRISMA-DTA flow diagram of study selection.



Source: Prepared by the authors.

The main characteristics are shown in Table 1.

Table 1 - Summary of characteristics: sample size, AI model, and performance metrics. Accuracy values indicate each system's ability to detect referenceable RD using specialized classification.

Author	N	AI	Sens. (%)	Spec. (%)	AUC
Brant *et al.*, 2025	4,357	ARDA	97	96.4	98
Bhuiyan *et al.*, 2022	88,702	CNN	97.6	99.4	99.8
Ruamviboonsuk *et al.*, 2022	7,651	CNN	91.4	95.4	94.7
Zhang *et al.*, 2022	1,089	CNN	92.9	93.3	95.8
Lee *et al.*, 2021	311,604	CNN	80.5	93.6	NR
Ryu *et al.*, 2021	301	CNN	97	99	91.9
Ipp *et al.*, 2021	893	CNN	95.5	89	92
Heisler *et al.*, 2020	380	CNN	90.4	93.3	NR
Bellemo *et al.*, 2019	1,574	CNN	95	89	97.3
Abràmoff *et al.*, 2018	819	IDx	87.2	90.7	NR
De Fauw *et al.*, 2018	14,884	CNN	94.5	94	99.2
Li *et al.*, 2018	71,043	CNN	97	91.4	98.9
Ting *et al.*, 2017	494,661	CNN	98.3	NR	NR
Gulshan *et al.*, 2016	128,175	Inception v3	96	93	97.4

Abbreviations: AI = artificial intelligence; Sens = sensitivity; Spec = specificity; AUC = area under the curve; ARDA = Automated Retinal Disease Assessment; CNN = convolutional neural network; IDx = IDx-DR system; NR = not reported.
Source: Prepared by the authors.

4 DISCUSSION

This review shows that AI systems for DR screening perform comparably to human experts. CNN-based models exhibit high sensitivity and specificity, with AUCs above 0.95 (Ting *et al.*, 2017). High sensitivity may increase false positives, reinforcing two-step workflows in which AI performs screening and human evaluators review uncertain cases (Lee *et al.*,

2021). In Brazil, validated AI integrated with primary care and teleophthalmology could improve the prioritization of referrals in the SUS. Limitations include the predominance of diagnostic accuracy studies, scarce evidence on clinical impact and cost-effectiveness, and methodological heterogeneity that may bias combined results (Abràmoff *et al.*, 2018). Many models rely on selected datasets, raising concerns about publication bias and limited generalizability (Lee *et al.*, 2021). Future research should prioritize multimodal and explainable AI, along with implementation and economic studies – especially in middle-income settings such as Brazil.

5 CONCLUSION

Current evidence shows that AI is a reliable and scalable tool for automated DR screening. In Brazil, adoption within the SUS requires local validation, clear pathways, trained staff, and attention to ethical and regulatory issues.

REFERENCES

ABRÀMOFF, M. D. *et al.* Pivotal clinical trial of an AI-based autonomous diagnostic system for detecting diabetic retinopathy in primary care settings. **Digital Medicine**, v. 1, n. 39, 2018. DOI: 10.1038/s41746-018-0040-6. Available at: https://www.nature.com/articles/s41746-018-0040-6. Access at: 18 March 2025.

BELLEMO, V. *et al.* Artificial intelligence using deep learning to screen for referable and vision-threatening diabetic retinopathy in Africa: a clinical validation study. **The Lancet Digital Health**, v. 1, n. 1, p. e35–e44, 2019. DOI: 10.1016/S2589-7500(19)30004-4. Available at: https://www.thelancet.com/journals/landig/article/PIIS2589-7500(19)30004-4/fulltext. Access at: 2 feb. 2025.

BHUIYAN, A. *et al.* Automated screening of diabetic retinopathy for primary care settings using deep learning. **Intelligence-Based Medicine**, v. 5, art. 100045, 2021. DOI: 10.1016/j.ibmed.2021.100045. Available at: https://pubmed.ncbi.nlm.nih.gov/35528965/. Access at: 11 Jan. 2025

BRANT, A. *et al.* Performance of a deep learning algorithm for diabetic retinopathy in India. **JAMA Network Open**, v. 8, n. 3, e250984, 2025. DOI: 10.1001/jamanetworkopen.2025.0984. Available at: https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2831702. Access at: 21 Apr. 2025.

DE FAUW, J. *et al.* Clinically applicable deep learning for diagnosis and referral in retinal diseases. **Nature Medicine**, v. 24, n. 9, p. 1342–1350, 2018. DOI: 10.1038/s41591-018-0107-6. Available at: https://www.nature.com/articles/s41591-018-0107-6. Access at: 9 Dec. 2024.

GULSHAN, V. *et al.* Development and validation of a deep learning algorithm for detecting diabetic retinopathy in retinal fundus photographs. **JAMA Network Open**, v. 316, n. 22, p. 2402–2410, 2016. DOI: 10.1001/jama.2016.17216. Available at: https://jamanetwork.com/journals/jama/fullarticle/2588763. Access at: 17 Oct. 2024

HEISLER, M. *et al.* Joint deep learning for diabetic retinopathy detection using optical coherence tomography angiography. **Translational Vision Science & Technology**, v. 9, n. 2, art. 20, 2020. DOI: 10.1167/tvst.9.2.20. Available at: https://pubmed.ncbi.nlm.nih.gov/32818081/. Access at: 28 Nov. 2024.

IPP, E. *et al.* Critical evaluation of an artificial intelligence system for autonomous detection of referable and vision-threatening diabetic retinopathy. **JAMA Network Open**, v. 4, n. 11, e2134254, 2021. DOI: 10.1001/jamanetworkopen.2021.34254. Available at: https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2786132. Access at: 6 Mar. 2025.

LEE, A. Y. *et al.* Multicenter, comparative, real-world validation study of seven automated artificial intelligence systems for screening diabetic retinopathy. **Diabetes Care**, v. 44, n. 5, p. 1168–1175, 2021. DOI: 10.2337/dc20-1877. Available at: https://diabetesjournals.org/care/article/44/5/1168/138752/Multicenter-Head-to-Head-Real-World-Validation. Access at: 14 Feb. 2025.

LI, Z. *et al.* An automated classification system for detecting vision-threatening referable diabetic retinopathy based on color fundus photographs. **Diabetes Care**, v. 41, n. 12, p. 2509–2516, 2018. DOI: 10.2337/dc18-0147. Available at: https://pubmed.ncbi.nlm.nih.gov/30275284/. Access at: 30 Oct. 2024.

RUAMVIBOONSUK, P. *et al.* Real-time screening for diabetic retinopathy using deep learning in a national multisite screening program: a prospective interventional cohort study. **The Lancet Digital Health**, v. 4, n. 4, p. e235–e244, 2022. DOI: 10.1016/S2589-7500(22)00017-6. Available at: https://pubmed.ncbi.nlm.nih.gov/35272972/. Access at: 19 May 2025.

RYU, G. *et al.* A deep learning model to identify diabetic retinopathy using optical coherence tomography angiography. **Scientific Reports**, v. 11, n. 1, art. 23024, 2021. DOI: 10.1038/s41598-021-02479-6. Available at: https://www.nature.com/articles/s41598-021-02479-6. Access at: 8 Jan. 2025.

TING, D. S. W. *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 Network Open**, v. 318, n. 22, p. 2211–2223, 2017. DOI: 10.1001/jama.2017.18152. Available at:

https://jamanetwork.com/journals/jama/fullarticle/2665775. Access at: 22 Sept. 2024.

ZHANG, G. *et al.* Automated multidimensional deep learning platform for referable diabetic retinopathy detection: a retrospective multicenter study. **BMJ Open**, v. 12, n. 7, e060155, 2022. DOI: 10.1136/bmjopen-2021-060155. Available at: https://bmjopen.bmj.com/content/12/7/e060155. Access at: 3 June 2025.