Advancing healthcare with artificial intelligence: diagnostic accuracy of machine learning algorithm in diagnosis of diabetic retinopathy in the Brazilian population

Mateus A. dos Reis[1,2*], Cristiano A. Künas[3], Thiago da Silva Araújo[3], Josiane Schneiders[1], Pietro B. de Azevedo[2], Luis F. Nakayama[4,5], Dimitris R. V. Rados[1,6], Roberto N. Umpierre[6,7], Otávio Berwanger[8], Daniel Lavinsky[1,9], Fernando K. Malerbi[4], Philippe O. A. Navaux[3] and Beatriz D. Schaan[1,10,11]

Abstract

Background In healthcare systems in general, access to diabetic retinopathy (DR) screening is limited. Artificial intelligence has the potential to increase care delivery. Therefore, we trained and evaluated the diagnostic accuracy of a machine learning algorithm for automated detection of DR.

Methods We included color fundus photographs from individuals from 4 databases (primary and specialized care settings), excluding uninterpretable images. The datasets consist of images from Brazilian patients, which differs from previous work. This modification allows for a more tailored application of the model to Brazilian patients, ensuring that the nuances and characteristics of this specific population are adequately captured. The sample was fractionated in training (70%) and testing (30%) samples. A convolutional neural network was trained for image classification. The reference test was the combined decision from three ophthalmologists. The sensitivity, specificity, and area under the ROC curve of the algorithm for detecting referable DR (moderate non-proliferative DR; severe non-proliferative DR; proliferative DR and/or clinically significant macular edema) were estimated.

Results A total of 15,816 images (4590 patients) were included. The overall prevalence of any degree of DR was 26.5%. Compared with human evaluators (manual method of diagnosing DR performed by an ophthalmologist), the deep learning algorithm achieved an area under the ROC curve of 0.98 (95% CI 0.97–0.98), with a specificity of 94.6% (95% CI 93.8–95.3) and a sensitivity of 93.5% (95% CI 92.2–94.9) at the point of greatest efficiency to detect referable DR.

Conclusions A large database showed that this deep learning algorithm was accurate in detecting referable DR. This finding aids to universal healthcare systems like Brazil, optimizing screening processes and can serve as a tool for improving DR screening, making it more agile and expanding care access.

Keywords Diabetic retinopathy, Artificial intelligence, diabetes mellitus

*Correspondence:
Mateus A. dos Reis
mateusaugustodosreis@gmail.com
Full list of author information is available at the end of the article

Background

The prevalence of type 2 diabetes has been increasing steadily in recent years and, by 2021, about 537 million people are estimated to have this condition [1] and by 2050, more than 1.31 billion (1.22–1.39) people are projected to have diabetes [2]. In Brazil, 12% of the population is diagnosed with diabetes [1, 3], making it the 6th country with the highest number of adults with diabetes in the world [1]. Diabetes is associated with microvascular and macrovascular complications, among which diabetic retinopathy (DR) stands out [4]. DR is the leading cause of blindness in working-age individuals, with a prevalence in people with diabetes of 22.2% worldwide [5] and 36.3% in Brazil [6], which may contribute to reduced quality of life depending on the severity of the disease [7].

The retina is the deepest layer of texture covering the rear of the eye, recorded by color fundus photographs (CFPs) or fundoscopy. Vessel detection and segmentation are essential in DR diagnosis [8], since the initial changes are asymptomatic [9]. Although early diagnosis and therapy can prevent severe vision loss in 90% of cases, only a small part of these patients are screened at the recommended frequency [10]. In Brazil, a study with patients with type 2 diabetes showed that only 11.5% of individuals treated under the Family Health Program, 14.9% treated in basic health units, and 35% in tertiary care had undergone fundoscopy in the last year [11]. On the other hand, another recently published study showed a screening rate of 63.9% in patients with type 1 diabetes [12]. The manual method of diagnosing DR is performed by an ophthalmologist who examines the human retinal fundus image. This manual process is very consuming in terms of the time and experience of an expert ophthalmologist, which makes developing an automated method to aid in the diagnosis of DR an essential and urgent need [13]. The lack of ophthalmologists to serve the number of patients who need to be screened is a limitation to this process [14].

Artificial intelligence is transforming ophthalmology and has been leveraged for CFP to accomplish core tasks including segmentation, classification, prediction and seem to be an alternative to solve this problem [15, 16], since the method has been tested on other populations and has shown good accuracy [17–20]. Combined with remote retinography, algorithm-based screening improves the process, selecting only higher-risk individuals to in person evaluation with an ophthalmologist [21, 22]. The deep learning system developed by Abramoff et al. has obtained a US Food and Drug Administration approval for the diagnosis of DR. It was evaluated in a prospective, although observational setting, achieving 87.2% sensitivity and 90.7% specificity [23].

Training these algorithms with diverse populations and large datasets is important to avoid biases, since previous studies suggest that the contrast between the retinal fundus and DR lesions can vary considerably between different ethnicities [24, 25]. Also, reproducibility (in different populations) is central to increasing the confidence in the scientific findings [26]. In Brazil, few studies evaluated the use of artificial intelligence to diagnose DR [27, 28].

The aim of this study is to train and evaluate the diagnostic performance of an algorithm, compared to the gold standard (evaluation by an ophthalmologist), in DR screening, using a large dataset of CFPs of Brazilian subjects, ensuring that the nuances and characteristics of this specific population are adequately captured.

Methods

Datasets

This study sample combined prospective and retrospective data from individuals from three public and academic institutions who applied for a CFP. A convenience sample with consecutive patients was used. Each dataset has individual characteristics, as follows:

- Endocrinology Unit of a tertiary public hospital in Porto Alegre (HCPA – Hospital de Clínicas de Porto Alegre): data from 2019 to 2021 were collected in the specialized hospital ambulatory setting prospectively and included patients more than seven years-old. The camera used was Canon CR-2 (Canon Inc., Melville, NY, USA) and two images of the posterior segment of each eye - one centered on the macula and the other centered on the disc (45° field of view) - after mydriasis induced by tropicamide 1% eye drops were captured.
- TeleOftalmo project from Rio Grande do Sul combines the state's health authority and the Federal University telehealth initiative (TelessaúdeRS-UFRGS) and provides remote ophthalmological care to primary care patients [29, 30]. Retrospective data from individuals older than eight years-old with and without diabetes were collected from seven cities between 2020 and 2021. The camera used was ZEISS Visucam (Oberkocjen, Germany) and after patients with diabetes underwent mydriasis with tropicamide 1% eye drops, posterior segment images (30° photo centered on the disc and 45° photo centered on the macula) and photos of the nasal, temporal, superior, and inferior sector were captured.
- Ophthalmology Unit of the specialized ambulatory service from Federal University of São Paulo (UNIFESP) provides retrospective data from patients older than 10 years with diabetes between 2010 and 2020 [31]. The camera used was Canon CR-2 (Canon Inc., Melville, NY, USA) and Nikon NF-505 (Tokyo,

Japan) and an image of posterior segment of each eye centered on the macula with mydriasis induced by tropicamide 0.5% was captured.

At all three sites, the images were collected by a nurse technician or researcher who had been trained in the procedure.

The study protocol was approved by the institutional Research Ethics Committees (2019−0113) and conducted in accordance with the Declaration of Helsinki. Data from all sets were de-identified and were in full compliance with the General Data Protection Law and local regulations.

The study is reported according to the Standards for Reporting of Diagnostic Accuracy Studies (STARD) [32] [see Additional file 1].

Grading

After anonymization, all images were assessed for quality by the ophthalmologists and the images with interference factors, such as overexposure, inadequate focus, insufficient lighting, and excessive artifacts, were discarded. In the tertiary hospital of Porto Alegre in 2019 and Tele-Oftalmo, before the images were analyzed by the ophthalmologists, the EyeQ algorithm [33], which performs retinal image quality assessment, was used to select only images of sufficient quality.

The CFPs were classified into the following five categories: no retinopathy, mild non-proliferative retinopathy, moderate non-proliferative retinopathy, severe non-proliferative retinopathy, and proliferative retinopathy or macular edema, according to the International Clinical Diabetic Retinopathy (ICDR) scale [34]. There are many DR classifications applied in distinct countries and screening programs, with the ICDR scale as the most applied in open-access ophthalmological datasets [35].

The classification of DR was done by imaging. No information other than eye images was available to the evaluator. This classification was performed independently by two ophthalmologists (reference standard). Clinical information and reference standard results were not available to the assessors of the reference standard. After grading, the images were divided into two groups, non-referable DR (no DR; mild non-proliferative DR and no macular edema or clinically insignificant macular edema) and referable DR (moderate non-proliferative DR; severe non-proliferative DR; proliferative DR and/or clinically significant macular edema). The kappa value for agreement between two ophthalmologists for referable DR and non-referable DR was 0.813 (near perfect). All disagreements were adjudicated by a third grader.

Neural network model

The goal of the neural network was to produce a binary prediction for each image: non-referable and referable DR.

For initial access to the feasibility of using a machine learning algorithm (index test) as a screening tool, 10 neural networks were trained. The codebase was inspired by Voets et al. [36] and the initial model was previously described by this study group [37]. The model used the Inception v3 architecture as a pre-trained model and was adapted and fine-tuned for our datasets. This approach leverages the knowledge encoded in the pre-trained model to solve a different but related problem more efficiently than training a new model from scratch. The same neural network architecture was used as in the original study by Gulshan et al. [15] and its replication by Voets et al. [36]. This consistency allows for a direct comparison with previous studies, highlighting that the primary differences lie in the data used and the population studied. The network was initialized with weights from the ImageNet dataset for all layers [38], except for the fully connected layer on top, which received training. After loading the weights, the fully connected dense layer with two units was added using the Sigmoid activation function [39].

Adam optimizer [40] was used for learning network weights during the training process. The initial value of the learning rate was 0.01 and the end value was 0.0037. The binary cross-entropy function was used as the loss function to estimate the logarithmic loss between the actual and predicted labels.

All images were processed in line with Voets [36], locating the center and radius of the eye fundus and resizing each image to 299×299 pixels, with the center of the fundus in the middle of the image. This resizing was done to match the default input size for the Inception V3 model. While larger image dimensions could be used, this would significantly increase training time. This resizing is particularly relevant for high resolution images, as compressing them to 299×299 pixels may result in some loss of details. However, this trade-off is necessary to maintain compatibility with the pre-trained Inception V3 architecture and to ensure efficient training. The algorithm was trained on each set for 200 epochs with batch size of 8. The datasets of the tertiary hospital in Porto Alegre, UNIFESP, and TeleOftalmo were used. The model was trained with each dataset separately and with grouped datasets into a single set. Each dataset was divided into two subsets—training and testing—in a stratified 70%/30% ratio.

Clinical information and reference standard results were not available to the performers/readers of the index test.

Statistical analysis

The distribution of demographic and clinical characteristics was presented using mean±standard deviation (SD) for continuous variables and by number and percentage for categorical data. For referable/non-referable DR classification, we reported the area under the receiver operating characteristic (ROC) curve, sensitivity and specificity. The sensitivity was calculated as the number of correctly predicted positive examples divided by the total number positive examples. The specificity was calculated as the number of correctly predicted negative examples divided by the total number negative examples. The performance of the algorithm in detecting referable/non-referable DR was measured by the area under the ROC curve generated by plotting sensitivity versus 1 *minus* specificity. Based on the two operating points, 2×2 tables were developed to characterize the sensitivity and specificity of the algorithm in relation to the gold standard which was defined as the majority decision of the ophthalmologists' readings. The 95% confidence intervals were calculated. Three operational cutoff points were used to evaluate the performance of the algorithm: with high specificity, with greater sensitivity and maximum gain point. We excluded from analysis uninterpretable images due to poor quality. We use mean-value imputation to address missing data (age) in baseline characteristics. This study used a convenience sample, with no sample calculation conducted. Statistical significance was set at $P<.05$.

Results

We used images of 5308 individuals and included 4590 patients (15816 images) in the analysis after quality evaluation, of which 4191 (26.5%) were classified as referable. Figure 1 shows the flow of participants during the study.

Table 1 presents the characteristics of the gradable images by dataset and by grouped dataset. In grouped dataset, the median age was 60 (7–97) years, 9684 (61.2%) were women and 4191 (26.5%) of the CFPs presented referable DR.

The set of images of the tertiary hospital in Porto Alegre in 2019 included 651 patients and 3626 CFPs [2522 fully gradable; 641 (25.4%) referable]. The set of the same hospital in 2021 had 412 patients and 1857 CFPs [1555 fully gradable; 422 (27.1%) referable]. The TeleOftalmo set included 2993 patients and 47,299 CFPs [9160 fully gradable; 2578 (28.1%) referable], and the UNIFESP set had 1252 patients and 2579 CFPs [2579 fully gradable; 550 (21.3%) referable].

Figure 2 shows the performance of the algorithm in detecting referable DR for the gradable images of all sets.
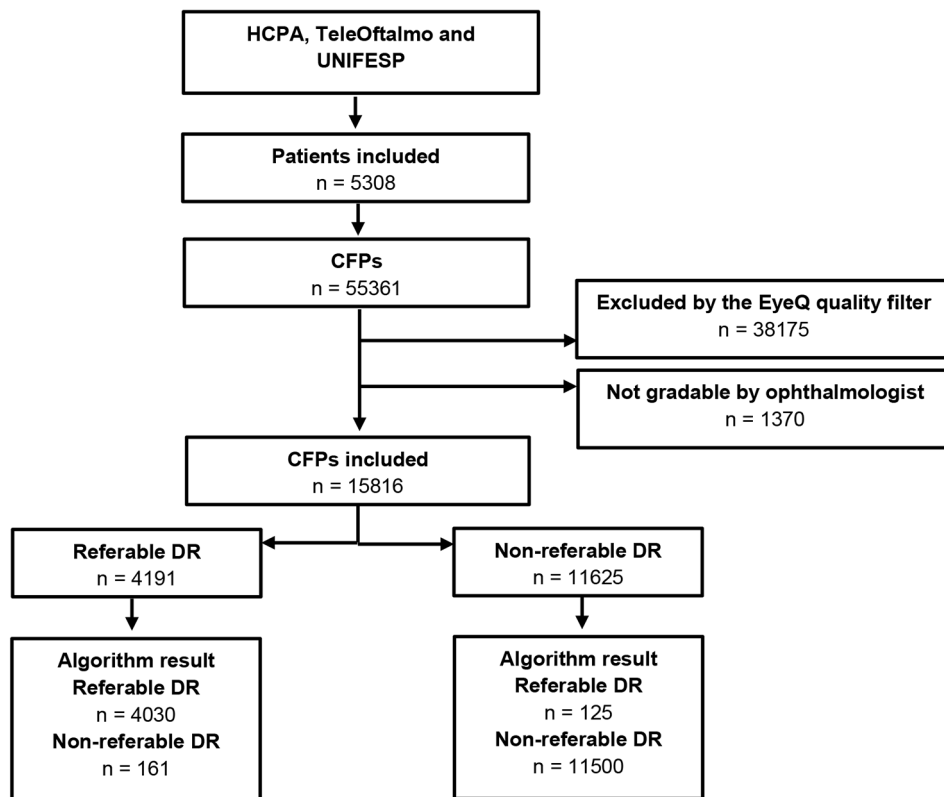

Fig.1 Standards for reporting of diagnostic accuracy studies (STARD) diagram for the algorithm output: referable diabetic retinopathy. DR: Diabetic Retinopathy; HCPA: Hospital de Clínicas de Porto Alegre; UNIFESP: Federal University of São Paulo; CFPs: Color Fundus Photographs

Table 1 Baseline characteristics

	Tertiary hospital in Porto Alegre 2019	Tertiary hospital in Porto Alegre 2021	TeleOftalmo	UNIFESP	All
Patients	651	412	2993	1252	5308
Total CFPs	3636	1857	47,299	2579	55,371
Included CFPs	2522 (69.3)	1555 (83.7)	9160 (19.3)	2579 (100)	15,816 (28.5)
Age (years-old) among included images	56 (12–97)	56 (7–87)	60 (8–96)	64 (10–92)	60 (7–97)
Women among included images	1556 (61.7)	905 (58.2)	5655 (61.7)	1568 (60.8)	9684 (61.2)
Disease severity distribution classified by majority decision of ophthalmologists (reference standard) in included CFPs					
No DR	1573 (62.4)	975 (62.7)	5651 (61.7)	1922 (74.5)	10,121 (63.9)
Mild non-proliferative DR	328 (13.0)	177 (11.4)	999 (10.9)	107 (4.1)	1611 (10.9)
Moderate non-proliferative DR	373 (14.8)	221 (14.2)	1993 (21.7)	181 (7.0)	2768 (17.5)
Severe non-proliferative DR	52 (2.0)	51 (3.3)	185 (2.0)	127 (4.9)	415 (2.6)
Proliferative DR	196 (7.8)	131 (8.4)	332 (3.6)	242 (9.4)	901 (5.7)
Clinically significant edema	153 (6.0)	119 (7.6)	965 (10.5)	275 (10.6)	1512 (9.5)
Referable DR	641 (25.4)	422 (27.1)	2578 (28.1)	550 (21.3)	4191 (26.5)

DR: diabetic retinopathy; CFPs: color fundus photographs; the data are presented as n or n (%)
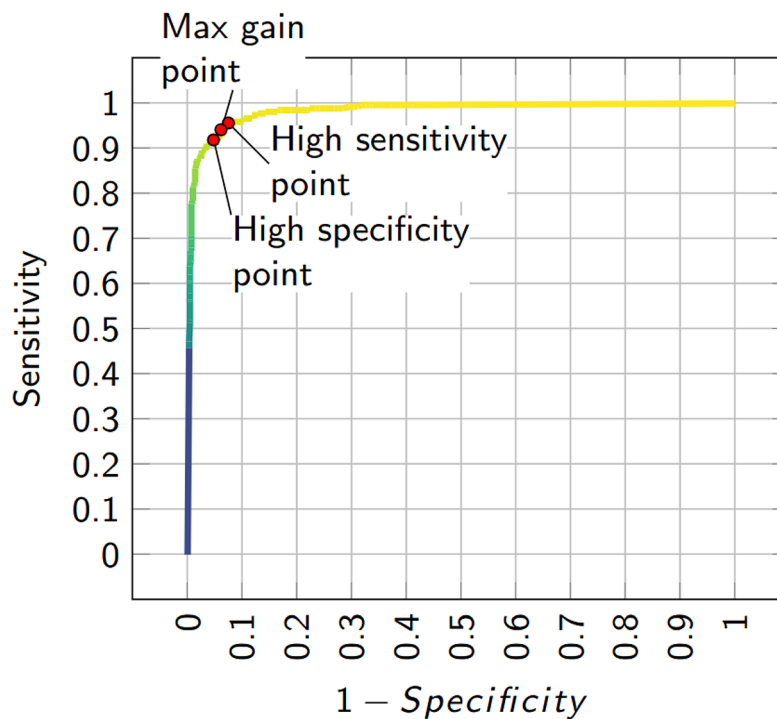
AUC = 97.9% - The high specificity point reaches 95.8% specificity and 90.8% sensitivity. The maximum gain point reaches 94.6% specificity and 93.5% sensitivity. The high sensitivity point reaches 90.9% specificity and 95.1% sensitivity.

Fig. 2 Performance of the algorithm in detecting referable diabetic retinopathy for the gradable images of all sets

Using the operational cut-off point with high specificity, the sensitivity was 90.8% (95% CI 89.2–92.4), and the specificity was 95.8% (95% CI 95.2–96.5). The point with the highest sensitivity, showing an output that would be used for a screening tool, had a sensitivity of 95.1% (95% CI 93.9–96.3) and specificity of 90.9% (95% CI 89.9–91.8). The maximum gain point had a specificity of 94.6% (95% CI 93.8–95.3) and a sensitivity of 93.5% (95% CI 92.2–94.9). Additional files 2–5 show the performance of

the algorithm in detecting referable DR for the gradable images in each set evaluated.

Table 2 presents the ROC curve, sensitivity, and specificity of each image set evaluated. The algorithm achieved an area under the ROC curve of 0.98 (0.97–0.98) in grouped dataset.

No significant adverse events occurred from performing the index test or the reference standard.

Discussion

The present results showed excellent accuracy, with similarly high sensitivity and specificity, suggesting that machine learning may be an alternative to improve the workflow of DR screening in other settings other than those already tested. Sensitivity in diagnosing a disease is the most important metric of a screening test, thus, a screening program requires high sensitivity values (>80%) [41]. However, having both, high sensitivity and high specificity is the best of both worlds: our study showed a sensitivity of 95.1% at the high sensitivity point and 90.8% at the high specificity point. A previous study undertaken with Brazilian individuals for the clinical validation of an artificial intelligence algorithm, for example, had reached a high sensitivity (97.8%) but a lower specificity (61.4%) [28]. Another study also reached high sensitivity and specificity values, but with a different methodology, as images obtained exclusively with a portable retinal camera were assessed [42]. The present study provides a comprehensive evaluation of the use of a machine learning algorithm for the detection of referable DR, making it the study in Brazil with the largest image set, the better results and with all parts of the training and validation processes performed on images of Brazilian subjects published to date.

Artificial intelligence by machine learning has become a tool that helps in reading and analyzing images [43]. Many groups around the world had previously studied the automated and semi-automated evaluation of DR [18, 28, 44–48]. DR screening is a challenge in many countries [49, 50], including Brazil, which is a continent-sized country with screening rates far below the ideal [14, 15]. The development of an algorithm with images of the Brazilian population can greatly increase the scope of DR screening, allowing only patients with changes in the

initial examination to be referred to an ophthalmologist, reducing the number of referrals, so that a greater number of patients have access to specialized services, when necessary. This process, compared with human classification, can reduce screening costs when deployed as semi-autonomous screening [21, 49, 50]. In countries with structured DR screening programs, medium-term results showed a reduction in amaurosis in the population with diabetes [11].

Our study presented values close to studies performed with other populations, with a sensitivity of 95.1% and specificity of 90.9% at the high sensitivity point. The meta-analysis conducted by Wu et al. [51] showed robust performance in detecting different categories of DR, with a pooled sensitivity of 93–97% and a pooled specificity of 90–98%.

When evaluated individually, the sets of the tertiary hospital in Porto Alegre in 2019 and 2021 and UNIFESP presented lower specificity in the performance of the algorithm compared with TeleOftalmo, probably because the TeleOftalmo set is larger and the algorithm may be better fitted to its population. Studies evaluating the performance of deep learning have shown that the initial dataset for algorithm development should be large and have diverse training data in terms of patient demographics and ethnicity, image acquisition methods, and image quality [44], which is exactly what we did in the present study.

Our image set included populations from different Brazilian states, different regions of Rio Grande do Sul, and different types of cameras for performing CFP with different imaging characteristics, which was a strength of this study. These characteristics increase the external validity of our results and qualify them for future use in screening strategies for this important complication of diabetes at the national level.

This study has limitations. First, the ophthalmologists classified macular edema considering the presence of hard exudates, microaneurysms, or hemorrhage in the macular region, and studies have shown that optical coherence tomography can detect earlier changes in the vascular and retinal morphology, making it an important tool for the management and follow-up of retinal diseases, such as macular edema [52, 53]. However, the high

Table 2 Area under the ROC curve, sensitivity, and specificity of each image set evaluated

	Tertiary hospital in Porto Alegre 2019	Tertiary hospital in Porto Alegre 2021	TeleOftalmo	UNIFESP	All
Area under the ROC curve, (95% CI)	0.92 (0.89–0.95)	0.95 (0.92–0.97)	0.98 (0.98–0.99)	0.94 (0.92–0.97)	0.98 (0.97–0.98)
Sensitivity (%) (95% CI)	95.3 (92.4–98.3)	95.3 (91.7–99.0)	95.2 (93.7–96.7)	95.3 (92.0–98.5)	95.1 (93.9–96.3)
Specificity (%) (95% CI)	56.7 (52.7–60.8)	72.4 (67.7–77.2)	92.6 (91.5–93.7)	80.5 (77.3–83.6)	90.9 (89.9–91.9)

ROC: receiver operating characteristic; data are presented as percentage and confidence intervals (CIs)

cost of this method (equipment and human workforce) makes its use unfeasible in a continental and low/middle-income country such as Brazil. Second, the clinical information of patients was limited, which hinders further analysis of the population and its associations with the findings presented. Third, we included only images with good quality in the training and testing set. Evaluating the performance of the algorithm in a real-life study, without selecting good quality images, is necessary.

Conclusion

This study showed that deep machine learning algorithm is reproducible in a diverse population, with elevated sensitivity and specificity for the detection of referable DR. This classification is useful in universal healthcare systems (such as the Brazilian example) as a tool to improve the screening flow of DR in the country. Further research is needed to determine the real-life applicability of the algorithm, its cost-effectiveness to assess whether it can improve the care of patients with diabetes, and the regulation of its use on a large scale.

Abbreviations

CFP	Color fundus photograph
CIs	Confidence intervals
DR	Diabetic retinopathy
UNIFESP	Federal University of São Paulo
HCPA	Hospital de Clínicas de Porto Alegre
ICDR	International Clinical Diabetic Retinopathy
ROC	Receiver operating characteristic
SD	Standard deviation
STARD	Standards for Reporting of Diagnostic Accuracy Studies

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13098-024-01447-0.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Supplementary Material 5

Author contributions
Reis, MA: study concept, methodology, data curation, writing of the original draft, and editingSchneiders, J; Azevedo, PB; Nakayama, LF; Malerbi, FK: methodologyLavinsky, D: study concept, methodologyKunas, CA; Araujo, TS; Navaux, POA: data analysis and interpretationUmpierre, RN; Berwanger, O; Rados, DRV: revising and editing of the articleSchaan, BD: study concept, supervision, and writing, revising, and editing of the articleReis, MA: takes responsibility for the integrity of the data and the accuracy of the data analysisAll authors reviewed the manuscript.

Data availability
The study dataset is available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate
The study protocol was approved by the institutional Research Ethics Committees (2019−0113) and conducted in accordance with the Declaration of Helsinki. Data from all sets were de-identified and were in full compliance with the General Data Protection Law and local regulations.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Graduate Program in Medical Sciences: Endocrinology, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil
[2]Universidade Feevale, Novo Hamburgo, RS, Brazil
[3]Institute of Informatics, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil
[4]Department of Ophthalmology and Visual Sciences, Universidade Federal de São Paulo, São Paulo, Brazil
[5]Laboratory for Computational Physiology, Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, MA, USA
[6]TelessaúdeRS Project, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil
[7]Department of Social Medicine, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil
[8]The George Institute for Global Health, Imperial College London, London, UK
[9]Department of Ophthalmology, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil
[10]Institute for Health Technology Assessment (IATS) – CNPq, Porto Alegre, Brazil
[11]Endocrinology Unit, Hospital de Clínicas de Porto Alegre, Porto Alegre, RS, Brazil

References
1. Sun H, Saeedi P, Karuranga S, Pinkepank M, Ogurtsova K, Duncan BB, et al. IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. Diabetes Res Clin Pract. 2022;183:109119.
2. GBD 2021 Diabetes Collaborators. Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the global burden of Disease Study 2021 [published correction appears in Lancet. 2023;402(10408):1132. Doi: 10.1016/S0140-6736(23)02044-5]. Lancet. 2023;402(10397):203–34. https://doi.org/10.1016/S0140-6736(23)01301-6.
3. Telo GH, Cureau FV, de Souza MS, Andrade TS, Copês F, Schaan BD. Prevalence of diabetes in Brazil over time: a systematic review with meta-analysis. Diabetol Metab Syndr. 2016;8:65.

Reis dos *et al. Diabetology & Metabolic Syndrome*   (2024) 16:209

Page 8 of 9

4. Committee ADAPP, American Diabetes Association Professional Practice Committee. 12. Retinopathy, Neuropathy, and Foot Care: *Standards of Medical Care in Diabetes—2022*. Diabetes Care. 2022.pp.S185–S194. https://doi.org/10.2337/dc22-s012

5. Teo ZL, Tham Y-C, Yu M, Chee ML, Rim TH, Cheung N, et al. Global prevalence of Diabetic Retinopathy and Projection of Burden through 2045. Ophthalmology. 2021;1580–91. https://doi.org/10.1016/j.ophtha.2021.04.027.

6. Chagas TA, Dos Reis MA, Leivas G, Santos LP, Gossenheimer AN, Melo GB, et al. Prevalence of diabetic retinopathy in Brazil: a systematic review with meta-analysis. Diabetol Metab Syndr. 2023;15:34.

7. Ben ÂJ, de Souza CF, Locatelli F, Rosses APO, Szortika A, de Araujo AL, et al. Health-related quality of life associated with diabetic retinopathy in patients at a public primary care service in southern Brazil. Arch Endocrinol Metab. 2021;64:575–83.

8. Rehman A, Harouni M, Karimi M, Saba T, Bahaj SA, Awan MJ. Microscopic retinal blood vessels detection and segmentation using support vector machine and K-nearest neighbors. Microsc Res Tech. 2022;85(5):1899–914. https://doi.org/10.1002/jemt.24051.

9. Liew G, Michaelides M, Bunce C. A comparison of the causes of blindness certifications in England and Wales in working age adults (16–64 years), 1999–2000 with 2009–2010. BMJ Open. 2014;4:e004015.

10. Flaxel CJ, Adelman RA, Bailey ST, Fawzi A, Lim JI, Vemulakonda GA, et al. Diabet Retinopathy Preferred Pract Pattern® Ophthalmol. 2020;127:P66–145.

11. Schneiders J, Telo GH, Bottino LG, Pasinato B, Neyeloff JL, Schaan BD. Quality indicators in type 2 diabetes patient care: analysis per care-complexity level. Diabetol Metab Syndr. 2019;11:34.

12. Foppa L, Alessi J, Nemetz B, et al. Quality of care in patients with type 1 diabetes during the COVID-19 pandemic: a cohort study from Southern Brazil. Diabetol Metab Syndr. 2022;14(1):75. https://doi.org/10.1186/s13098-022-00845-6.

13. Oulhadj M, Riffi J, Khodriss C, et al. Diabetic retinopathy prediction based on vision transformer and modified capsule network. Comput Biol Med. 2024;175:108523. https://doi.org/10.1016/j.compbiomed.2024.108523.

14. Resnikoff S, Lansingh VC, Washburn L, Felch W, Gauthier T-M, Taylor HR, et al. Estimated number of ophthalmologists worldwide (International Council of Ophthalmology update): will we meet the needs? Br J Ophthalmol. 2020;104:588–92.

15. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of Diabetic Retinopathy in Retinal Fundus photographs. JAMA. 2016;316:2402–10.

16. Driban M, Yan A, Selvam A, Ong J, Vupparaboina KK, Chhablani J. Artificial intelligence in chorioretinal pathology through fundoscopy: a comprehensive review. Int J Retina Vitreous. 2024;10(1):36. https://doi.org/10.1186/s40942-024-00554-4. Published 2024 Apr 23.

17. Gargeya R, Leng T. Automated identification of Diabetic Retinopathy using deep learning. Ophthalmology. 2017;124:962–9.

18. Ting DSW, Pasquale LR, Peng L, Campbell JP, Lee AY, Raman R, et al. Artificial intelligence and deep learning in ophthalmology. Br J Ophthalmol. 2019;103:167–75.

19. Gulshan V, Rajan RP, Widner K, Wu D, Wubbels P, Rhodes T, et al. Performance of a deep-learning algorithm vs manual grading for detecting Diabetic Retinopathy in India. JAMA Ophthalmol. 2019;137:987–93.

20. van Der Heijden AA, Abramoff MD, Verbraak F, van Hecke MV, Liem A, Nijpels G. Validation of automated screening for referable diabetic retinopathy with the IDx-DR device in the Hoorn Diabetes Care System. Acta Ophthalmol. 2018;96:63–8.

21. Xie Y, Nguyen QD, Hamzah H, Lim G, Bellemo V, Gunasekeran DV, et al. Artificial intelligence for teleophthalmology-based diabetic retinopathy screening in a national programme: an economic analysis modelling study. Lancet Digit Health. 2020;2:e240–9.

22. Ben ÂJ, Neyeloff JL, de Souza CF, Rosses APO, de Araujo AL, Szortika A, et al. Cost-utility analysis of opportunistic and systematic Diabetic Retinopathy screening strategies from the perspective of the Brazilian Public Healthcare System. Appl Health Econ Health Policy. 2020;18:57–68.

23. Abramoff MD, Lavin PT, Birch M, et al. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. NPJ Digit Med. 2018;1:1–8. https://doi.org/10.1038/s41746-018-0040-6.

24. Ting DSW, Cheung CY-L, 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;318:2211–23.

25. Wong TY, Bressler NM. Artificial Intelligence with Deep Learning Technology looks into Diabetic Retinopathy Screening. JAMA. 2016;2366. https://doi.org/10.1001/jama.2016.17563.

26. Munafò MR, Nosek BA, Bishop DVM, Button KS, Chambers CD, du Sert NP, Simonsohn U, Wagenmakers EJ, Ware JJ, Ioannidis JPA. A manifesto for reproducible science. Nat Hum Behav. 2017;1:0021. https://doi.org/10.1038/s41562-016-0021.

27. Cleland CR, Rwiza J, Evans JR, et al. Artificial intelligence for diabetic retinopathy in low-income and middle-income countries: a scoping review. BMJ Open Diabetes Res Care. 2023;11(4):e003424. https://doi.org/10.1136/bmjdrc-2023-003424.

28. Malerbi FK, Andrade RE, Morales PH, Stuchi JA, Lencione D, de Paulo JV, et al. Diabetic Retinopathy Screening using Artificial Intelligence and Handheld Smartphone-based retinal camera. J Diabetes Sci Technol. 2022;16:716–23.

29. Lutz de Araujo A, Moreira T, de Varvaki Rados C, Gross DR, Molina-Bastos PB, Katz CG. The use of telemedicine to support Brazilian primary care physicians in managing eye conditions: the TeleOftalmo Project. PLoS ONE. 2020;15:e0231034.

30. Gonçalves MR, Umpierre RN, D'Avila OP, Katz N, Mengue SS, Siqueira ACS, et al. Expanding Primary Care Access: a Telehealth Success Story. Ann Fam Med. 2017;15:383.

31. Nakayama LF, Goncalves M, Zago Ribeiro L, Santos H, Ferraz D, Malerbi F et al. A Brazilian Multilabel Ophthalmological Dataset (BRSET). PhysioNet; 2023. doi:https://doi.org/10.13026/XCXW-8198.

32. Šimundić A-M. Measures of diagnostic accuracy: Basic definitions. EJIFCC. 2009;19:203–11.

33. Fu H, Wang B, Shen J, Cui S, Xu Y, Liu J et al. Evaluation of retinal image quality assessment networks in different color-spaces. 2019. https://doi.org/10.48550/ARXIV.1907.05345

34. Wilkinson CP, Ferris FL 3rd, Klein RE, Lee PP, Agardh CD, Davis M, et al. Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. Ophthalmology. 2003;110:1677–82.

35. Nakayama LF, Gonçalves MB, Ferraz DA, Santos HNV, Malerbi FK, Morales PH, et al. The challenge of Diabetic Retinopathy standardization in an ophthalmological dataset. J Diabetes Sci Technol. 2021. https://doi.org/10.1177/19322968211029943.

36. Voets M, Møllersen K, Bongo LA. Reproduction study using public data of: Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. PLoS ONE. 2019;14:e0217541.

37. Moreira F, Schaan B, Schneiders J, Reis M, Serpa M, Navaux P. Impacto Da Resolução na Detecção De Retinopatia Diabética com uso de Deep Learning. Anais Principais do Simpósio Brasileiro De Computação Aplicada à Saúde (SBCAS 2020). Sociedade Brasileira de Computação - SBC; 2020. https://doi.org/10.5753/sbcas.2020.11546.

38. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S et al. ImageNet Large Scale Visual Recognition Challenge. arXiv [cs.CV]. 2014.Available: http://arxiv.org/abs/1409.0575

39. Han J, Moraga C. The influence of the sigmoid function parameters on the speed of backpropagation learning. From Natural to Artificial neural computation. Springer Berlin Heidelberg; 1995. pp. 195–201.

40. Kingma DP, Ba J, Adam. A method for stochastic optimization. 2014. https://doi.org/10.48550/ARXIV.1412.6980

41. Vujosevic S, Aldington SJ, Silva P, Hernández C, Scanlon P, Peto T, et al. Screening for diabetic retinopathy: new perspectives and challenges. Lancet Diabetes Endocrinol. 2020;8:337–47.

42. Malerbi FK, Nakayama LF, Melo GB, et al. Automated identification of different severity levels of Diabetic Retinopathy using a Handheld Fundus Camera and single-image protocol. Ophthalmol Sci. 2024;4(4):100481. https://doi.org/10.1016/j.xops.2024.100481.

43. Haug CJ, Drazen JM. Artificial Intelligence and Machine Learning in Clinical Medicine, 2023. N Engl J Med. 2023;388:1201–8.

44. Grzybowski A, Singhanetr P, Nanegrungsunk O, Ruamviboonsuk P. Artificial Intelligence for Diabetic Retinopathy Screening using Color retinal photographs: from development to Deployment. Ophthalmol Ther. 2023. https://doi.org/10.1007/s40123-023-00691-3.

45. Shah A, Clarida W, Amelon R, Hernaez-Ortega MC, Navea A, Morales-Olivas J, et al. Validation of automated screening for Referable Diabetic Retinopathy with an Autonomous Diagnostic Artificial Intelligence System in a Spanish Population. J Diabetes Sci Technol. 2021;15:655–63.

46. Bellemo V, Lim ZW, Lim G, Nguyen QD, Xie Y, Yip MYT, 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:e35–e44

47. Zhang Y, Shi J, Peng Y, Zhao Z, Zheng Q, Wang Z, Liu K, Jiao S, Qiu K, Zhou Z, Artificial intelligence-enabled screening for diabetic retinopathy: A real-world, multicenter and prospective study. BMJ Open Diabetes Res. Care 2020, 8, e001596.

48. Heydon, Egan P, Bolter C, Chambers L, Anderson R, Aldington J, Stratton S, Scanlon IM, Webster PH, Mann L et al. S.; Prospective evaluation of an artificial intelligence-enabled algorithm for automated diabetic retinopathy screening of 30,000 patients. Br. J. Ophthalmol. 2020, 105, 723–728.

49. Ruamviboonsuk P, Tiwari R, Sayres R, Nganthavee V, Hemarat K, Kongprayoon A, et al. Real-time diabetic retinopathy screening by deep learning in a multisite national screening programme: a prospective interventional cohort study. Lancet Digit Health. 2022;4:e235–44.

50. Hao S, Liu C, Li N, Wu Y, Li D, Gao Q, et al. Clinical evaluation of AI-assisted screening for diabetic retinopathy in rural areas of midwest China. PLoS ONE. 2022;17:e0275983.

51. Tufail A, Rudisill C, Egan C, Kapetanakis VV, Salas-Vega S, Owen CG, et al. Automated Diabetic Retinopathy Image Assessment Software: diagnostic accuracy and cost-effectiveness compared with human graders. Ophthalmology. 2017;124:343–51.

52. Srisubat A, Kittrongsiri K, Sangroongruangsri S, Khemvaranan C, Shreibati JB, Ching J, et al. Cost-utility analysis of deep learning and trained human graders for Diabetic Retinopathy Screening in a Nationwide Program. Ophthalmol Ther. 2023;12:1339–57.

53. Wu J-H, Liu TYA, Hsu W-T, Ho JH-C, Lee C-C. Performance and Limitation of Machine Learning Algorithms for Diabetic Retinopathy Screening: Meta-analysis. J Med Internet Res. 2021;23:e23863.

54. Vujosevic S, Muraca A, Alkabes M, Villani E, Cavarzeran F, Rossetti L, et al. Early microvascular and neural changes in patients with type 1 and type 2 diabetes mellitus without clinical signs of diabetic retinopathy. Retina. 2019;39:435–45.

55. Lavinsky F, Lavinsky D. Novel perspectives on swept-source optical coherence tomography. Int J Retina Vitreous. 2016;2:25.

Publisher's Note