

Automated Identification of Different Severity Levels of Diabetic Retinopathy Using a Handheld Fundus Camera and Single-Image Protocol

Fernando K. Malerbi, MD, PhD,¹ Luis Filipe Nakayama, MD,¹ Gustavo Barreto Melo, MD, PhD,¹ José A. Stuchi, MSC, PhD,² Diego Lencione, MSC,² Paulo V. Prado, MSC,² Lucas Z. Ribeiro, MD,¹ Sergio A. Dib, MD, PhD,¹ Caio V. Regatieri, MD, PhD¹

Purpose: To evaluate the performance of artificial intelligence (AI) systems embedded in a mobile, handheld retinal camera, with a single retinal image protocol, in detecting both diabetic retinopathy (DR) and more-than-mild diabetic retinopathy (mtmDR).

Design: Multicenter cross-sectional diagnostic study, conducted at 3 diabetes care and eye care facilities.

Participants: A total of 327 individuals with diabetes mellitus (type 1 or type 2) underwent a retinal imaging protocol enabling expert reading and automated analysis.

Methods: Participants underwent fundus photographs using a portable retinal camera (Phelcom Eyer). The captured images were automatically analyzed by deep learning algorithms retinal alteration score (RAS) and diabetic retinopathy alteration score (DRAS), consisting of convolutional neural networks trained on EyePACS data sets and fine-tuned using data sets of portable device fundus images. The ground truth was the classification of DR corresponding to adjudicated expert reading, performed by 3 certified ophthalmologists.

Main Outcome Measures: Primary outcome measures included the sensitivity and specificity of the AI system in detecting DR and/or mtmDR using a single-field, macula-centered fundus photograph for each eye, compared with a rigorous clinical reference standard comprising the reading center grading of 2-field imaging protocol using the International Classification of Diabetic Retinopathy severity scale.

Results: Of 327 analyzed patients (mean age, 57.0 ± 16.8 years; mean diabetes duration, 16.3 ± 9.7 years), 307 completed the study protocol. Sensitivity and specificity of the AI system were high in detecting any DR with DRAS (sensitivity, 90.48% [95% confidence interval (CI), 84.99%–94.46%]; specificity, 90.65% [95% CI, 84.54%–94.93%]) and mtmDR with the combination of RAS and DRAS (sensitivity, 90.23% [95% CI, 83.87%–94.69%]; specificity, 85.06% [95% CI, 78.88%–90.00%]). The area under the receiver operating characteristic curve was 0.95 for any DR and 0.89 for mtmDR.

Conclusions: This study showed a high accuracy for the detection of DR in different levels of severity with a single retinal photo per eye in an all-in-one solution, composed of a portable retinal camera powered by AI. Such a strategy holds great potential for increasing coverage rates of screening programs, contributing to prevention of avoidable blindness.

Financial Disclosure(s): F.K.M. is a medical consultant for Phelcom Technologies. J.A.S. is Chief Executive Officer and proprietor of Phelcom Technologies. D.L. is Chief Technology Officer and proprietor of Phelcom Technologies. P.V.P. is an employee at Phelcom Technologies. *Ophthalmology Science* 2024;4:100481 © 2024 by the American Academy of Ophthalmology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).



Supplemental material available at www.ophtalmologyscience.org.

Diabetic retinopathy (DR) meets the criteria for a disease that warrants screening because of its asymptomatic phase before vision loss, its significance as a public health problem, the availability of tests, the existence of established and effective treatment if performed in a timely manner, and improved prognosis with early intervention.^{1,2} However, despite DR being one of the leading causes of vision loss

among adults, having affected > 100 million adults worldwide in 2020, low rates of patients with diabetes undergo the recommended annual DR screening, even in developed countries.^{2,3} Among the causes for such low adherence, socioeconomic and geographic barriers have been identified, as well as delayed referrals from primary care practitioners and inadequate patient education.⁴

Recent modalities, such as telemedicine, artificial intelligence (AI) systems, and portable retinal cameras, have been proposed to increase the uptake rates of DR screening.⁵ Among these, handheld retinal cameras are interesting tools for such tasks, because of their affordability and portability, enabling a wider coverage of screening programs. These devices can be utilized by primary care physicians or nonmedical personnel after proper training.⁶

The role of telemedicine in DR screening is continually expanding; the imaging protocol for DR evaluation has undergone a substantial evolution, from the 7-field ETDRS protocol to the more recent 2-field protocol, currently used for expert reading in many established programs.⁷ However, some barriers for wide-scale implementation of telemedicine DR screening are still present; some examples are the lack of adherence and operational challenges such as poor network access in rural areas.^{8,9} Hence, further work is necessary to continually optimize practices and improve long-term patient outcomes. In that sense, simplifying the imaging protocol holds the potential to increase adherence and improve the effectiveness of screening programs.

Recently, deep learning (DL) algorithms have emerged as a promising tool for the detection of DR, offering diagnostic performance comparable to human experts and the potential to scale up screening programs efficiently.^{10,11} The first such algorithm that received United States Food and Drug Administration (FDA) approval detects more-than-mild DR (mtmDR) because, from an ophthalmic perspective, only patients with mtmDR are currently recommended to receive specialized evaluation and consideration for treatment.¹⁰ However, the presence of DR is consistently associated with other complications of diabetes¹²; thus, the detection of any DR may also be part of a strategy to stratify the systemic risk of patients with diabetes. In that sense, DR screening strategies could be tailored to local circumstances, depending on the desired outcome and several additional factors, such as local epidemiology, available infrastructure, and workforce, allowing for effective implementation.

The aim of this study was to evaluate the performance of AI systems for the detection of both DR and mtmDR on fundus images obtained with a mobile, handheld retinal camera, utilizing a single retinal image protocol.

Methods

Study Design and Population

This was an observational cross-sectional study, in compliance with the tenets of the Declaration of Helsinki. The algorithmic performance was validated using a prospectively collected data set of retinal images from 327 individuals with diabetes enrolled at 3 different centers. One center is a tertiary referral ophthalmological hospital located in São Paulo, Brazil, whereas the other 2 are diabetes centers in São Paulo and Sergipe, Brazil, where DR evaluation is regularly performed. Informed consent was obtained from each patient before the study enrollment. The study was approved by the Research Ethics Committee of the Federal University of São Paulo (33.842.220.7.0000.5505). Inclusion criteria were patients with type 1 or 2 diabetes mellitus who agreed with the study terms. Exclusion criteria were a history of any other

ocular disease that could impair DR classification, such as other maculopathies, uveitis, congenital malformations, media opacities that precluded retinal images of enough quality, and poor patient collaboration for the imaging protocol.

Smartphone-Based Retinal Camera

The Eyer (Phelcom Technologies, LLC) is an FDA-approved smartphone-based camera built using a Samsung Galaxy S10 (Android 13) smartphone. The camera captures retinal fundus photographs with a 45° field angle. With a 12-megapixel sensor, it delivers an image at 1600 × 1600 pixels. It has an autofocus range from −20 to +20 diopters.

Image Acquisition

All participants underwent pharmacological mydriasis through the administration of 2 drops of 0.5% tropicamide at 5-minute intervals, followed by image acquisition, consisting of 2 fundus images per eye, one macula centered and the other one optic disc centered. Images were acquired by previously trained health care professionals who already had experience in ocular imaging, using a standardized protocol.¹³ All images were anonymized, deidentified, and reviewed to ensure the removal of any personal health information.

Grading Protocol

The ground truth was DR classification according to expert reading. Labeling was performed independently by 2 masked, certified ophthalmologists, with a third senior retinal specialist adjudicating in discordant cases. Classification of DR severity was performed according to the International Classification of Diabetic Retinopathy, based on 2 images per eye after mydriasis.¹⁴ Images were deemed gradable for human readers if ≥ 80% of the image area was visible and if the assessment of at least the third retinal vascular branch was possible.¹⁵

Automated Detection of DR

Although the ground truth relied on expert reading using a 2-image protocol, we intended to assess the performance of automated detection of DR using a single image per eye. In order to achieve this goal, we employed 2 different DL systems, along with a combination of both, for the analysis of different severity levels of DR, after automated image quality evaluation. Such quality assessment was performed by a convolutional neural network (CNN), which was previously trained, based on factors including blurriness, exposure, presence of artifacts, media opacities, or incorrect fields of view.¹⁶ Images classified as gradable after the automatic quality evaluation underwent further automatic DR detection by 2 different DL systems, whose specific details are outlined below.

The first DL system utilized in this study for the detection of DR is the retinal alteration score (RAS), which has been designed and previously validated for the detection of retinal changes, and is described elsewhere.¹⁷ In essence, RAS is a modified version of the CNN Xception. It was initially trained using transfer learning on the EyePACS data set and then fine-tuned using a data set of 10,569 fundus images captured with the Phelcom Eyer device (resolution 1600 × 1600 × 3 red, green and blue channels). Images from the EyePACS data set had similar magnification and field of view as images obtained with the Eyer device. The training data were separated into 2 classes: images from normal eyes and images with retinal alterations.

Another DL-based approach employed in this study is the diabetic retinopathy alteration score (DRAS). It utilized a modified version of the EfficientNetV2S CNN, with different input and output parameters while maintaining the same intermediate convolutional

layers. The input was modified to receive images of size $599 \times 599 \times 3$ red, green and blue channels. Images were resized to reduce data dimensionality and also to enable faster convergence during the training of the network, in the best possible trade-off regarding the final performance. Additionally, the last 3 layers were dropped, and new layers of convolutional, batch normalization, activation, global average pooling, dense, and output layers were added. The network's output represents the probability of alteration for the DR and non-DR classes. Therefore, the class to which the evaluated image belongs is identified based on the neuron with the highest value. DRAS gives an output as a numerical score (Table S1, available at www.ophtalmologyscience.org) ranging from 0 (low probability of DR) to 1 (high probability of DR).

The DRAS network went through a training process in which each image was individually evaluated, allowing its internal parameters to be progressively adjusted to obtain an output from the last layer that closely aligns with the corresponding image class. To train the algorithm, transfer learning was employed using the EyePACS data set along with an internal Phelcom data set comprising 17,330 DR images captured exclusively using the Eyer device (resolution $1600 \times 1600 \times 3$ red, green and blue channels). For validation purposes, 30% of these images were used to periodically evaluate the performance of the network. To add more diversity, data augmentation was applied to images, with rotation, width and height shift, zoom, and brightness values randomly applied. This system's output was assessed for classifying patients regarding the presence of DR.

Finally, images with a positive output from the RAS system were further analyzed by DRAS to provide the mtmDR output. (Fig 1) The combined networks (RAS and DRAS) analysis gives an output as a numerical score (Table S2, available at www.ophtalmologyscience.org) ranging from 0 (low probability of mtmDR) to 1 (high probability of mtmDR).

The implementation of the RAS and DRAS models is based on default Xception and EfficientNet v2s CNNs, respectively (Appendix S1, available at www.ophtalmologyscience.org). The only differences created by our group are that the input and output layers have been modified, so the input consists of resized images (of 599×599) and the output consists of 2 classes (DR and non-DR classes).

The automatic DR detection was based on a single fundus image per eye, macula-centered for both outcomes: any DR and mtmDR. After resizing, images obtained by the device were fed into the DL systems, without any other preprocessing.

Statistical Analysis

Data were collected in MS Excel 2010 files (Microsoft Corporation). Statistical analyses were performed using SPSS 19.0 for Windows (SPSS Inc). Individuals' characteristics and quantitative variables are presented in terms of mean and standard deviation. The chi-square test was used for the comparison of DR severity among races. The intergraders agreement was evaluated using the kappa (κ) statistic test, which ranges from 0 (no agreement or agreement that can be expected from random chance) to 1.00 (perfect agreement).¹⁸ The weighted κ was calculated based on International Classification of Diabetic Retinopathy classification subgroups. The 5% level of significance was used. Sensitivity, specificity, positive predictive value, and negative predictive value and their 95% confidence intervals (CIs) were calculated for the device outputs; for those calculations and also for the calculation of sample size, the populational prevalences were estimated according to available data in the literature.¹⁹ Artificial intelligence output was compared against human reading as the ground truth (see Tables S1 and S2). Diagnostic accuracy is reported according to the Standards for Reporting of Diagnostic Accuracy Studies.²⁰

Results

This study included retinal fundus photographs from 327 patients with a mean age of 57.0 (standard deviation, 16.8; range, 9–90) years and 45.3% male patients. Diabetes duration averaged $16.3 (\pm 9.7)$ years. Patients demographics and comorbidities are summarized in Table 1. Regarding DR classification, 44% of patients had no retinopathy, 26.47% had nonproliferative DR, and 29.31% had proliferative DR.

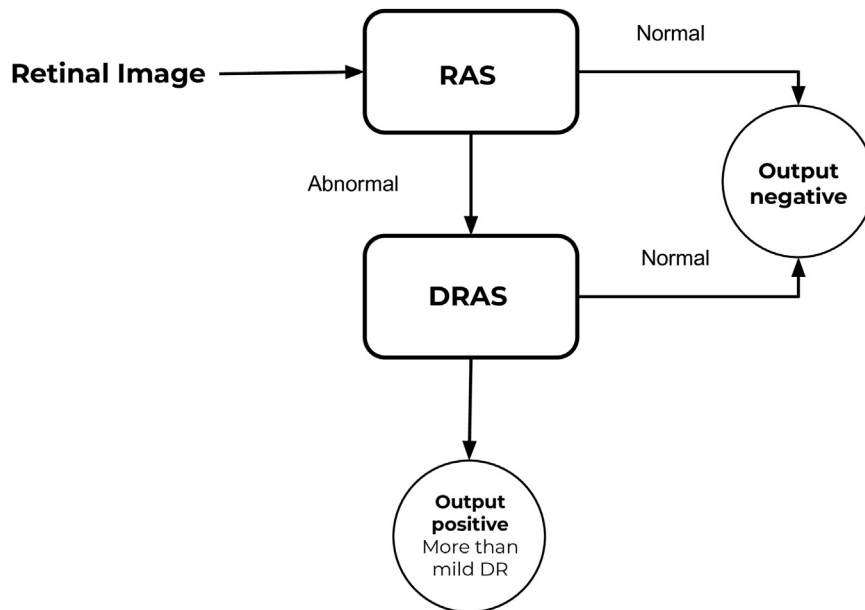


Figure 1. Artificial intelligence algorithms combination for detection of more-than-mild diabetic retinopathy (DR). DRAS = diabetic retinopathy alteration score; RAS = retinal alteration score.

Table 1. Demographic Description and Clinical Characteristics of Enrolled Patients*

	Value $\pm$ SD or (%)
Male sex	148 (45.26%)
Age, yrs	57.03 $\pm$ 16.82
Race [†]	
Mixed	133 (40.67%)
White	107 (32.72%)
Black	70 (21.42%)
Asian	10 (3.06%)
Indigenous	3 (0.92%)
Diabetes diagnosis, yrs	16.35 $\pm$ 9.69
Insulin use, yrs	11.67 $\pm$ 9.47
BMI, kg/m ²	28.11 $\pm$ 5.37
Hypertension	221 (68%)

BMI = body mass index; SD = standard deviation.

*Values are presented as mean $\pm$ SD; percentages are indicated when adequate.

[†]Race distribution was not associated with diabetic retinopathy severity across the sample ($P = 0.54$).

The intergraders agreement regarding different categories of DR severity was very high (weighted κ , 0.895). After initial automatic assessment for quality, 20 patients (6.1%) who had their images graded by the human experts could not have them graded by the DL systems due to insufficient quality (Fig 2), resulting in a total of 307 individuals who had their images with sufficient quality for classification by both the specialist reader and the AI.

The sensitivity and specificity, per the human grading standard, for the system to detect any DR was 90.48% (95% CI, 84.99%–94.46%) and 90.65% (95% CI, 84.54%–94.93%), respectively. Positive predictive value and negative predictive

value for any DR were 73.49% (95% CI, 62.24%–82.34%) and 97.08% (95% CI, 95.41%–98.15%), respectively.

The sensitivity and specificity, per the human grading standard, for the system to detect mtmDR was 90.23% (95% CI, 83.87%–94.69%) and 85.06% (95% CI, 78.88%–90.00%), respectively. Positive predictive value and negative predictive value for mtmDR were 28.42% (95% CI, 21.71%–36.24%) and 99.25% (95% CI, 98.74%–99.55%), respectively.

The confusion matrices for both outcomes are displayed on Table 2. The area under the receiver operating characteristic curve was 0.948 for any DR and 0.895 for mtmDR (Fig 3).

A qualitative post hoc analysis performed on incorrectly classified cases (false-positives and false-negatives, see Table 2) pointed to subtle losses of image quality in the majority of such cases: blurring, nonhomogenous fundus illumination, images with areas of low visibility, excessive brightness or overexposure, subtle losses of focus and/or image sharpness. Moreover, in some false negative cases, we have noticed the presence of microaneurysms or small hemorrhages; in selected false positive cases, we have observed the presence of drusen or pigment clumping in the macular area, tigroid or tessellated patterns of the fundus, or the persistence of myelinated fibers.

Discussion

The present DR screening strategy with an automated detection system for the analysis of a single image per eye, obtained with a low-cost, portable retinal camera, achieved high sensitivity and specificity values for both outcomes: any DR and mtmDR. Of note, the achieved sensitivity values (90.48% vs. 90.23% for any DR and mtmDR, respectively) and specificity values (90.65% vs. 85.06% for

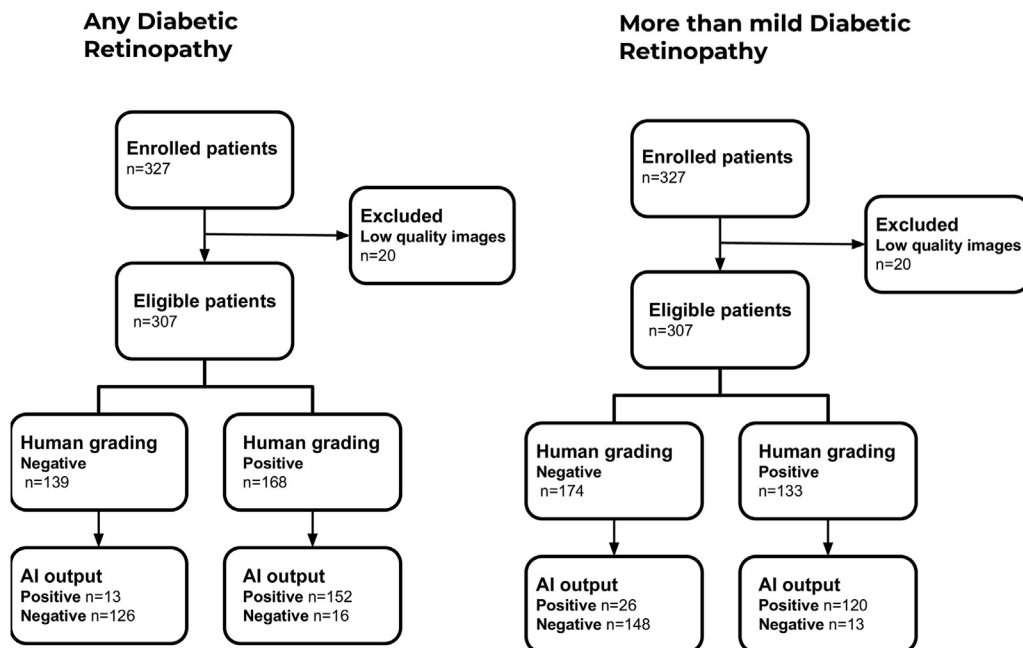


Figure 2. Waterfall diagram. Standards for Reporting of Diagnostic Accuracy Studies diagram for the algorithm. Eligible patients were analyzed by both the automatic systems and the human graders. Left: any diabetic retinopathy output. Right: more-than-mild diabetic retinopathy output. AI = artificial intelligence.

Table 2. Confusion Matrices for the Outcomes: Any DR and mtmDR*

	Expert Reading Positive	Expert Reading Negative	Total
Any DR predicted			
Positive	152 (TP)	13 (FP)	165
Negative	16 (FN)	126 (TN)	142
Total	168	139	307
mtmDR predicted			
Positive	120 (TP)	26 (FP)	146
Negative	13 (FN)	148 (TN)	161
Total	133	174	307

DR = diabetic retinopathy; FN = false negative; FP = false positive; mtmDR = more-than-mild diabetic retinopathy; TN = true negative; TP = true positive.

*Displayed results are valid for patients who underwent human and automatic analysis.

any DR and mtmDR, respectively) were higher than the prespecified primary endpoint goals established for the first FDA approval of an automatic system for the detection of DR (sensitivity > 85% and specificity > 82.5%).¹⁰ The achieved values also compare well with those reported in the literature, both for strategies with traditional, tabletop retinal cameras and with portable devices.^{5,21–25}

Limited access to ophthalmologists in many parts of the world poses a significant challenge in preventing avoidable blindness secondary to diabetes.²⁶ In order to overcome challenges brought by the increasing numbers of diabetes cases worldwide, along with social and economic barriers to screening, telemedicine and AI have been established as cost-effective solutions.^{5,17} The employment of a smartphone-based device with an offline, embedded AI system could allow incremental coverage because of the portability and the potential of point-of-care results, even in

areas without internet connection, further improving adherence rates and aiding in the prevention of visual loss.^{9,27}

The portable device employed in the present study has previously been clinically validated for the detection of DR and tested in a real-world, high-burden setting.^{27,28} The adoption of low-cost, portable devices, such as the one reported herein, allows an easier screening process, especially in low-resource settings and hard-to-reach populations.²⁷ Regarding the evolution of the imaging protocol for DR screening, we believe that the high performance achieved by automated evaluation of a single, macula-centered retinal image per eye, has the potential of making the process more friendly both for the operator and the patient; besides saving time, it is more likely to obtain good quality images in a single-field protocol than if multiple fields are required.

Other authors have recently assessed the performance of AI systems for DR evaluation using portable retinal cameras and yielding variable outcomes, including the detection of any DR, referable DR and sight-threatening DR; the main results of those studies are displayed in Table 3.^{5,23–25} The performance of our systems was comparable and often superior to the ones evaluated in those studies. Besides individual cameras' and AI systems' characteristics, such heterogeneity of performances may also be explained by different study designs; uneven sample sizes and variable data sets composition, especially the proportion of patients with and without the condition of interest, may also have influenced such variable results. Our sample was composed of 327 patients with diabetes, 26.47% with nonproliferative DR, and 29.31% with proliferative DR.

The choice of different algorithm outputs among the mentioned studies also makes their comparison challenging, and this aspect is very relevant in terms of clinical application: the desired outcome has to suit local circumstances, the available health care workforce, financial constraints, and the intended goals of each program. The outcome of detecting mtmDR was chosen when the first autonomous system for DR evaluation was approved by the FDA.¹⁰ In our study, a good performance was achieved not only for mtmDR but also for the detection of any DR. Such “disease” versus “no disease” strategy has already been recognized as suitable as an assistive approach for reducing the burden on expert manual grading of retinal images, with a potentially favorable cost-to-benefit ratio regarding the number of human graders needed.²⁹ Another reason for choosing the “any DR” outcome is related to the suggestion that a lower threshold for referrals is necessary when handheld cameras are employed for DR screening.⁶ Our research group has recently observed that eventual disagreements in the detection of microaneurysms, small hemorrhages, and intraretinal microvascular abnormalities may contribute to some discordance between portable devices and the traditional, tabletop retinal camera.²⁷ Finally, separating patients with diabetes with DR from those without DR may also be a valuable tool if the risk for systemic events is to be evaluated.¹²

Ensuring high sensitivity and specificity rates is crucial for the success of a screening strategy. Although the systems reported herein attained a high performance, we also

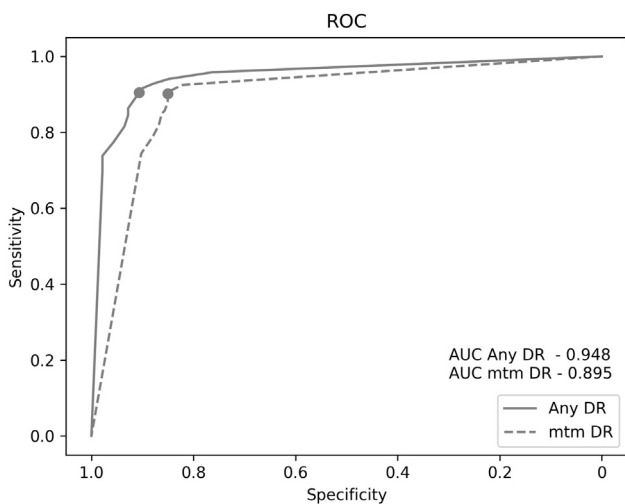


Figure 3. Receiver operating characteristic (ROC) curve of the artificial intelligence systems for detection of any diabetic retinopathy (DR) and more-than-mild DR (mtmDR). Area under the curve (AUC) for any DR: 0.948; AUC for mtmDR: 0.895.

Table 3. Performance of AI Systems for the Detection of DR with Portable Retinal Cameras (Selected Studies)^{5,23-25}

Authors and Year	Population/Setting	Sample Size (Rate of DR)	Output	Sensitivity/ Specificity (%)	Camera, AI System
Lupidi et al, ⁵ 2023	Specialized Retina Clinic (Italy)	256 (50%)	Any DR	96.8/96.8	Optomed Aurora (Optomed, Oulu, Finland); Selena+
Ruan et al, ²³ 2022	Hospital-based (China)	315 (N/A)	Referable DR	88.2/40.7	Optomed Aurora (Optomed, Oulu, Finland); Phoebus
Rajalakshmi et al, ²⁴ 2018	Hospital-based (India)	296 (65%)	Any DR Sight-threatening DR	95.8/80.2 99.1/80.4	Remidio (Remidio Innovative Solutions, India); EyeArt
Nunez do Rio et al, ²⁵ 2022	Community screening, rural and urban areas (India)	11 199 (3.8%*)	Referable DR	72.08/85.65	Zeiss Visuscout (Carl Zeiss Meditec, Jena, Germany); VISUHEALTH-AI DR

AI = artificial intelligence; DR = diabetic retinopathy; N/A = not available.
*Referable DR.

described some incorrectly classified cases. We believe that even subtle losses of image quality, corresponding to small changes in the system's input, may have affected AI performance in a small fraction of the studied data set, occasionally resulting in false-negative or false-positive results. Based on the qualitative analysis of incorrectly classified cases, we believe future advances in automated detection of DR will include the improvement of automated quality assessment, to prevent images with suboptimal quality from being assessed by the algorithms for DR detection.^{7,30} In addition, DL systems for DR detection could be improved with model retraining and data augmentation of the training data set.

Besides a high performance, several other aspects must be implemented in order to guarantee that automatic systems are fully integrated into a clinical workflow and will actually lead to health improvements, and such factors depend on local circumstances such as regulatory issues, the legal framework, disease prevalence, availability of workforce, and economic constraints. It is important to recognize that, even if a DL system generates increased screenings and better referral adherence, access to subsequent specialty care might be unavailable within the current health care infrastructure.³¹ Further studies are needed for the evaluation of clinical outcomes and health economic metrics.^{31,32} Additionally, raising awareness about DR among patients and local health care workers is essential for the overall success of screening initiatives.⁹

We believe the main strength of this study lies in its design; besides being a multicenter study, it relied on a consistent and robust labeling performed by a reading center, and the data set was collected from an adequately balanced sample. Another important strength of the study is reporting the performance of automatic analysis evaluating a single image per eye for 2 different outcomes, representing different severity levels of DR; besides employing an automatic algorithm for the identification of retinal changes previously reported by our group,¹⁷ in the present study, we also evaluated a new algorithm that demonstrated high sensitivity and specificity for the detection of DR. The combination of both systems performed well in detecting mtmDR. Furthermore, we employed a previously validated

device, not only regarding comparison to tabletop retinal cameras but also for being used in a real-world, high-burden setting.^{27,28}

However, our study has limitations. The most important one is related to an analysis undertaken only on images with good quality, obtained by trained photographers, after pupil dilation, which precludes the extrapolation of our results in a real-world setting. Of note, some automatic systems of quality evaluation have been reported in this field.^{7,30} In addition, ours was not a real-world sample, because patients were recruited from reference centers, as opposed to patients attending screening programs. Finally, a good performance is a necessary but not sufficient condition for the success of the deployment of an AI-assisted screening program: results must be generalizable to different health care settings and among diverse demographic factors such as age, race, ethnicity, and socioeconomic status; clinical outcomes and economic analyses are also needed for real-world implementation.³¹ Nevertheless, we believe the present study adds value to the field of DR screening as a proof of concept, with further studies needed for the validation of the strategy with real-world data, such as the study authored by Bhaskaranand et al.³³

In conclusion, with a growing diabetes epidemic and the global challenges faced by health systems to avoid blindness caused by diabetes, it is clear that portable and cost-effective devices and automatic systems will play a role in maximizing the outcomes of screening programs. The reported AI system, which attained a high performance for both outcomes of any DR and mtmDR, with the potential of being embedded in a portable device, working offline, and relying on just a single retinal image per eye, is a promising alternative for health systems overwhelmed by such a burden. Future studies are needed to evaluate the cost-effectiveness of such strategy and the feasibility of its deployment in the real world.

Acknowledgments

The authors acknowledge Juliana Angélica Estevão de Oliveira, Talita Virgínia Fernandes de Oliveira, Stefano Neto Jai Hyun Choi, Edgar Menezes Neto, and Viviane Santos Cardoso for the collection and classification of retinal images.

Footnotes and Disclosures

Originally received: September 11, 2023.

Final revision: January 20, 2024.

Accepted: January 25, 2024.

Available online: February 7, 2024. Manuscript no. XOPS-D-23-00188R2.

¹ Federal University of São Paulo, São Paulo, Brazil.

² Phelcom Technologies, São Carlos, Brazil.

Disclosures:

All authors have completed and submitted the ICMJE disclosures form.

The authors have made the following disclosures:

F.K.M.: Medical consultant — Phelcom Technologies.

G.B.M.: Grant — SJJ Solutions and Molecular Partners; Consultant — SJJ Solutions; Support for attending meetings and/or travel — SJJ Solutions; Participation on a Data Safety Monitoring Board or Advisory Board — Roche, Bayer, and West Pharmaceuticals.

J.A.S.: Chief Executive Officer — Phelcom Technologies; Stock — Phelcom Technologies.

D.L.: Chief Technology Officer — Phelcom Technologies; Stock — Phelcom Technologies.

P.V.P.: Employee — Phelcom Technologies.

S.A.D.: Support — From Novo Nordisk for American Diabetes Association Meeting 2023 in San Francisco.

HUMAN SUBJECTS: Human subjects were included in this study. This was an observational cross-sectional study, in compliance with the tenets of the Declaration of Helsinki. Informed consent was obtained from each

patient before the study enrollment. The study was approved by the Research Ethics Committee of the Federal University of São Paulo (33.842.220.7.0000.5505).

No animal subjects were included in this study.

Author Contributions:

Conception and design: Malerbi, Stuchi, Lencione, Prado

Data collection: Malerbi, Nakayama, Melo, Stuchi, Prado, Ribeiro, Regatieri

Analysis and interpretation: Malerbi, Nakayama, Melo, Stuchi, Lencione, Prado, Ribeiro, Dib, Regatieri

Obtained funding: N/A

Overall responsibility: Malerbi, Nakayama, Melo, Stuchi, Lencione, Prado, Ribeiro, Dib, Regatieri

Abbreviations and Acronyms:

AI = artificial intelligence; **CI** = confidence interval; **CNN** = convolutional neural network; **DL** = deep learning; **DR** = diabetic retinopathy; **DRAS** = Diabetic Retinopathy Alteration Score; **FDA** = Food and Drug Administration; **mtmDR** = more-than-mild diabetic retinopathy; **RAS** = retinal alteration score.

Keywords:

Artificial intelligence, Diabetes, Handheld camera, Portable, Retina.

Correspondence:

Fernando K. Malerbi, Rua Botucatu, 822, CEP 04023-062, São Paulo, Brazil. E-mail: fernandokmalerbi@gmail.com.

References

- Speechley M, Kunnilathu A, Aluckal E, et al. Screening in public health and clinical care: similarities and differences in definitions, types, and aims—a systematic review. *J Clin Diagn Res*. 2017;11:LE01–LE04.
- Grauslund J. Diabetic retinopathy screening in the emerging era of artificial intelligence. *Diabetologia*. 2022;65:1415–1423.
- Padilla Conde T, Robinson L, Vora P, et al. Effectiveness of telemedicine diabetic retinopathy screening in the USA: a protocol for systematic review and meta-analysis. *Syst Rev*. 2023;12:48.
- Jani PD, Forbes L, Choudhury A, et al. Evaluation of diabetic retinal screening and factors for ophthalmology referral in a telemedicine network. *JAMA Ophthalmol*. 2017;135:706–714.
- Lupidi M, Danieli L, Fruttini D, et al. Artificial intelligence in diabetic retinopathy screening: clinical assessment using handheld fundus camera in a real-life setting. *Acta Diabetol*. 2023;60:1083–1088.
- Jacoba CMP, Salongcay RP, Rageh AK, et al. Comparisons of handheld retinal imaging with optical coherence tomography for the identification of macular pathology in patients with diabetes. *Ophthalmic Res*. 2023;66:903–912.
- Huemer J, Wagner SK, Sim DA. The evolution of diabetic retinopathy screening programmes: a chronology of retinal photography from 35 mm slides to artificial intelligence. *Clin Ophthalmol*. 2020;14:2021–2035.
- Chen JS, Lin MC, Yiu G, et al. Barriers to implementation of teleretinal diabetic retinopathy screening programs across the University of California. *Telemed J E Health*. 2023;29:1810–1818.
- Malerbi FK, Melo GB. Feasibility of screening for diabetic retinopathy using artificial intelligence, Brazil. *Bull World Health Organ*. 2022;100:643–647.
- 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;1:39.
- Ipp E, Liljenquist D, Bode B, et al; EyeArt Study Group. Pivotal evaluation of an artificial intelligence system for autonomous detection of referable and vision-threatening diabetic retinopathy. *JAMA Netw Open*. 2021;4:e2134254.
- Pearce I, Simó R, Lövestam-Adrian M, et al. Association between diabetic eye disease and other complications of diabetes: implications for care. A systematic review. *Diabetes Obes Metab*. 2019;21:467–478.
- Malerbi FK, Morales PH, Farah ME, et al. Comparison between binocular indirect ophthalmoscopy and digital retinography for diabetic retinopathy screening: the multicenter Brazilian Type 1 Diabetes Study. *Diabetol Metab Syndr*. 2015;7:116.
- Wilkinson CP, Ferris III FL, Klein RE, et al. Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. *Ophthalmology*. 2003;110:1677–1682.
- Shi C, Lee J, Wang G, et al. Assessment of image quality on color fundus retinal images using the automatic retinal image analysis. *Sci Rep*. 2022;12:10455.
- Stuchi JA, Canto NG, Boccato L, Attux R. A frequency-domain approach with learnable filters for image classification.

- <https://ssrn.com/abstract=4333420>; 2023. Accessed December 21, 2023.
17. Malerbi FK, Andrade RE, Morales PH, et al. Diabetic retinopathy screening using artificial intelligence and handheld smartphone-based retinal camera. *J Diabetes Sci Technol*. 2022;16:716–723.
18. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)*. 2012;22:276–282.
19. Teo ZL, Tham YC, Yu M, et al. Global prevalence of diabetic retinopathy and projection of burden through 2045: systematic review and meta-analysis. *Ophthalmology*. 2021;128:1580–1591.
20. Šimundić AM. Measures of diagnostic accuracy: basic definitions. *EJIFCC*. 2009;19:203–211.
21. Wu JH, Liu TYA, Hsu WT, et al. Performance and limitation of machine learning algorithms for diabetic retinopathy screening: meta-analysis. *J Med Internet Res*. 2021;23:e23863.
22. Tufail A, Rudisill C, Egan C, et al. Automated diabetic retinopathy image assessment software: diagnostic accuracy and cost-effectiveness compared with human graders. *Ophthalmology*. 2017;124:343–351.
23. Ruan S, Liu Y, Hu WT, et al. A new handheld fundus camera combined with visual artificial intelligence facilitates diabetic retinopathy screening. *Int J Ophthalmol*. 2022;15:620–627.
24. Rajalakshmi R, Subashini R, Anjana RM, Mohan V. Automated diabetic retinopathy detection in smartphone-based fundus photography using artificial intelligence. *Eye (Lond)*. 2018;32:1138–1144.
25. Nunez do Rio JM, Nderitu P, Bergeles C, et al. Evaluating a deep learning diabetic retinopathy grading system developed on mydriatic retinal images when applied to non-mydriatic community screening. *J Clin Med*. 2022;11:614.
26. Sengupta S, Sindal MD, Besirli CG, et al. Screening for vision-threatening diabetic retinopathy in South India: comparing portable non-mydriatic and standard fundus cameras and clinical exam. *Eye (Lond)*. 2018;32:375–383.
27. de Oliveira JAE, Nakayama LF, Zago Ribeiro L, et al. Clinical validation of a smartphone-based retinal camera for diabetic retinopathy screening. *Acta Diabetol*. 2023;60:1075–1081.
28. Malerbi FK, Andrade RE. Real-World diabetic retinopathy screening with a handheld fundus camera in a high-burden setting. *Acta Ophthalmol*. 2022;100:e1771.
29. Ribeiro L, Oliveira CM, Neves C, et al. Screening for diabetic retinopathy in the central region of Portugal. Added value of automated ‘disease/no disease’ grading. *Ophthalmologica*. 2014;233:96–103.
30. Ruamviboonsuk P, Tiwari R, Sayres R, 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–e244.
31. Yuan A, Lee AY. Artificial intelligence deployment in diabetic retinopathy: the last step of the translation continuum. *Lancet Digit Health*. 2022;4:e208–e209.
32. Xie Y, Nguyen QD, Hamzah H, 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–e249.
33. Bhaskaranand M, Ramachandra C, Bhat S, et al. The value of automated diabetic retinopathy screening with the EyeArt System: a study of more than 100,000 consecutive encounters from people with diabetes. *Diabetes Technol Ther*. 2019;21:635–643.