

Diabetic Retinopathy Screening Using Artificial Intelligence and Handheld Smartphone-Based Retinal Camera

Journal of Diabetes Science and Technology
2022, Vol. 16(3) 716–723
© 2021 Diabetes Technology Society
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1932296820985567
journals.sagepub.com/home/dst


Fernando Korn Malerbi, MD, PhD^{1,2}, Rafael Ernane Andrade, MD, PhD^{1,3},
Paulo Henrique Morales, MD, PhD^{1,2}, José Augusto Stuchi, MSC⁴,
Diego Lencione, MSC⁴, Jean Vitor de Paulo, MSC, PhD⁴,
Mayana Pereira Carvalho, MD⁵, Fabrícia Silva Nunes, MD⁵,
Roseanne Montargil Rocha, RN, PhD⁵, Daniel A. Ferraz, MD, PhD^{2,6},
and Rubens Belfort Jr, MD, PhD^{1,2}

Abstract

Background: Portable retinal cameras and deep learning (DL) algorithms are novel tools adopted by diabetic retinopathy (DR) screening programs. Our objective is to evaluate the diagnostic accuracy of a DL algorithm and the performance of portable handheld retinal cameras in the detection of DR in a large and heterogenous type 2 diabetes population in a real-world, high burden setting.

Method: Participants underwent fundus photographs of both eyes with a portable retinal camera (Phelcom Eyer). Classification of DR was performed by human reading and a DL algorithm (PhelcomNet), consisting of a convolutional neural network trained on a dataset of fundus images captured exclusively with the portable device; both methods were compared. We calculated the area under the curve (AUC), sensitivity, and specificity for more than mild DR.

Results: A total of 824 individuals with type 2 diabetes were enrolled at Itabuna Diabetes Campaign, a subset of 679 (82.4%) of whom could be fully assessed. The algorithm sensitivity/specificity was 97.8 % (95% CI 96.7-98.9)/61.4 % (95% CI 57.7-65.1); AUC was 0.89. All false negative cases were classified as moderate non-proliferative diabetic retinopathy (NPDR) by human grading.

Conclusions: The DL algorithm reached a good diagnostic accuracy for more than mild DR in a real-world, high burden setting. The performance of the handheld portable retinal camera was adequate, with over 80% of individuals presenting with images of sufficient quality. Portable devices and artificial intelligence tools may increase coverage of DR screening programs.

Keywords

artificial intelligence, Covid-19, diabetic retinopathy, mobile healthcare, point-of-care, screening, telemedicine

Introduction

Diabetic retinopathy (DR) screening is considered one of the most cost-effective initiatives on diabetes care,¹ fitting into the World Health Organization criteria for the screening of chronic disease; its importance in preventing blindness is well established.² However, implementation of broad public health DR screening programs remains a substantial challenge where there are insufficient health care structure and resources in terms of funding, trained health care human resources, and facilities.³

Novel portable smartphone-based retinal cameras are readily available, low-cost, easy-to-use devices with validated sensitivity and specificity for DR screening; their implementation is feasible and realistic.⁴

¹Department of Ophthalmology and Visual Sciences, Federal University of São Paulo, São Paulo, Brazil

²Instituto Paulista de Estudos e Pesquisas em Oftalmologia, IPEPO, Vision Institute, São Paulo, Brazil

³Hospital de Olhos Beira Rio, Itabuna, BA, Brazil

⁴Phelcom Technologies, São Carlos, Brazil

⁵State University of Santa Cruz, Ilhéus, BA, Brazil

⁶NIHR Biomedical Research Centre for Ophthalmology, Moorfields Eye Hospital, NHS Foundation Trust and UCL Institute of Ophthalmology, London, UK

Corresponding Author:

Fernando Korn Malerbi, Federal University of São Paulo, Rua Botucatu, 822. São Paulo, SP 04039-032, Brazil.
Email: fernandokmalerbi@gmail.com

Artificial intelligence (AI) systems for DR screening/grading have the potential to increase further the accessibility of DR screening for people with diabetes and increase diagnostic accuracy, efficiency, productivity, reproducibility, and outcomes;⁵ a robust economic rationale has been found for the use of deep learning (DL) systems as assistive tools for DR screening.⁶ Such characteristics confer to AI systems a great relevance especially in low-income, underserved areas with insufficient access to ocular exams, as is the case of Itabuna, Northeast Brazil, where patients rely on the Itabuna Diabetes Campaign, a once-a-year initiative that offers DR screening, treatment, and counseling. Involving mass gatherings of patients and retina specialists, for most patients it was the sole opportunity for the annual eye exam. During the 2019 event, a strategy based on handheld portable cameras with an embedded AI algorithm compatible with telemedicine was evaluated; if successful, such a strategy could allow for a more dispersed, continuous screening intervention that would not require mass gatherings.

The study aims to evaluate the diagnostic accuracy of a semiautomated strategy of DR screening with mobile handheld retinal cameras and an AI algorithm, using a manual grading as the independent reference standard, in a real-life, high burden setting in an area with scarce healthcare resources in Northeast Brazil.

Methods

Study Design, Population, and Setting

This retrospective study enrolled a convenience sample of individuals aged over 18 years old with a previous type 2 diabetes mellitus (T2DM) diagnosis who attended the Itabuna Diabetes Campaign, an event that took place on November 23, 2019 at the city of Itabuna, Bahia State, Northeast Brazil (Latitude 14° 47'08" longitude 39° 16'49"W). This annual event mobilizes a significant amount of the city's inhabitants and involves diabetes awareness, counseling, screening, and treatment of diabetes complications. The study protocol was approved by the ethics Committee of Federal University of São Paulo (# 1260/2015) and was conducted in compliance with the Declaration of Helsinki, following the institutional ethics committees. After signing informed consent, participants answered a questionnaire for demographic and self-reported clinical characteristics and underwent ocular imaging.

Imaging

Smartphone-based hand-held devices (Eyer, Phelcom Technologies, São Carlos, Brazil) were used for fundus photography and image acquisition: two images of the posterior segment—one centered on macula and another disc centered (45° field of view)—were captured for each eye, after mydriasis induced by 1% tropicamide eye-drops.⁷ Image acquisition was performed by a team of nine examiners, including

medical students, with variable degrees of previous experience in this kind of procedure.

Image Grading

Remote image reading was performed at EyerCloud platform (Phelcom Technologies, São Carlos, Brazil) by a single retinal specialist (FKM) after anonymization. The photographs were evaluated for quality, regarding transparency of the media, focus, and image boundaries, and classified as gradable or ungradable images.⁷ The lack of images with representative fields was considered as protocol failure. Subsequently, DR and maculopathy classification was manually determined from fundus photographs for those individuals with gradable images. Classification of DR was given per individual, considering the most affected eye, according to a strict standardized protocol (Table 1)⁸; apparently present macular oedema was defined as apparent retinal thickening or hard exudates in posterior pole.⁹ For analysis, classification grades were combined into no or mild non-proliferative DR (NPDR) versus moderate NPDR, severe NPDR, proliferative DR, or apparently present macular oedema (more than mild DR, mtmDR) in at least one eye. Vision-threatening DR (vtDR) was considered as the presence of severe NPDR, proliferative DR, or apparently present macular oedema in at least one eye.⁹ Individuals with images of insufficient quality were excluded from analysis, except cases when one gradable eye was classified as vtDR. No information other than ocular images was available for the reader.

Automated Detection of DR

Images from 824 individuals, totalizing 3,255 color images, were graded by a DL-enhanced method (PhelcomNet), a modified version of the convolutional neural network (CNN) Xception,¹⁰ with different input and output parameters but the same intermediate convolutional layers. The input was changed to receive images of size $699 \times 699 \times 3$ RGB, and two fully connected layers of 2100 neurons were added at the top. Finally, two neurons with softmax activation classified the network input according to class. Softmax normalized the respective neuron input values, creating a probabilistic distribution in which the sum will be 1. Therefore, the neuron with the highest value identified the class to which the evaluated image belongs.

The PhelcomNet prediction score was defined as x , corresponding to the interval between 0 and 1, which indicates the likelihood of DR. In contrast to the reference standard, the device puts both no DR and mild DR into one grade; all the other levels are considered more than mild DR. Since there are up to four images per individual, each of the images has a score indicating a probability of retinal alteration. In order to obtain a single score for each individual considering all outputs, we used a linear combination y of each prediction for image x_1, x_2, x_3, x_4 ordered in descending order— x_1 being the highest and x_4 being the lowest score. In this case, four

Table 1. Diabetic Retinopathy Severity Levels and Distribution Among Patients.

Diabetic retinopathy severity level	Distribution (%)	Lesions
Absent	62.7	No alterations
Mild NPDR	11.0	At least one hemorrhage or microaneurysm
Moderate NPDR	10.0	Four or more hemorrhages in only one hemi-field ^a
Severe NPDR	5.1	Any of the following: - Four or more hemorrhages in the superior and inferior hemi-fields - Venous beading - Intraretinal microvascular abnormalities (IRMA)
Proliferative diabetic retinopathy	10.6	Any of the following: - Active neovessels - Vitreous hemorrhage

Abbreviation: NPDR, non-proliferative diabetic retinopathy.

^aSuperior and inferior hemi-fields separated by the line passing through the center of the macula and the optic disc.

linear combinations were employed, depending on the number of images per individual and defined as:

$$y_4 = 0.5x_1 + 0.25x_2 + 0.125x_3 + 0.125x_4$$

$$y_3 = 0.5x_1 + 0.25x_2 + 0.25x_3$$

$$y_2 = 0.5x_1 + 0.5x_2$$

$$y_1 = 1.0x_1$$

The inspiration for these linear combinations is that a larger weight for predictions with a higher probability makes it more significant than the others, since fundus images with higher DR evidence tend to be more important on human assessment.

The PhelcomNet went through a training process that evaluated each image individually, progressively adjusting its internal parameters values to obtain an output from the softmax layer closest to the class to which the image belongs. For the algorithm training, a dataset of 10,569 fundus images captured between 2019 and 2020, exclusively using the Phelcom Eyer device (resolution 1600 × 1600 × 3 RGB) was used. For validation, 20% of these images were used to periodically evaluate the performance of the network. Images were separated into two classes: those containing only images from normal eyes (Class 0), and those containing images with retinal alterations (Class 1).

To add more diversity to the data, transformations were applied to images before CNN evaluation, a technique known as data augmentation. Rotation, width and height shift, zoom, and brightness values were randomly applied.

GradCam Heatmap

In order to visualize the location of the most important regions obtained by CNN, as a means to discriminate between classes,

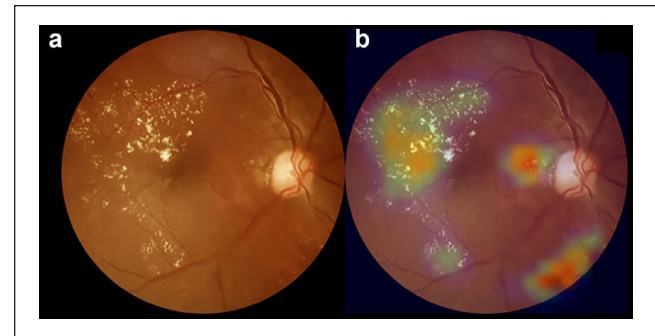


Figure 1. Example of heatmap visualization using gradient method. (a) Color fundus photograph depicting hard exudates and microaneurysms in the macular region, suggesting the possibility of diabetic macular oedema. (b) Overlay with the GradCam heatmap visualization can aid in making a diagnosis as the modifications are flagged in a color scale, from blue (low importance) to red (high importance).

the Gradient-Based Class Activation Map (GradCam) was used¹¹; it generates a heat map with the values obtained on the last convolutional layer (Figure 1). The training set was split between Classes 0 and 1, rendering the lesions as the most important discriminatory region for classes to be differentiated from each other; this is notably highlighted on the heat map. The equation to obtain the weights of importance of a neuron is described elsewhere.¹¹

Statistical Analysis

Data were collected in MS Excel 2010 files (Microsoft Corporation, Redmond, WA, USA). Statistical analyses were performed using SPSS 19.0 for Windows (SPSS Inc., Chicago, IL, USA). Individual's characteristics and quantitative variables are presented in terms of mean and standard deviation (SD). A paired two-tailed Student *t* test was used to compare continuous clinical variables between the two groups. Fisher's exact or chi-square tests were used for unpaired variables. The 5% level of significance was used. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV), and their 95% confidence intervals (CIs), were calculated for the device outputs with no or mild DR and mtmDR, compared with the corresponding reference standard classifications⁹; comparison was made against human reading as the ground-truth. The 0.75 threshold was chosen as the operating point (see Supplementary Material). Diagnostic accuracy is reported according to the Standards for Reporting of Diagnostic Accuracy Studies (STARD).¹²

Results

Nine hundred and forty individuals aged over 18 years old with a previous T2DM diagnosis were assessed. Average age

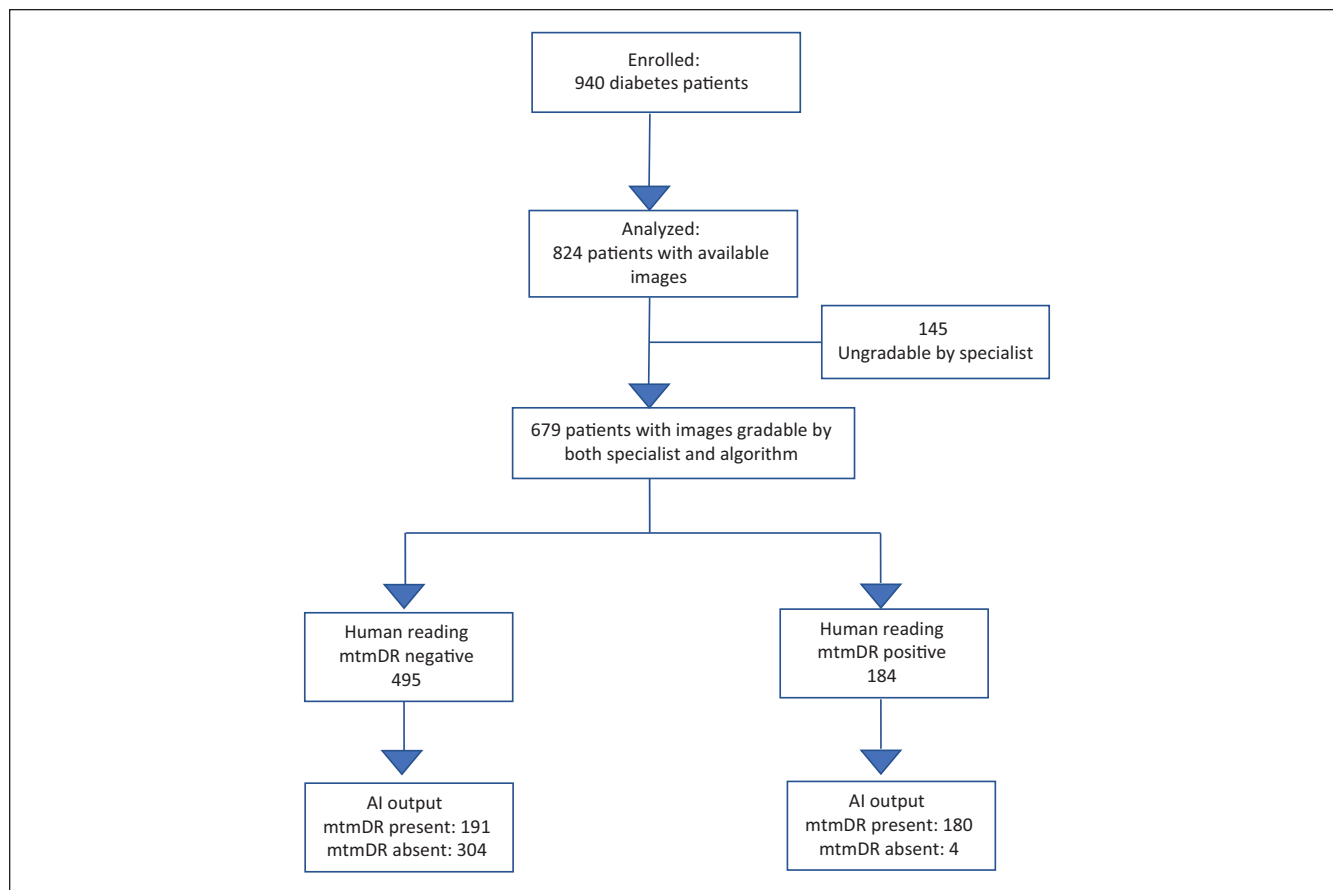


Figure 2. Waterfall diagram. Standards for Reporting of Diagnostic Accuracy Studies (STARD) diagram for the algorithm mtmDR output. mtmDR, more than mild diabetic retinopathy.

was 60.8 ± 11.4 years, and diabetes duration was 10.4 ± 8.7 years; the majority of participants were women (64.9%). Use of insulin was reported in 25.8% of individuals, systemic blood hypertension was present in 68.4%, and smoking habit was present in 48.4%. The vast majority of participants relied on the public health system (94.1%); most individuals were illiterate or had incomplete primary education (54.4%).

Digital fundus photography images were obtained for both eyes of 824 individuals, 145 of whom (17.6%) could not be graded by the specialist reader due to insufficient quality; age ($P < .001$) and diabetes duration ($P = .001$) were inversely associated with gradeability. The remaining 679 individuals had DR classified as follows: absent 426 (62.7%), mild NPDR 75 (11.0%), moderate NPDR 71 (10.0%), severe NPDR 35 (5.1%), and proliferative DR 72 (10.6%) (Table 1). Diabetic macular oedema was detected in 25.4% of individuals. Among individuals with ungradable images as per the specialist reader, the algorithm gave an output of insufficient quality for one participant. Hence, a total of 679 individuals had their images with sufficient quality for classification by both the specialist reader and the AI system (Figure 2).

The sensitivity/specificity, per the human grading standard, for the device to detect mtmDR was 97.8% (95% CI 96.7-98.9)/61.4% (95% CI 57.7-65.1). PPV and NPV for mtmDR were 48.5% (95% CI 44.75-52.27) and 98.7% (95% CI 97.85-99.55), respectively (Table 2). Diagnostic accuracy was 71.3% (95% CI 67.91-74.69). Area under the receiver operating characteristic (ROC) curve was 0.89 (Figure 3).

All cases of false negative or positive exams were individually reviewed. There were four false negative exams for the algorithm output according to the human grading reference standard. All of them were classified as moderate NPDR in the most affected eye, with two individuals having apparently present macular oedema.

Discussion

The results show that a portable smartphone-based retinal camera with an AI algorithm achieved a good diagnostic accuracy for detecting mtmDR, compared to an independent reference standard, in a high burden setting. A review of false negatives did not reveal any case of severe NPDR or proliferative DR. We believe that high sensitivity and the setting of

Table 2. Confusion Matrix for Reference Standard According to Human Grading and Device Output; Sensitivity, Specificity, and Predictive Values of the Algorithm.

Human reference standard	Device output			Sensitivity, specificity
	mtmDR	Not mtmDR	Total	
mtmDR	180	4	184	97.8% sensitivity
Not mtmDR	191	304	495	61.4% specificity
All	371	308	679	
Predictive value	48.5% PPV	98.7% NPV		

Abbreviations: mtmDR, more than mild diabetic retinopathy; NPV, negative predictive value; PPV, positive predictive value.

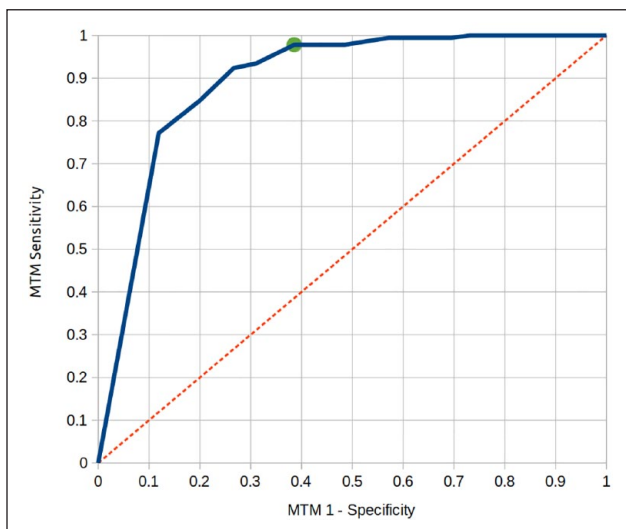


Figure 3. Receiver operating characteristic (ROC) curve of the artificial intelligence device for detection of more than mild diabetic retinopathy (mtmDR). Area under the curve = 0.89.

mtmDR as the cutoff point allows safe use in a screening strategy (Figure 3); the sensitivity reached by the current algorithm compares well with previous reports.^{4,13,14}

Implementation and maintenance of DR screening programs worldwide are challenged by financial and workforce constraints;⁶ under-resourced health systems, particularly in low- to middle-income countries that face severe worker and infrastructure shortages,^{15,16} need rational and cost-effective strategies to deal with the increasing demand brought by growing global prevalence of diabetes mellitus and to overcome social and economic barriers; telemedicine and AI have been established as successful and cost-effective strategies for DR screening,^{6,17} helping increase program coverage and assisting in the detection of vtDR.¹⁵

The high sensitivity found in the present study also allows for a semiautomated strategy, with algorithmic assistance increasing efficiency of non-specialist clinicians to diagnose at scale and reducing the workload of the specialist in a high burden setting, as it directs attention to concerning features consistently across large data sets, for example, through heatmaps (Figure 1).¹⁸ Interestingly, the current algorithm

comes embedded in the handheld device and works offline, which allows for a potential point-of-care DR detection.

The algorithm specificity was somewhat lower than previously reported,^{4,13,14} and a review of false positives revealed mainly fundus pigment changes or image artifacts flagged as pathological changes by the algorithm (Figure 4); such an outcome brings insights between the differences of a deep neuronal network and biological intelligence, the former resulting from a feed-forward approach and the latter being characterized by context-sensitive checking.¹⁹

Adequate image quality is a major factor upon which the success of the screening strategy is dependent,²⁰ and the quality of images obtained with portable devices relies heavily on the training of the operator,²¹ among other variables. Handheld portable retinal cameras are becoming increasingly popular because of its lower cost in comparison with traditional tabletop retinal cameras, and have been shown to help increase access to DR screening.²²⁻²⁴ In the present study, the team responsible for image acquisition was heterogeneous and composed of fully trained members as well as volunteers with no previous experience; there was a high flow of individuals due to the dynamics of the Itabuna Diabetes Campaign, a high-burden setting with more than 900 individuals in a period of six hours. Nevertheless, in the present study, the overall ungradability rate of 17.6% for the specialist reading compares favorably with a recent experience in rural India of 34%.²⁵ Variables related to the individual profile also have influence on image quality, namely age, diabetes duration, poor collaboration, poor mydriasis and media opacities, especially cataracts;²⁶ gradeability was inversely associated with age or diabetes duration.

Individuals enrolled in the present study adequately represent the population with T2DM treated in primary health care, characterized by the predominance of the elderly, mostly women and diabetes duration over 10 years. Arterial hypertension (68.4%) was the most prevalent comorbidity associated with T2DM, but smoking habit (48.4%) was reported at a higher rate than in previous reports.²⁷ The relatively high rate of apparently present diabetic macular oedema may reflect the low level of healthcare access in the region: Brazil contains the fifth largest population of patients with diabetes worldwide,²⁸ and data from a Brazilian multicenter study

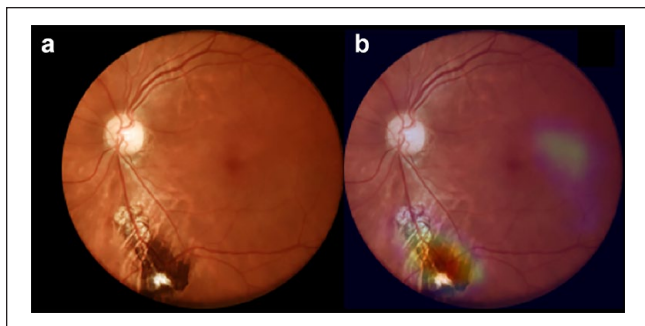


Figure 4. Example of a false positive case. (a) Color fundus photograph depicting inferior pigmentary change. (b) Overlay with the GradCam heatmap; the pigmentary alteration is flagged by the algorithm, leading to a false positive output.

indicate that less than half of T2DM patients followed at reference centers reached the glycemic goal.²⁷ In its Northeast region, access is low to both public healthcare and education,^{27,29} and the region presented the higher mortality rate due to diabetes in the country, possible due to lack of access to healthcare, according to a study on the burden of T2DM in Brazil.³⁰ Data on other diabetes complications are scarce in the region, but the reported prevalence of diabetic foot disease, foot ulcers, and amputation in a previous edition of the Itabuna Diabetes Campaign was 20.6%, 5.8%, and 1.0%, respectively;³¹ a study performed in a Northeastern Brazilian capital revealed that the most frequent cause of end-stage renal disease was diabetes.³² Most individuals in our sample relied exclusively on the public health system (94.1%), were illiterate, or had incomplete primary education (54.4%). Due to the lack of a standardized DR screening protocol in the Brazilian public health system, as an effort to raise access to DR screening and mitigate diabetic blindness, some initiatives based on volunteer work have emerged, and the Itabuna Diabetes Campaign is one successful example; besides diabetes complications treatment and prevention, it helps raise diabetes awareness. Its model, which relies on the retina status as a biomarker of other diabetes complications, was replicated in all five regions of Brazil: during November 2019, Diabetes Campaigns occurred in 24 cities distributed throughout 15 Brazilian states, in commemoration of the month of diabetes awareness and prevention, as established by the International Diabetes Federation³³; however, with the COVID-19 pandemic, an alternative to massive gatherings is needed.³⁴ The current strategy presented herein offers an alternative suitable for this new circumstance, with telemedicine allowing a more dispersed, continuous screening action that does not rely on the presence of specialists, which are scarce in underserved areas; additionally, portable cameras may be operated outside the clinic, even outdoors, rendering this strategy adequate during times of social distancing. A model based on semiautomated screening with mobile units, portable devices, and AI in a primary care setting could contribute to address this novel challenge.

This study has important strengths, as it brings the results of a real-life situation of high burden DR screening in an underserved region of Brazil and a system that involved portable retinal cameras and algorithmic analysis; to the best of our knowledge, this is the first report of such a strategy that could increase access of DR screening in a relevant and cost-effective manner in the country that hosts the fifth most significant population of patients with diabetes on the planet. Of note, such a strategy has already been successfully used in 2020, allowing the first major DR screening initiative after the pandemic.

Our study has several limitations, the most notable of which is that human grading was performed by only one specialist, a potential source of bias. Nevertheless, we believe that the goal of evaluating the potential of AI as an assistive tool for DR screening in a high-burden setting was accomplished; our conclusions point to the relevant role of technology in increasing access to quality healthcare. Additionally, cataracts were not documented or classified systematically, raising concerns on the extent of their influence on gradeability of retinal images. Furthermore, diabetic maculopathy was not evaluated with gold standard methods; instead, its presence was inferred in non-stereoscopic images. The relatively high rate of apparently present diabetic macular oedema found in our sample may be related to the methodology; it may also have been the result of overcall. Finally, the lack of comprehensive clinical and laboratory data is also a limitation of the current study.

Conclusion

In conclusion, the present study shows the feasibility and presents favorable results of a strategy based on retinal images acquired with low-cost, portable devices and automated algorithmic analysis for DR screening in a high burden setting. Such features, combined with telemedicine, may constitute a cost-effective model for middle- to low-income countries where there is insufficient access to ocular healthcare, in a sustainable, continuous manner, as opposed to episodic screening events. The high sensitivity achieved with the algorithm offers the possibility of point-of-care triage. Other challenges remain, such as establishing and complying with the legal and regulatory framework related to algorithm analysis in each healthcare environment. Additionally, DR screening should be envisioned as a part of promoting diabetes awareness, education, and optimizing clinical control; finally, as postulated decades ago, screening should only be offered as long as the health system can provide proper treatment to all detected cases.

Abbreviations

AI, Artificial intelligence; AUC, area under the curve; CI, confidence interval; CNN, convolutional neural network; DL, deep learning; DR, diabetic retinopathy; FAPESP, São Paulo Research Foundation; GradCam, Gradient-Based Class Activation Map;

mtmDR, more than mild DR; NPDR, non proliferative diabetic retinopathy; NPV, negative predictive value; PPV, positive predictive value; ROC, receiver operating characteristic; SD, standard deviation; STARD, Standards for Reporting of Diagnostic Accuracy Studies; T2DM, type 2 diabetes mellitus; vtDR, Vision-threatening DR.

Author Contributions

FKM and REA made the literature search, conceived and designed the study, oversaw the study implementation and collection of data, interpreted the data, contributed to the writing of the first draft of the manuscript and its subsequent revisions and figures, and contributed to its intellectual content. Both of them are the guarantor of this work and, as such, had full access to all underlying data in the study, and take responsibility for the integrity of data and the accuracy of the data analyses; PHM, JAS, DL, JVP, and RMR collected the data, oversaw the implementation of the study protocol, contributed to the writing of the first draft of the manuscript and figures, and contributed to its intellectual content; MPC and FSN collected and analyzed the data; and DAF and RBJ oversaw the study writing, interpreted the data and contributed to the intellectual content of the study and to the final writing.

Authors' Note

The research presented in this paper in the field of artificial intelligence (AI) was conducted at the Research and Development Department of Phelcom Technologies and was supported by Phelcom and FAPESP. Professor Rubens Belfort Jr is a CNPq - National Council for Scientific and Technological Development Researcher IA.

Ethics declarations

The study protocol was approved by the ethics Committee of Federal University of São Paulo (# 1260/2015) and was conducted in compliance with the Declaration of Helsinki, following the institutional ethics committees. Written informed consent was obtained from all participants.

Availability of data and materials

The data that support the findings of this research study are available on request.

Acknowledgements

The authors gratefully acknowledge all the volunteers of Itabuna Diabetes Campaign for their valuable contribution.

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: JAS is Chief Executive Officer and proprietary of Phelcom Technologies. DL is Chief Technology Officer and proprietary of Phelcom Technologies. JVP has a research scholarship conducted at Phelcom Technologies

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This

work was partially funded by the São Paulo Research Foundation (FAPESP) – project numbers 2018/23331-6 and 2017/16014-1.

ORCID iDs

Fernando Korn Malerbi  <https://orcid.org/0000-0002-6523-5172>
Jean Vitor de Paulo  <https://orcid.org/0000-0001-9155-9800>

Supplemental Material

Supplemental material for this article is available online.

References

- 1 Klonoff DC, Schwartz DM. An economic analysis of interventions for diabetes. *Diabetes Care*. 2000;23:390-404.
- 2 Scanlon PH. The English national screening programme for diabetic retinopathy 2003–2016. *Acta Diabetol*. 2017;54:515-525.
- 3 Wong TY, Sun J, Kawasaki R, et al. Guidelines on diabetic eye care. The international council of ophthalmology recommendations for screening, follow-up, referral, and treatment based on resource settings. *Ophthalmology*. 2018;125:1608-1622.
- 4 Vujosevic S, Aldington SJ, Silva P, et al. Screening for diabetic retinopathy: new perspectives and challenges. *Lancet Diabetes Endocrinol*. 2020;8:337-347.
- 5 Shah A, Clarida W, Amelon R, et al. Validation of automated screening for referable diabetic retinopathy with an autonomous diagnostic artificial intelligence system in a Spanish population [published online ahead of print March 16, 2020]. *J Diabetes Sci Technol*. doi:10.1177/1932296820906212
- 6 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*. 2020;2:E240-E249.
- 7 Malerbi FK, Morales PH, Farah ME, et al.; Brazilian Type 1 Diabetes Study Group. 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.
- 8 Philip S, Fleming AD, Goatman KA, et al. The efficacy of automated “disease/no disease” grading for diabetic retinopathy in a systematic screening programme. *Br J Ophthalmol*. 2007;91:1512-1517.
- 9 Wilkinson CP, Ferris FL III, Klein RE, et al.; Representing the Global Diabetic Retinopathy Project Group. Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. *Ophthalmology*. 2003;110:1677-1682.
- 10 Chollet F. Xception: deep learning with depthwise separable convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017:1251-1258. https://openaccess.thecvf.com/content_cvpr_2017/papers/Chollet_Xception_Deep_Learning_CVPR_2017_paper.pdf. Accessed September 7, 2020.
- 11 Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-cam: visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE international conference on computer vision; 2017:618-626. https://openaccess.thecvf.com/content_iccv_2017/html/Selvaraju_Grad-CAM_Visual_Explanations_ICCV_2017_paper.html. Accessed September 7, 2020.

- 12 Šimundić AM. Measures of diagnostic accuracy: basic definitions. *EJIFCC* 2009;19:203-211.
- 13 Abramoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ Digital Med* 2018;1:39.
- 14 Verbraak FD, Abramoff MD, Bausch GCF, et al. Diagnostic accuracy of a device for the automated detection of diabetic retinopathy in a primary care setting. *Diabetes Care*. 2019;42:651-656.
- 15 Bellemo V, Lim ZW, Lim G, et al. Artificial intelligence using deep learning to screen for referable and vision-threatening diabetic retinopathy in Africa: a clinical validation study. *Lancet Digital Health*. 2019;1: e35-e44.
- 16 Pasternak J. What is the future of the Brazilian Public Health System? *Einstein (São Paulo)* 2018;16:eED4811.
- 17 Ben AJ, Neyeloff JL, de Souza CF, 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.
- 18 Sayres R, Taly A, Rahimy E, et al. Using a deep learning algorithm and integrated gradients explanation to assist grading for diabetic retinopathy. *Ophthalmology*. 2019;126:552-564.
- 19 Segal M. A more human approach to artificial intelligence. *Nature*. 2019;571:S18.
- 20 Li HK, Horton M, Bursell SE, et al. Telehealth practice recommendations for diabetic retinopathy. *Telemed J E Health* 2011;17: 814-837.
- 21 Davila JR, Sengupta SS, Niziol LM, et al. Predictors of photographic quality with a handheld nonmydriatic fundus camera used for screening of vision-threatening diabetic retinopathy. *Ophthalmologica*. 2017;238:89-99.
- 22 Hilgert GR, Trevizan E, de Souza JM. Use of a handheld fundus camera as a screening tool for diabetic retinopathy. *Rev Bras Oftalmol*. 2019;78: 321-326
- 23 Kim TN, Myers F, Reber C, et al. A smartphone-based tool for rapid, portable, and automated wide-field retinal imaging. *Transl Vis Sci Technol*. 2018;7:21.
- 24 Malerbi FK, Fabbro ALD, Vieira-Filho JP, Franco LJ. The feasibility of smartphone based retinal photography for diabetic retinopathy screening among Brazilian Xavante Indians. *Diabetes Res Clin Pract*. 2020;168.
- 25 Rani PK, Bhattarai Y, Sheeladevi S, ShivaVaishnavi K, Ali MH, Babu JG. Analysis of yield of retinal imaging in a rural diabetes eye care model. *Indian J Ophthalmol*. 2018; 66:233-237.
- 26 Sosale AR. Screening for diabetic retinopathy—is the use of artificial intelligence and cost-effective fundus imaging the answer? *Int J Diabetes Dev C*. 2019;39:1-3
- 27 Coutinho WF, Silva WS Jr. Diabetes care in Brazil. *Ann Glob Health*. 2015;81:735-741.
- 28 International Diabetes Federation (IDF). Atlas—9th edition. 2019. <http://www.diabetesatlas.org>. Accessed May 5, 2020.
- 29 Wink M Jr, Paese L. Inequality of educational opportunities: evidence from Brazil. *Economia*. 2019;20:109-120.
- 30 Costa AF, Flor LS, Campos MR, et al. Burden of type 2 diabetes mellitus in Brazil. *Cad Saúde Pública*. 2017;33:e00197915.
- 31 Andrade RE, Nascimento LF, Arruda DM, et al. Diabetes approach through the implementation of a strategic multidisciplinary model of care – “Mutirão do Diabético”. *Diabetes Clínica*. 2012;6:424-433.
- 32 Sarmiento LR, Fernandes PFCBC, Pontes MX, et al. Prevalence of clinically validated primary causes of end-stage renal disease (ESRD) in a State Capital in Northeastern Brazil. *Braz J Nephrol*. 2018;40:130-135.
- 33 Platon I. World Diabetes Day 2012—expanding the Circle of Influence. *Diabetes Res Clin Pract*. 2012;97:514-516.
- 34 Malerbi FK, Morales PHA, Regatieri CVS. Diabetic retinopathy screening and the COVID-19 pandemic in Brazil. *Arq Bras Oftalmol*. 2020;83(4): V-VI. Doi:10.5935/0004-2749.20200070.