

Diagnostic Accuracy of Community-Based Diabetic Retinopathy Screening With an Offline Artificial Intelligence System on a Smartphone

Sundaram Natarajan, MD; Astha Jain, MD; Radhika Krishnan, MD; Ashwini Rogye, B Optom; Sobha Sivaprasad, FRCOphth

IMPORTANCE Offline automated analysis of retinal images on a smartphone may be a cost-effective and scalable method of screening for diabetic retinopathy; however, to our knowledge, assessment of such an artificial intelligence (AI) system is lacking.

OBJECTIVE To evaluate the performance of Medios AI (Remidio), a proprietary, offline, smartphone-based, automated system of analysis of retinal images, to detect referable diabetic retinopathy (RDR) in images taken by a minimally trained health care worker with Remidio Non-Mydriatic Fundus on Phone, a smartphone-based, nonmydriatic retinal camera. Referable diabetic retinopathy is defined as any retinopathy more severe than mild diabetic retinopathy, with or without diabetic macular edema.

DESIGN, SETTING, AND PARTICIPANTS This prospective, cross-sectional, population-based study took place from August 2018 to September 2018. Patients with diabetes mellitus who visited various dispensaries administered by the Municipal Corporation of Greater Mumbai in Mumbai, India, on a particular day were included.

INTERVENTIONS Three fields of the fundus (the posterior pole, nasal, and temporal fields) were photographed. The images were analyzed by an ophthalmologist and the AI system.

MAIN OUTCOMES AND MEASURES To evaluate the sensitivity and specificity of the offline automated analysis system in detecting referable diabetic retinopathy on images taken on the smartphone-based, nonmydriatic retinal imaging system by a health worker.

RESULTS Of 255 patients seen in the dispensaries, 231 patients (90.6%) consented to diabetic retinopathy screening. The major reasons for not participating were unwillingness to wait for screening and the blurring of vision that would occur after dilation. Images from 18 patients were deemed ungradable by the ophthalmologist and hence were excluded. In the remaining participants (110 female patients [51.6%] and 103 male patients [48.4%]; mean [SD] age, 53.1 [10.3] years), the sensitivity and specificity of the offline AI system in diagnosing referable diabetic retinopathy were 100.0% (95% CI, 78.2%-100.0%) and 88.4% (95% CI, 83.2%-92.5%), respectively, and in diagnosing any diabetic retinopathy were 85.2% (95% CI, 66.3%-95.8%) and 92.0% (95% CI, 97.1%-95.4%), respectively, compared with ophthalmologist grading using the same images.

CONCLUSIONS AND RELEVANCE These pilot study results show promise in the use of an offline AI system in community screening for referable diabetic retinopathy with a smartphone-based fundus camera. The use of AI would enable screening for referable diabetic retinopathy in remote areas where services of an ophthalmologist are unavailable. This study was done on patients with diabetes who were visiting a dispensary that provides curative services to the population at the primary level. A study with a larger sample size may be needed to extend the results to general population screening, however.

JAMA Ophthalmol. 2019;137(10):1182-1188. doi:10.1001/jamaophthalmol.2019.2923
Published online August 8, 2019.

← Invited Commentary
page 1188

Author Affiliations: Aditya Jyot Foundation for Twinkling Little Eyes, Mumbai, India (Natarajan, Jain, Krishnan, Rogye); Moorfields Eye Hospital, London, United Kingdom (Sivaprasad).

Corresponding Author: Astha Jain, MD, Aditya Jyot Foundation for Twinkling Little Eyes, Mumbai, 153, Major Parmeswaran Road, Wadala, Mumbai 400031, India (drradhika.ajftle@gmail.com).

Diabetic retinopathy (DR) is a major cause of preventable blindness in the working-age population in many countries of the world.¹ People with diabetes usually remain asymptomatic until an advanced stage of DR. Therefore, screening for sight-threatening complications is necessary to initiate timely treatment.^{1,2} Prevalence of blindness attributable to DR has decreased in countries such as the United Kingdom because of effective population-based screening programs that use desktop retinal cameras to capture 1-field, 2-field, or 3-field mydriatic digital retinal photographs,^{2,3} followed by primary or secondary human grading and an arbitration process.¹⁻³ These are costly, time-consuming, and complex screening programs that require considerable training in the use of cameras, as well as experienced retinal graders, limiting the relevance of these models in developing countries. There is a substantial unmet need for accurate and simple-to-use screening modalities that can be used globally to screen people for DR. Smartphone-based retinal imaging is emerging as a cost-effective way of screening for retinopathy in the community.^{4,5} Similarly, automated analysis of retinal images captured using standard retinal cameras has the promise of being cost-effective and scalable within population-based DR screening programs.^{6,7} Incorporating similar automated analysis into low-cost, smartphone-based devices has been shown to be acceptable for screening.⁸

To date, most automated algorithms use deep-learning and neural networks that require a processor-intensive environment for inferencing, resulting in images needing to be transferred to the cloud. However, there are many parts of the world where access to a stable internet connection is not assured. This study validates the performance of an offline automated analysis algorithm that runs directly off a smartphone. To our knowledge, this is the first study evaluating an offline artificial intelligence (AI) algorithm to detect DR using an affordable, easy-to-use, smartphone-based imaging device.

Methodology

Fundus images were captured using the Remidio Non-Mydriatic Fundus on Phone (Remidio Innovative Solutions Pvt Ltd). The images so captured were subjected to automated analysis by the Medios AI (Remidio), a proprietary offline automated analysis of retinal images on a smartphone to detect referable diabetic retinopathy (RDR) on images taken by a health care worker on a smartphone-based, nonmydriatic retinal camera. These were also graded by a vitreoretinal resident physician and a vitreoretinal surgeon (A.J.) who were masked to the results from the AI system.

Institutional review board approval was obtained from the Aditya Jyot Eye Hospital Ethics Committee. Informed consent was obtained from all participants. The protocol adhered to the tenets of the Declaration of Helsinki. Both the offline automated analysis and the smartphone-based, nonmydriatic retinal imaging system are based on proprietary technologies. However, authors of the study have no financial interest in these technologies.

Key Points

Question What is the performance of an offline, automated artificial intelligence system of analysis to detect referable diabetic retinopathy on images taken by a health worker on a smartphone-based, nonmydriatic retinal camera?

Finding In this cross-sectional study, fundus images from 213 study participants were subjected to offline, automated analysis. The sensitivity and specificity of the analysis to diagnose referable diabetic retinopathy were 100.0% and 88.4%, respectively, and the sensitivity and specificity for any diabetic retinopathy were 85.2% and 92.0%, respectively.

Meaning This study suggests these methods might be used to screen for referable diabetic retinopathy using offline artificial intelligence and a smartphone-based, nonmydriatic retinal imaging system.

Capture of Retinal Images

This was a prospective, cross-sectional study of diagnostic accuracy. Patients with diabetes mellitus who were visiting the various dispensaries administered by the Municipal Corporation of Greater Mumbai in Mumbai, India, on a particular day were screened for DR using the portable, smartphone-based, nonmydriatic retinal imaging system. Preliminary data, such as age, sex, duration since diabetes onset, and postprandial blood glucose level, were collected. Patients' eyes were dilated using single drop of tropicamide eyedrops, 1%, which has previously been found to cause minimal risk of angle-closure glaucoma.⁹ Fundus imaging was then done by a health care worker with no professional experience in the use of fundus cameras. An anterior segment photograph was first captured, followed by 3 fields of the fundus (namely, the posterior pole, including the disc and macula, and the nasal and temporal fields), as per the Early Treatment Diabetic Retinopathy Study protocol (Figure 1). The offline AI algorithm on the smartphone flags images of poor quality, prompting the operator to take additional pictures of the same retinal view until the images were deemed acceptable by the AI system.

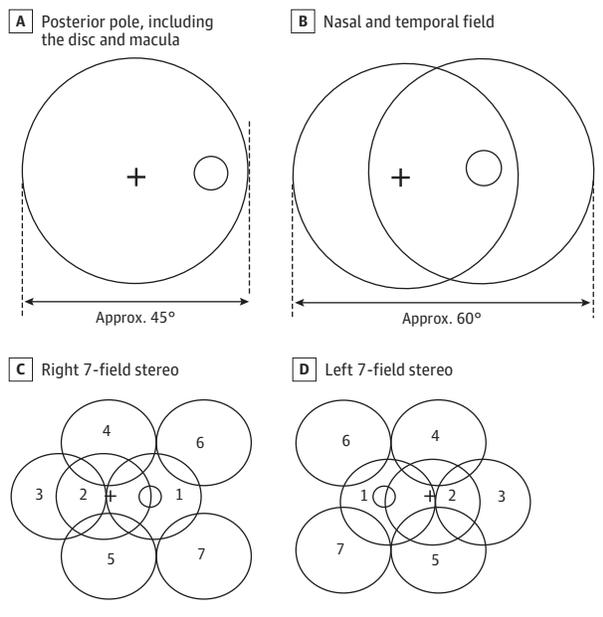
Grading by Human Graders

The images were stored on a Health Insurance Portability and Accountability Act-compliant cloud server (Amazon Web Services) and graded by a vitreoretinal resident and a vitreoretinal surgeon (A.J.) at the Aditya Jyot Eye Hospital and Aditya Jyot Foundation for Twinkling Little Eyes in Mumbai, India, who were masked to the AI grading results. In case of a discrepancy between the grading of the resident and surgeon, the diagnosis of the surgeon was considered final. The grading of retinopathy was done according to the International Clinical DR severity scale.¹⁰ The final diagnosis for each patient was determined by the stage of DR of the more affected eye per the International Clinical DR severity scale. Patients whose image of 1 or both eyes was considered ungradable were excluded from the AI analysis.

Grading by Offline AI System

The offline automated analysis application is integrated into the smartphone-based, nonmydriatic retinal imaging system. It is

Figure 1. Seven-Segment Early Treatment Diabetic Retinopathy Study Protocol



a component of the camera control app and thus seamlessly integrates into the image acquisition workflow. It can be broadly divided into 2 core components. First, an algorithm checks the quality of the captured images. A second DR assessment mechanism generates a diagnosis by detecting DR lesions. This mechanism relies on 2 convolutional neural networks.

Captured images are first processed by a cropping algorithm. This removes the black border surrounding the circular field of view generated by the fundus camera. They are then down-sampled to a standardized image size. A first neural network assesses the quality of the image. This network is based on the MobileNet architecture. It has been trained with fundus images flagged as ungradable, as well as images of sufficient quality. The user is advised to recapture the image if the result of the neural network is negative.

Two neural networks have been trained separately to detect DR. They consist of binary classifiers based on the Inception-V3 (Google) architecture that separate healthy images from images with referable DR (defined as moderate nonproliferative DR and cases of greater severity). No images with mild nonproliferative DR have been used during training. The training set consisted of 34 278 images from the Eye Picture Archive Communication System (EyePACS) data set, 14 266 images taken with a Kowa VX-10a mydriatic camera at Diacon Hospital in Bangalore, India, and 4350 nonmydriatic images taken in screening camps by the smartphone-based, nonmydriatic retinal imaging system. The data set has been curated to contain as many referral cases as healthy ones. It has also been curated to contain images taken in a variety of conditions, including with nonmydriatic and low-cost cameras.

Three different data sets were used for internally validating networks and ensembling them. These are separate from the training data. Results on these data sets are shown in

Table 1. Data set 1 consists of images taken with the mydriatic version of the smartphone-based, nonmydriatic retinal imaging system at Dr Mohan's Diabetes Specialties Center in Chennai, India. Data set 2 consists of images taken with the mydriatic mode (one of several modes available on this more recent device) of the smartphone-based, nonmydriatic retinal imaging system at Diacon Hospital in Bangalore, India. These institutions only provided images with their diagnosis and were not involved in computing the results.

One of the 2 networks has been trained directly on the captured images, while the other works on images that underwent an image-processing algorithm to boost their contrast. The contrast-enhancement algorithm has been empirically optimized to make DR lesions stand out in the input images. Both outputs of each network are then fed to a linear classifier that computes the final assessment of an image. This follows the ensemble learning paradigm. It improves the accuracy by combining several classifiers trained under different strategies. Common data-augmentation techniques, such as rotations, flipping, and zooming, were applied to both networks. Final referral recommendations are given on a patient level. A patient was considered to have referable disease if any image was flagged as referable by the algorithm.

Class activation mapping¹¹ was also implemented. This gives a visual feedback to the physician by displaying the areas of the fundus image that have triggered a positive diagnosis. Examples of outputs are given in **Figure 2**.

The whole system has been implemented directly on the iPhone 6 (Apple) using high-performance image processing techniques based on CoreML version 2.0 (Apple) and Open Graphics Library ES 2.0 (Silicon Graphics), leveraging on the graphics processing unit of the device instead of an internet connection to a remote server. Both the image processing algorithms and the neural networks run in seconds on an iPhone 6 (Apple).

The same images of these patients were graded by the offline automated analysis algorithm to have either referable DR or no DR. The AI algorithm was run offline on the smartphone by the operator immediately after image acquisition. Adjudication of images that presented results that differed between the resident, surgeon, and AI system was handled by the vitreoretinal surgeon (A.J.) at Aditya Jyot Foundation for Twinkling Little Eyes.

The offline automated analysis is designed to binary-type only RDR and no DR. It does not grade the stages of DR, such mild nonproliferative DR, moderate nonproliferative DR, severe nonproliferative DR, and proliferative DR.

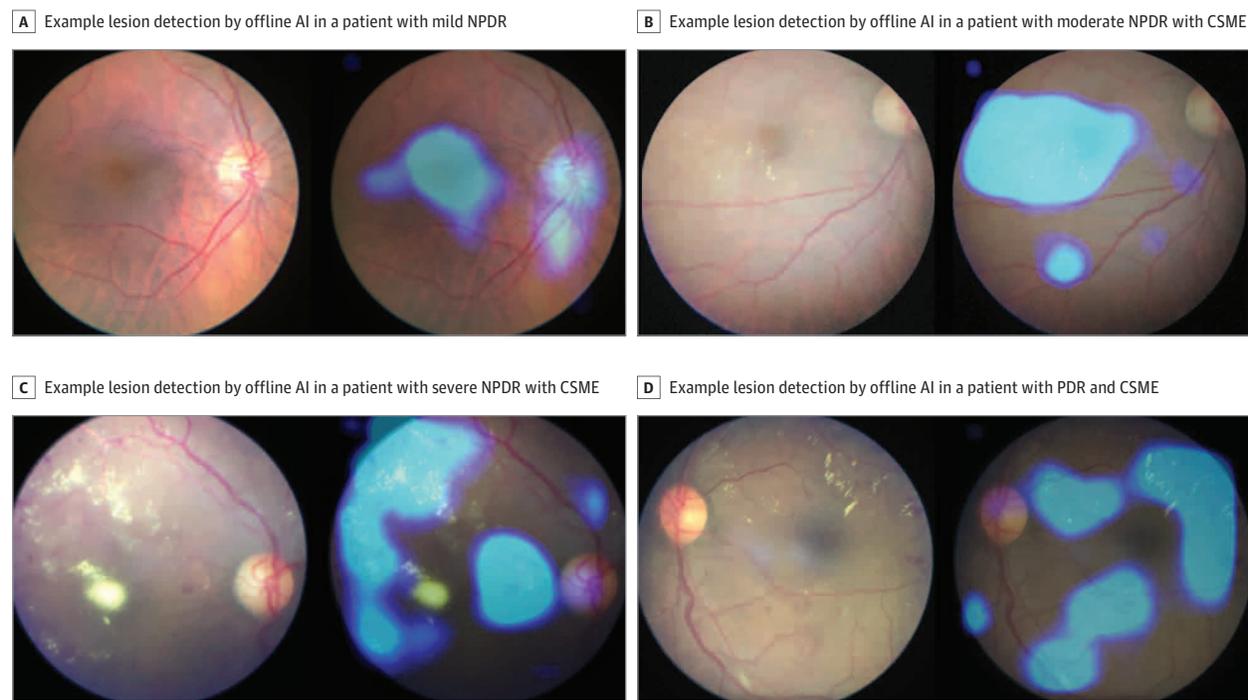
Statistical Analysis

Sensitivity and specificity statistics were computed for both any DR as well as RDR, assuming the ground truth to be the evaluation of the same patient images by the vitreoretinal resident and surgeon. The eye with the more severe retinopathy grading was considered the patient-level grading. The minimum number of patients needed to be screened in an opportunistic screening context, such as an outreach center in India, was calculated, assuming a margin of error of 7% for this study on either side of the mean (given a previously published mandate from the US Food and Drug Administration of an end point of at least 86% diagnostic sensitivity of RDR).¹²

Table 1. Medios Artificial Intelligence Internal Validation: Performance Results on Data Sets

Data Set	Images, No.	Patients, No.	Referable Diabetic Retinopathy		Any Diabetic Retinopathy	
			Sensitivity, %	Specificity, %	Sensitivity, %	Specificity, %
1	3038	301	95.9	81.3	86.2	99.1
2	1054	165	100.0	78.7	77.1	91.3

Figure 2. Class Activation Mapping in Mild, Moderate, and Severe Nonproliferative Diabetic Retinopathy (NPDR) and Proliferative Diabetic Retinopathy (PDR)



AI indicates artificial intelligence; CSME, clinically significant macular edema.

At the 95% CI, and a population of 18.4 million in Mumbai (census of 2011), with a maximum DR prevalence of 20%,¹³ this resulted in a minimum sample size of 200 patients.

Results

Of the 255 patients seen at the dispensary on the day the study was conducted, 231 consented for DR screening. The major reasons given by the 24 nonparticipating patients for refusal were the unwillingness to wait for screening and the blurring of vision that would occur after dilation. Images of one or both eyes from 18 patients were deemed ungradable by the ophthalmologist, and these individuals were therefore excluded. Hence, a total of 213 patients were analyzed for assessment of RDR using AI, with 1 or more ophthalmologists grading the same images regarded as the ground truth (defined as a direct observation serving as the gold standard). Among the 213 included patients, there were 110 female patients and 103 male patients. The mean (SD) age of the participants was 53.1 (10.3) years. The mean (SD) postprandial blood glucose level was 207.8 (74.7) mg/dL (to convert to millimoles per liter, multi-

ply by 0.0555), and the mean (SD) duration since diabetic disease onset was 5.5 (4.75) years.

A total of 187 patients were diagnosed as having no DR by ophthalmologist grading. Of these, 172 patients were correctly diagnosed by AI, whereas 15 patients (8.0%) were incorrectly diagnosed as having RDR. Fifteen patients (8.0%) were identified as having RDR by ophthalmologist grading, and all 15 (100.0%) were correctly diagnosed by the AI. Among 12 individuals with cases of mild nonproliferative DR who were diagnosed by the ophthalmologists, 8 patients (67%) were diagnosed as having RDR by the AI, while 4 (33%) were diagnosed as not having DR. This gave a sensitivity and specificity of diagnosing RDR as 100% (95% CI, 78.2%-100%) and 88.4% (95% CI, 83.16%-92.53%), respectively; the same values for any DR were 85.2% (95% CI, 66.3%-95.8%) and 92.0% (95% CI, 97.1%-95.4%), respectively (Table 2). There was excellent intergrader agreement between the vitreoretinal resident and the vitreoretinal surgeon (eyewise grading: minimum $\kappa = 0.89$ [SE, 0.05]; clinically significant macular edema grading: minimum $\kappa = 0.77$ [SE, 0.06]).

Since the image capture was done by a nonspecialist, there were images that did not meet the minimum image quality

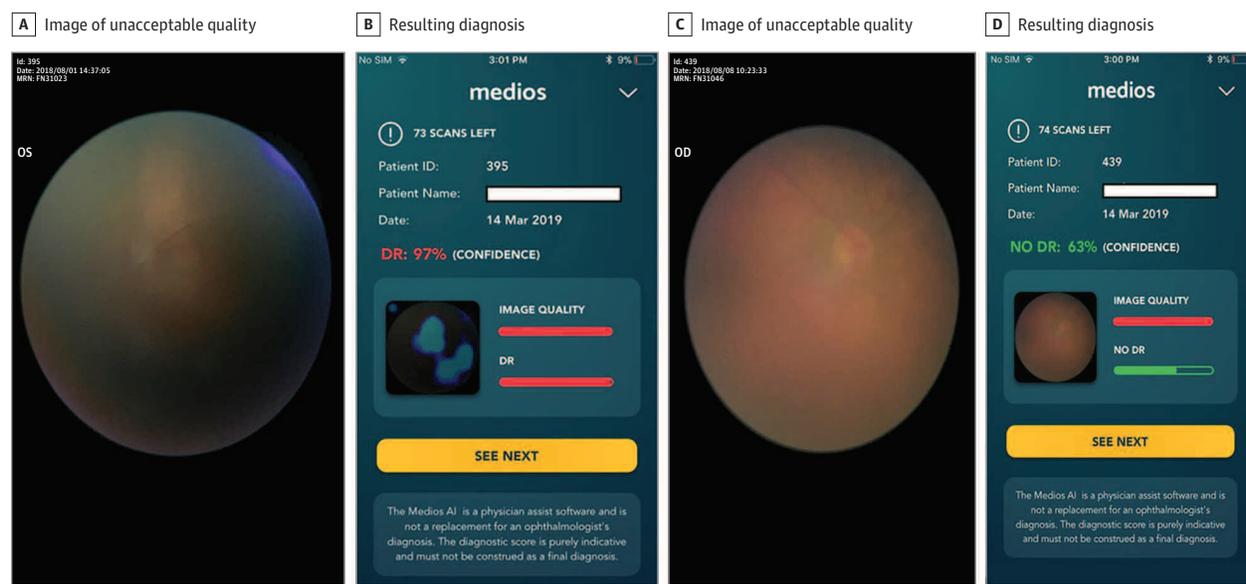
Table 2. Medios Offline Artificial Intelligence Diagnoses vs Ophthalmologist Diagnoses

Diabetic Retinopathy Diagnosis	Medios Artificial Intelligence				By Eye	
	By Patient, After Excluding Poor-Quality Images		By Patient, With Poor-Quality Images Included		Referable Diabetic Retinopathy	No Diabetic Retinopathy
	Referable Diabetic Retinopathy	No Diabetic Retinopathy	Referable Diabetic Retinopathy	No Diabetic Retinopathy		
Ophthalmologist grading ^a						
None	15	172	28	159	35	317
Nonproliferative						
Mild	8	4	8	4	11	8
Moderate	12	0	12	0	17	0
Severe	2	0	2	0	4	0
Proliferative	1	0	1	0	2	0
Diagnostic accuracy, %						
Referable						
Sensitivity	100.0		100.0		100.0	
Specificity	88.4		81.9		87.6	
Any		NA		NA		NA
Sensitivity	85.2		85.2		81.0	
Specificity	92.0		85.0		90.1	

Abbreviation: NA, not applicable.

^a Ground truth.

Figure 3. Artificial Intelligence Analysis of Poor-Quality Images



requirement of the AI. A separate analysis was performed wherein the health care worker was asked to use all images taken in every patient, including the images not meeting the minimum quality standards of the AI, in assessing the diagnostic accuracy of the AI (Figure 3). The sensitivity for detection of RDR continued to remain 100.0% (95% CI, 78.2%-100.0%), while the specificity dropped to 81.9% (95% CI, 75.9%-87.0%) as a result of an increase in the number of no DR cases graded as RDR by the AI, consequent to the inclusion of the poorer-quality images (Table 2). In fact, eyewise analysis of the same data set with the inclusion of poor-quality images showed sensitivity of detection of RDR of

100.0% (95% CI, 85.2%-100.0%), specificity of detection of RDR of 87.6% (95% CI, 83.8% to 90.8%), sensitivity of detection of any DR of 81.0% (95% CI, 65.9% to 91.4%), and specificity of detection of any DR of 90.1% (95% CI, 86.4% to 93.0%), as shown in Table 2.

Discussion

This study evaluates the diagnostic accuracy of an offline AI algorithm for detection of RDR on images taken from a smartphone-based portable camera. To our knowledge, this is the

first study assessing an offline AI algorithm on a smartphone for detection of RDR.

In a developing country such as India, nearly 70% of the population resides in rural areas, and a ratio of only 1 ophthalmologist per 100 000 people¹⁴ is available for the care of the entire population. This study shows how an offline AI algorithm can help address this lack of specialist access through automatic, instant grading of the retinal images, highlighting a possible solution for implementation of large-scale models for screening for RDR. Various software have been used previously for automated detection of DR.^{6,7,12} The current study uses the offline automated analysis that gives results in real time.

The sensitivity for detection of RDR remained at 100.0% in both the eyewise analysis as well as the patientwise analysis, pointing to an inherent robustness of the offline algorithm in screening for RDR. Similar high sensitivity was found in the study by Gulshan et al⁶ albeit in an in-clinic retrospective study, while in the EyePACS-1 data set, the sensitivity was 97.5% and the specificity was 93.4%. In the Messidor data set, the sensitivity was 96.1% and the specificity was 93.9% for detecting RDR. A study using the cloud-based software EyeArt (Eyenuk) showed a sensitivity and specificity of 95.8% and 80.2%, respectively, for detecting any DR and 99.1% and 80.4%, respectively, in detecting sight-threatening DR,⁸ using the same smartphone-based retinal imaging system technology, albeit via an earlier model of the device offering mydriatic imaging alone. In a multiethnic study, Ting et al⁷ showed a sensitivity and specificity of 90.5% and 91.6%, respectively, for RDR and 100% and 91.1%, respectively, for vision-threatening DR using a conventional, desktop fundus camera.

The specificity seen in this study for detecting RDR may have been slightly lower because many mild nonproliferative DR cases have been identified as RDR by the offline automated analysis. This is because the offline AI was purposely not trained on mild nonproliferative DR images, to ensure high specificity in no DR and RDR diagnoses. Retinal lesions other than DR, such as retinitis pigmentosa, drusen, and retinal pigment epithelium changes, were overdiagnosed by the AI as RDR. Although incorrectly labeled as RDR, these cases likely would warrant a referral to an ophthalmologist.

The major advantage of the offline, automated analysis over the previously used deep-learning algorithms^{6,7,12,15,16} may be that it can be used offline on a smartphone and would not require internet access for real-time transfer of images, which enables results to be given to patients immediately. The AI also provides a lesion detection map on the images (Figure 2), which guides the health worker and educates the patient.

Unlike previous retrospective studies that have assessed the performance of the AI algorithms in an in-clinic setting with images that are usually of excellent quality, this study involved the use of the smartphone-based, nonmydriatic retinal imaging system in the field by a health care worker who was trained for less

than 2 weeks on how to use the device. Thus, not all the images were of excellent quality, which is representative of what would typically be expected in large-scale, opportunistic community screenings. Even when the offline AI system was subjected to images with quality deemed unacceptable by the AI (Figure 3), the sensitivity for RDR and any DR of the offline automated analysis remained unchanged, while the specificity for RDR and any DR decreased by 7% as a result of some cases of no DR being incorrectly graded as RDR. The superiority end point deemed by the FDA in the pivotal clinical trials evaluating the IDx AI algorithm was a sensitivity of 85% and a specificity of 82.5%.¹² The offline automated analysis algorithm provides a sensitivity and specificity to detect RDR of 100.0% and 88.4%, and this was above the defined thresholds.

Limitations

One limitation of this study is the small sample size of the test population on which the offline automated analysis was tested compared with other studies in literature.^{6,7,12} Nevertheless, to estimate the specificity and sensitivity with a lower margin of error of less than 3% (at the 95% CI), a larger sample size of nearly 1050 patients will be needed, assuming a prevalence of up to 20% of DR in the Mumbai population.¹³ Another drawback is that the current version of the offline AI does not permit grading of retinopathy according to the International Clinical DR severity scale or the National Health Service classification. Hence, this study is unable to assess the sensitivity and specificity of the offline AI for detection of individual grades of DR.

In this study, the analysis included imaging pupils as small as 3 mm. However, nonmydriatic imaging protocols in an Indian population have shown a large percentage of ungradable images, owing to the comorbidity of cataract and smaller mesopic pupil sizes.¹⁷ This necessitated dilated retinal photography in screening for RDR, especially when implementing 3-segment digital retinal imaging protocols.¹⁸

Conclusions

The Municipal Corporation of Greater Mumbai, India, where the study was conducted, is the largest municipal corporation in the country, with 174 dispensaries and 210 health posts across the city. Of these, all 174 dispensaries run diabetes management services through monitoring and basic treatment. Of the 255 patients eligible for inclusion, 231 patients (90.6%) visiting the centers received examinations using a dilated imaging protocol, pointing to convenience, affordability, and access to instant reporting driving the demand for a proper DR screening with dilation. A validation of these findings on a larger data set that enables precise assessment of specificity and sensitivity with a standard error less than 3% is currently in progress.

ARTICLE INFORMATION

Accepted for Publication: May 26, 2019.

Published Online: August 8, 2019.

doi:10.1001/jamaophthalmol.2019.2923

Open Access: This article is published under the JN-OA license and is free to read on the day of publication.

Author Contributions: Drs Natarajan and Jain had full access to all the data in the study and take

responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Natarajan, Jain, Krishnan, Sivaprasad.

Acquisition, analysis, or interpretation of data:

Rogye.

Drafting of the manuscript: Jain, Rogye.

Critical revision of the manuscript for important intellectual content: Natarajan, Krishnan, Sivaprasad.

Administrative, technical, or material support: Krishnan.

Supervision: Natarajan, Jain, Sivaprasad.

Conflict of Interest Disclosures: Dr Sivaprasad reported grants and personal fees from Novartis, Allergan, Boehringer Ingelheim, Roche, and Optos and grants, personal fees, and nonfinancial support from Bayer outside the submitted work. No other disclosures were reported.

Additional Contributions: We acknowledge Medios Technologies, Singapore, for providing the AI software for conducting the study and Florian M. Savoy, MSc, Medios Technologies, for authoring the paragraphs describing the technical design of the AI software in the methodology section. We acknowledge Anand Sivaraman, PhD, Remidio Innovative Solutions Pvt Ltd, for technical discussions on study design, training on the NM FOP 10 device, and inputs on statistical analysis. We acknowledge Diacon Hospital, Bangalore, India, and Dr Mohan's Diabetes Specialities Center, Chennai, India, for providing images for internal validation. The individuals named were not compensated.

REFERENCES

- Squirrell DM, Talbot JF. Screening for diabetic retinopathy. *J R Soc Med*. 2003;96(6):273-276. doi:10.1177/014107680309600604
- Bachmann MO, Nelson SJ. Impact of diabetic retinopathy screening on a British district population: case detection and blindness prevention in an evidence-based model. *J Epidemiol Community Health*. 1998;52(1):45-52. doi:10.1136/jech.52.1.45
- Scanlon PH. The English national screening programme for diabetic retinopathy 2003-2016. *Acta Diabetol*. 2017;54(6):515-525. doi:10.1007/s00592-017-0974-1
- Rajalakshmi R, Arulmalar S, Usha M, et al. Validation of smartphone based retinal photography for diabetic retinopathy screening. *PLoS One*. 2015;10(9):e0138285. doi:10.1371/journal.pone.0138285
- Sengupta S, Sindal MD, Baskaran P, Pan U, Venkatesh R. Sensitivity and specificity of smartphone based retinal imaging for diabetic retinopathy: a comparative study. *Ophthalmol Retina*. 2019;3(2):146-153. doi:10.1016/j.oret.2018.09.016
- Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316(22):2402-2410. doi:10.1001/jama.2016.17216
- Ting DS, Cheung CY, Lim G, 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(22):2211-2223. doi:10.1001/jama.2017.18152
- Rajalakshmi R, Subashini R, Anjana RM, Mohan V. Automated diabetic retinopathy detection in smartphone-based fundus photography using artificial intelligence. *Eye (Lond)*. 2018;32(6):1138-1144. doi:10.1038/s41433-018-0064-9
- Pandit RJ, Taylor R. Mydriasis and glaucoma: exploding the myth, a systematic review. *Diabet Med*. 2000;17(10):693-699. doi:10.1046/j.1464-5491.2000.00368.x
- Wilkinson CP, Ferris FL III, Klein RE, et al; Global Diabetic Retinopathy Project Group. Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. *Ophthalmology*. 2003;110(9):1677-1682. doi:10.1016/S0161-6420(03)00475-5
- Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A. Learning deep features for discriminative localization. Proceedings of the IEEE conference on computer vision and pattern recognition. http://cnlocalization.csail.mit.edu/Zhou_Learning_Deep_Features_CVPR_2016_paper.pdf. Published 2016. Accessed July 8, 2019.
- Abramoff M, Lavin PT, Birch M, Shah N, Folk J. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ Digital Medicine*. 2018;1(39):1-8. doi:10.1038/s41746-018-0040-6
- Sunita M, Singh AK, Rogye A, et al. Prevalence of diabetic retinopathy in urban slums: the Aditya Jyot Diabetic Retinopathy in Urban Mumbai slums study, report 2. *Ophthalmic Epidemiol*. 2017;24(5):303-310. doi:10.1080/09286586.2017.1290258
- Lundquist BM, Sharma N, Kewalramani K. Patient perceptions of eye disease and treatment in Bihar, India. *J Clin Experiment Ophthalmol*. 2012;3:2. doi:10.4172/2155-9570.1000213
- Walton OB IV, Garoon RB, Weng CY, et al. Evaluation of automated teleretinal screening program for diabetic retinopathy. *JAMA Ophthalmol*. 2016;134(2):204-209. doi:10.1001/jamaophthalmol.2015.5083
- Kumar PNS, Deepak RU, Sathar A, Sahasranamam V, Kumar RR. Automated detection system for diabetic retinopathy using two field fundus photography. *Procedia Comput Sci*. 93: 486-494.
- Gupta V, Bansal R, Gupta A, Bhansali A. Sensitivity and specificity of nonmydriatic digital imaging in screening diabetic retinopathy in Indian eyes. *Indian J Ophthalmol*. 2014;62(8):851-856. doi:10.4103/0301-4738.141039
- Vujosevic S, Benetti E, Massignan F, et al. Screening for diabetic retinopathy: 1 and 3 nonmydriatic 45-degree digital fundus photographs vs 7 standard early treatment diabetic retinopathy study fields. *Am J Ophthalmol*. 2009;148(1):111-118. doi:10.1016/j.ajo.2009.02.031

Invited Commentary

Smartphone-Based, Artificial Intelligence-Enabled Diabetic Retinopathy Screening

T. Y. Alvin Liu, MD

Artificial intelligence, specifically deep learning (DL), has garnered substantial attention in the medical world in recent years because it has been used to accurately classify medical images, often at a human expert level, across different medical disciplines. Within ophthalmology, the prospect of applying DL in screening for diabetic

retinopathy (DR) is particularly exciting, because DR is the most common cause of visual acuity loss among the working-age population in the United States¹ and many countries around the world, and early detection of DR is cost-effective.² While numerous studies have demonstrated the ability of DL to detect DR in color fundus photographs with high sensitivity and specificity, most of these studies were performed using retro-

spectively collected data. Also, there is evidence to suggest that artificial intelligence-based diagnostic tools, when deployed in a clinical practice setting, can have greatly diminished performance because of variation in disease prevalence and image quality.³

In a landmark prospective clinical trial, Abramoff et al⁴ addressed the feasibility and efficacy of deploying DL-enabled DR screening in the primary care setting. Their proprietary system achieved a sensitivity of 87.2% and a specificity of 90.7%, exceeding all prespecified superiority end points, and received US Food and Drug Administration approval for the first fully automatic diagnostic system for the detection of more-than-mild nonproliferative diabetic retinopathy. A study in this issue of *JAMA Ophthalmology* by