Medios– An offline, smartphone-based artificial intelligence algorithm for the diagnosis of diabetic retinopathy

Bhavana Sosale, Aravind R Sosale, Hemanth Murthy[1], Sabyasachi Sengupta[2], Muralidhar Naveenam[1]

Purpose: An observational study to assess the sensitivity and specificity of the Medios smartphone-based offline deep learning artificial intelligence (AI) software to detect diabetic retinopathy (DR) compared with the image diagnosis of ophthalmologists. **Methods:** Patients attending the outpatient services of a tertiary center for diabetes care underwent 3-field dilated retinal imaging using the Remidio NM FOP 10. Two fellowship-trained vitreoretinal specialists separately graded anonymized images and a patient-level diagnosis was reached based on grading of the worse eye. The images were subjected to offline grading using the Medios integrated AI-based software on the same smartphone used to acquire images. The sensitivity and specificity of the AI in detecting referable DR (moderate non-proliferative DR (NPDR) or worse disease) was compared to the gold standard diagnosis of the retina specialists. **Results:** Results include analysis of images from 297 patients of which 176 (59.2%) had no DR, 35 (11.7%) had mild NPDR, 41 (13.8%) had moderate NPDR, and 33 (11.1%) had severe NPDR. In addition, 12 (4%) patients had PDR and 36 (20.4%) had macular edema. Sensitivity and specificity of the AI in detecting referable DR was 98.84% (95% confidence interval [CI], 97.62–100%) and 86.73% (95% CI, 82.87–90.59%), respectively. The area under the curve was 0.92. The sensitivity for vision-threatening DR (VTDR) was 100%. **Conclusion:** The AI-based software had high sensitivity and specificity in detecting referable DR. Integration with the smartphone-based fundus camera with offline image grading has the potential for widespread applications in resource-poor settings.

Key words: Artificial intelligence, deep learning, diabetic retinopathy

Around five million Indians have vision-threatening diabetic retinopathy (VTDR).[1,2] Smartphone-based fundus cameras and evolution of artificial intelligence (AI) can make screening scalable.[3–8] High computational power and Internet access—a prerequisite for cloud-based AI—is often lacking in developing countries. Medios Technologies, Singapore to our knowledge is the first company to develop an offline AI algorithm to address this obstacle.

Our aim is to evaluate the performance of an offline AI-based software (Medios Technologies, Singapore) loaded on a smartphone-based fundus camera in the detection of diabetic retinopathy (DR) compared with the image diagnosis of ophthalmologists.

Methods

The study was approved by the institutional ethics committee and carried out as per the declaration of Helsinki. Patients who consented to have their eye dilated and photographed during routine care were included.

This was a cross-sectional observational study of 304 diabetic patients attending the outpatient department (OPD) of a university-recognized tertiary center for diabetes care

and research in Bangalore, India during the month of October 2018. All subjects, above 18 years of age, with type 1 or 2 diabetes or secondary diabetes were invited to participate. Eyes with significant media opacity, such as corneal opacity or advanced cataracts, that precluded retinal imaging were excluded.

At the time of the patient's hospital visit, routine clinical care procedures like the collection of demographic data, medical history, vital measurements, anthropometric measurements, and general physical exams were carried out.

Retinal image acquisition

A drop of 1% tropicamide solution was used to dilate the pupils to a minimum size of 5 mm. Retinal images were captured using the smartphone-based "Remidio Non Mydriatic Fundus on Phone Camera (NM FOP 10)" (Remidio Innovative Solutions Pvt. Ltd., Bangalore, India) by a trained technician.[3] Three fields of view (FOVs) were captured from each eye—posterior pole (macula centered), nasal field, and superotemporal field. The technician was trained to recognize characteristics of an excellent image and urged to capture more than one image per FOV if required to obtain excellent images.

Department of Diabetology, Diacon Hospital, [1]Department of Vitreo-Retinal Surgery, Retina Institute of Karnataka, [2]Department of Vitreo-Retinal Surgery, Future Vision Eye Care, Mumbai, Maharashtra, India

Correspondence to: Dr. Bhavana Sosale, 360, Diacon Hospital, 19th Mail, 1st Block, Rajajinagar, Bengaluru - 560 010, Karnataka, India. E-mail: bhavanasosale@gmail.com

The Remidio NM FOP 10 device uses an iPhone 6 smartphone's camera to capture images of the retina either by using an infrared light emitting diode (IRLED) based live view (nonmydriatic mode) or using a warm white LED live view (the mydriatic mode). As per the Apple's official website, the iPhone 6 specifications include a camera with a 8 MP (2448 × 3264 pixels) resolution, a screen display resolution of 750 × 1334 pixels, and a 1.4 GHz Cyclone processor paired with 1 GB of RAM. During the course of this study, the phone used an iOS 10 mobile operating system and came preloaded with the integrated Remidio NM FOP 10 app with the Medios AI. The manufacturer stated resolution of the Remidio NM FOP camera was a minimum of 80 line pairs/mm, conforming to the requirements of ISO10940 standard.

Image grading by a vitreoretinal specialist

The de-identified (anonymized) images with the subject ID were uploaded online from the NM FOP 10 device to an Amazon Web Services (AWS) hosted cloud service provided by the manufacturer. The images were accessed from the cloud by two fellowship-trained vitreoretinal surgeons with more than 20 years of experience in treating DR. Both the retinal surgeons, grading and adjudicating the images, were affiliated to a different hospital; hence, they remained unbiased, being masked to the clinical and AI diagnosis.

The retinal surgeons individually graded the set of three retinal photographs from every eye using the International Clinical Diabetic Retinopathy Severity Scale System (ICDRS).[9] Images were graded as no DR, mild nonproliferative DR (mild NPDR), moderate nonproliferative DR (moderate NPDR), severe nonproliferative DR (severe NPDR), and proliferate DR (PDR). The diagnosis of diabetic macular edema (DME) was recorded as present or absent. The eye with the more severe stage of retinopathy was considered as the final diagnosis for that patient, in cases where each eye had a different stage of disease severity. Patients whose images were considered ungradable by the retina specialists were excluded from the final analysis. Whenever the two graders differed on the diagnosis, a consensus was reached by revisiting the images, discussing them, and reaching a mutual agreement. The adjudicated patient diagnosis obtained from the retina specialists was considered as the gold standard for comparisons.

The clinical diagnosis was not considered as the reference standard in this study. The retina specialists affiliated to the tertiary diabetes hospital conducting the study were many—each covering the outpatient clinic on rotation on different days of the week. This made adjudication of clinical diagnosis (necessary to overcome the interobserver variability) impossible. Based on the studies on intergrader variability and evaluation of machine learning models, an adjudicated image diagnosis was considered as the ground truth.[10]

Image analysis using AI-based offline software

The AI-based automated image analysis software used for the study was designed by Medios Technologies, Singapore, a subsidiary of Remidio Innovative Solutions. The Medios AI algorithm is based on convolutional neural networks (CNN). The AI consists of a first neural network for image quality assessment and two other neural networks that detect DR lesions. The network responsible for quality assessment is based on a MobileNet architecture. It consists of a binary classifier and a message prompts the user to recapture the image if it fails the quality check (QC).

The neural network has been trained to separate healthy fundus images (No DR) with images with referable DR (defined as moderate NPDR and above). This maximizes the sensitivity for referable DR and the specificity to rule out all grades of DR. A comprehensive dataset consisting of images taken in a variety of conditions has been used for training, with a proportion of it taken using nonmydriatic and/or low-cost cameras. These include 4350 nonmydriatic images taken during screening camps with the Remidio Fundus-on-Phone; 14,266 images captured with a KOWA Vx-10 mydriatic camera; and 34,278 images come from the EyePACS dataset. A final per-patient DR diagnosis was computed from the outputs of the neural networks and applied to all images of that patient. A patient was deemed as referable if the prediction for one or more images was positive.

The Medios software is integrated with the Remidio NM-FOP application loaded on the smartphone used to acquire images. Thanks to leveraging on the high-performance capabilities of the smartphone with CoreML and OpenGL, image processing is done directly on the graphics processing unit (GPU) instead of relying on a connection to a server on the Internet.

In this study, the AI algorithm was run offline by the technician on the smartphone itself after the images were acquired. The technician was trained to recapture images if the AI gave an alert of "poor image quality." The AI QC was first to run on images and then the diagnosis of the AI was recorded as a binary output, i.e., DR present and No DR. All images captured during this study met the quality standards of the AI and were included in the analysis.

Outcome measure

The primary aim of this study was to determine the sensitivity, specificity, positive predictive value (PPV), and negative PV (NPV) of the AI algorithm in detecting RDR (referable DR was defined as moderate NPDR or more severe disease, or the presence of DME) compared with the gold standard diagnosis by retina specialists. The secondary aims were to assess the sensitivity and specificity of the AI algorithm in the diagnosis of "any DR" and VTDR. Any DR was defined as mild NPDR or more severe disease or the presence of DME, while VTDR was defined as severe NPDR or more severe disease or the presence of DME. An adjudicated image diagnosis was considered as the ground truth.

Statistical analysis

Continuous variables are presented as mean with standard deviation (SD) and categorical variables are presented as proportions (n, %). The sensitivity, specificity, PPV, and NPV of the AI in the detection of referable DR, any DR and VTDR were calculated along with 95% confidence interval (CI). The area under the receiver operating curve (AUROC) was plotted. Cohen's kappa (κ) was measured to assess intergrader variability. All data were stored in Microsoft Excel and analyzed using the Stata software (StataCorp 14.2, Texas, USA).

Results

The study population had a mean age of 55 ± 11 years, duration of diabetes 11 ± 8 years, hemoglobin A1c (HbA1c) 8 ± 2%, and

body mass index (BMI) 27 ± 4 kg/m². Females constituted 42% (*n* = 128) of the study population. The final analysis included images from 297 patients. The images obtained from either one or both eyes of the 7 (2.3%) patients were considered ungradable by the retina specialists and excluded. There was no evidence of DR in 176 participants (59.2%), but mild NPDR was seen in 35 (11.7%), moderate NPDR in 41 (13.8%), severe NPDR in 33 (11.1%), and PDR in 12 (4%) patients. DME was present is 36 (20.4%) individuals with DR with different grades of NPDR or PDR. The intergrader agreement between the retina specialists was 0.89 for grading of DR and 0.9 for grading of DME.

The sensitivity and specificity of the AI algorithm in detecting RDR was 98.84% (95% CI, 97.62–100%) and 86.73% (95% CI, 82.87–90.59%), respectively, while the PPV was 75.22% (95% CI, 70.31–80.13%) and NPV was 99.46% (95% CI, 98.62–100%). The AUROC was 0.92 [Fig. 1]. An example of the output from the Medios AI algorithm along with the respective image with a diagnosis of referable DR is shown in Fig. 2.

For any DR, the sensitivity and specificity of the AI algorithm was 86.78% (95% CI, 82.92–90.63%) and 95.45% (95% CI, 93.09–97.82%), respectively, while the PPV and NPV were 92.92% (95% CI, 90.00–95.84%) and 91.30% (95% CI, 88.10–94.51%). The AUROC was 0.91 [Fig. 1]. An example of the output from the Medios AI algorithm

along with the respective image with a diagnosis of any DR is shown in Fig. 3. The sensitivity for the diagnosis of VTDR was 100% [Fig. 4].

The number of false positives was eight (AI had labeled eight cases of No DR as having the disease). Six out of these eight images were found to have artifacts that had perhaps been misidentified by the algorithm.

Discussion

This study aimed to evaluate the performance of Medios AI. It demonstrated that the AI algorithm has very high sensitivity and specificity to detect RDR, any DR, as well as VTDR compared with manual, adjudicated grading by fellowship-trained vitreoretinal surgeons. The algorithm being integrated with the image acquisition and storage application of an existing commercially available smartphone-based imaging device, i.e., the Remidio NM FOP 10, made it user-friendly. Seamless integration of the AI with the NM FOP 10 manufacturers' application made the image workflow simple and time-efficient so that the reports could be produced in real-time by the technician using the device.

DR screening with teleophthalmology is widely practiced in India and across the world with the use of reading centers. Limited access to trained readers and interobserver variability associated with human grading led to the development of automated systems for DR. Advances in deep learning has led
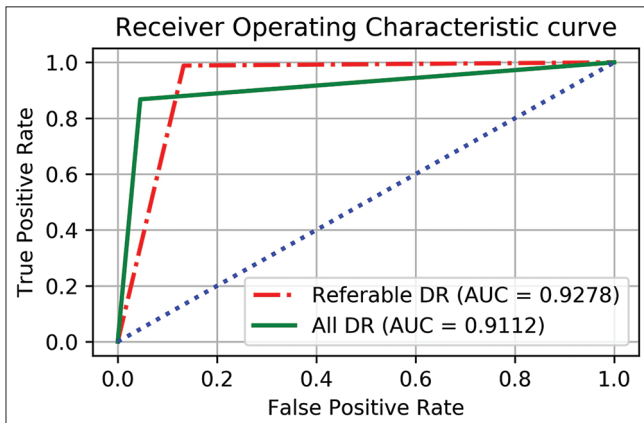

Figure 1: Area under the receiver operating curve for referable diabetic retinopathy and any diabetic retinopathy
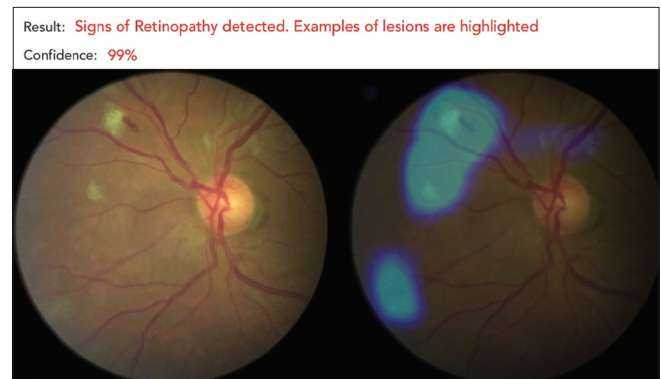

Figure 2: Example of the output from the Medios AI algorithm along with the respective image with a diagnosis of referable diabetic retinopathy
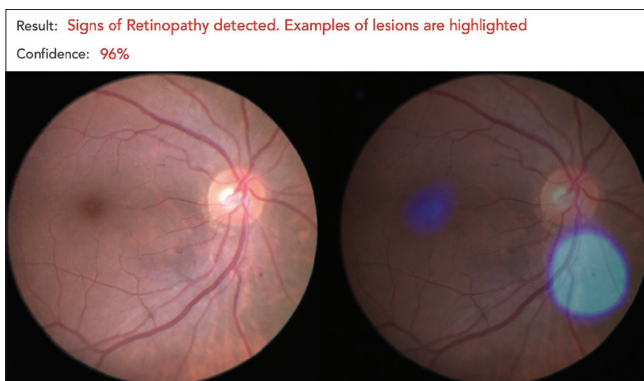

Figure 3: Example of the output from the Medios AI algorithm along with the respective image with a diagnosis of any diabetic retinopathy

Figure 4: Example of the output from the Medios AI algorithm along with the respective image with a diagnosis of proliferative diabetic retinopathy

to the development of several algorithms like the Google AI, Eye Nuk, and IDx-DR for the detection of DR. The CNN-based software designed by Google was used on images from patients presenting for DR screening in three tertiary eye hospitals in India, and it showed a high sensitivity and specificity (>90%) for detecting DR.[11] In a prospective study, the sensitivity and specificity of the Google AI for RDR at the first study site was 88.9% (95% CI, 85.8–91.5) and 92.2% (95% CI, 90.3–93.8); and 92.1% (95% CI, 90.1–93.8) and 95.2% (95% CI, 94.2–96.1) at second study site.[12]

Using another algorithm based on deep machine learning on the publicly available fundus image datasets, Abramoff et al. also found a high sensitivity (97%) and specificity (87%), similar to our results.[13] In a large pivotal study conducted by Abramoff et al., the sensitivity and specificity of the IDx-DR system in identifying RDR met the United States Food and Drug Administration (FDA) cut-offs for superiority. This study made IDx-DR the first FDA approved AI algorithm for the diagnosis of RDR [Table 1].[14]

Tufail et al. studied the performance of three different automated image analysis software and reported a sensitivity and specificity (>90%) using the EyeArt (Eyenuk Inc., Woodland Hills, CA) and Retmarker (Coimbra, Portugal) software.[15] The Retmarker software has also been used to detect DR in Indian eyes.[16] Roy et al. analyzed 5780 eyes of 1445 patients through the Retmarker software and found a high sensitivity (>90%) and low specificity (11–61%). Walton et al. published outcomes on the sensitivity and specificity of another algorithm, the Intelligent Retinal Imaging System (IRIS) for DR screening. In this retrospective study, the sensitivity and specificity of IRIS was 66.4% and 72.8%.[17]

The EyeArt algorithm of Eyenuk has been evaluated in several studies. Rajalakshmi et al. evaluated the EyeArt software for the detection of RDR [Table 1] using images captured with the Remidio FOP.[18] Bhaskaranand et al. published results using the EyeArt software, with mydriatic and nonmydriatic images from 1,01,710 eyes [Table 1].[19] The outcomes from mydriatic imaging were marginally better with improved sensitivity and greater AUROC. The Remidio FOP has nonmydriatic image acquisition capabilities and it will be interesting to see how the AI algorithm performs on nonmydriatic images going forward.

All the software programs that are cloud-based require high computational power and above all, Internet connectivity, for real-time reporting of results. The Medios AI, in contrast works offline, without Internet (or electricity). To our knowledge, this is one of the first few studies analyzing the accuracy of an offline AI-based software for DR screening.

A recent study by Natarajan et al., evaluated the performance of the Medios AI using images captured from the Remidio FOP, in 231 individuals with diabetes during a DR screening program [Table 1]. A community health care worker captured two-segment retinal images. An additional sensitivity analysis was performed to assess the AI's performance using images that are likely to be captured during large screening camps or high patient workflow by the unskilled workforce. Both good quality images and images that did not meet the minimum quality standards by the AI were included in this analysis. In the sensitivity analysis, the sensitivity of the AI for RDR remained unchanged at 100%, while the specificity dropped

Table 1: Performance of AI algorithms in the detection of referable diabetic retinopathy

	AI software	Sensitivity	Specificity
FDA cut-off for superiority		85%	82.5%
Present Study	Medios AI	98.8%	86.7%
Natarajan et al.[19]	Medios AI	100%	88.4%
Rajalakshmi et al.[17]	EyeArt	99.3%	68.8%
Bhaskaranand M et al.[18]	EyeArt	91%	91%
IDx[13]	IDx-DR	87.2%	90.7%

AI=Artificial intelligence, FDA=United States Food and Drug Administration

from 88.4% to 81.9%. There were a larger number of false positives outputs (attributed to poor image quality). While this may lead to increased referrals, patient safety was not compromised as the AI detected all individuals with RDR. This gives an idea of the real practical use of AI for DR screening.[20]

The IDX-DR is currently the only FDA approved algorithm for DR screening. While several studies have been conducted to evaluate various algorithms, the only ones currently in use for DR screening commercially in the US are IDx-DR, and EyeArt in the EU; neither are available in India. While we acknowledge that the comparison of different algorithms based on published results has its limitations (because of differences in study methods), we summarize the performance of the deep learning cloud-based algorithms currently used for DR screening in the USA and EU, and the offline Medios AI [Table 1].

In their recent paper on the current state of teleophthalmology in the United States, Rathi et al. describe applications of teleophthalmology in many diseases, including DR.[7] Authors highlight the upcoming role of automated DR screening using various algorithms to ease the burden of manual DR screening. Authors also state that given the increasing prevalence of DR, the emergence of automated screening serves as a promising tool to address this public health issue. In addition, we believe that it is extremely important to make upcoming AI-based algorithms offline for widespread adoption. An AI-based algorithm that gives consistent results with high accuracy may overcome human barriers like inter-grader variability, in addition to its ability to process millions of images quickly, maybe the best way forward for grading DR in the future. In India, and the developing world with limited resources, where access to the Internet and continuous electricity is a challenge in smaller towns and villages, these technologies can ensure that DR screening proceeds uninterrupted.

The advantages of this study are the use of three-field photography, grading, and adjudication by vitreoretinal surgeons as the gold standard. The drawbacks include a small sample size and the use of only mydriatic images. Larger studies, studies with nonmydriatic imaging will address if these results can be generalized. Studies evaluating the integration of the AI into the clinical workflow, and comparison with the clinical diagnosis from a comprehensive eye exam and real-world studies will provide more insight and understanding of this technology. We acknowledge that the algorithm is only trained to detect DR and currently works only integrated with the Remidio FOP camera. Additional work to focus on grading of DR, and detect other retinal disorders is required.

Conclusion

In conclusion, our preliminary results show that the novel AI algorithm has high sensitivity and specificity in detecting RDR as well as VTDR. This is probably the only software available in an offline mode that can deliver results instantly in real-time. If used on a larger scale, it has the potential of ensuring timely referrals. The results of the SMART Study (Simple Mobile-Based Artificial Intelligence Algorithm in the diagnosis of Diabetic Retinopathy) with a sample size of 900 patients using nonmydriatic images to test the robustness of the Medios algorithm is awaited. Multiple larger studies that show reproducibility and consistency can help validate the algorithm further. We conclude that our findings are encouraging, further work remains to improve the clinical validity of these algorithms.

Conflicts of interest

Medios Technologies only provided the AI for use for the study and played no role in the study design, funding, study process, data analysis, results or publication. Authors SB, SAR are related to one of the two co-founders of Medios Technologies, Singapore. We have no personal financial interests in the form of stocks and have received no remuneration or consultation fee from Medios Technologies, Singapore. The other authors have no conflict of interest to declare. No funding was received for this study.

References

1. Gadkari SS, Maskati QB, Nayak BK. Prevalence of diabetic retinopathy in India: The All India Ophthalmological Society diabetic retinopathy eye screening study 2014. Indian J Ophthalmol 2016;64:38-44.

2. Raman R, Rani PK, Reddi Rachepalle S, Gnanamoorthy P, Uthra S, Kumaramanickavel G, *et al.* Prevalence of diabetic retinopathy in India: Sankara Nethralaya diabetic retinopathy epidemiology and molecular genetics study report 2. Ophthalmology 2009;116:311-8.

3. 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;3:146-53.

4. Bastawrous A, Rono HK, Livingstone IA, Weiss HA, Jordan S, Kuper H, *et al.* Development and validation of a smartphone-based visual acuity test (Peek Acuity) for clinical practice and community-based fieldwork. JAMA Ophthalmol 2015;133:930-7.

5. Rajalakshmi R, Arulmalar S, Usha M, Prathiba V, Kareemuddin KS, Anjana RM, *et al.* Validation of smartphone based retinal photography for diabetic retinopathy screening. PLoS One 2015;10:e0138285.

6. Sharma A, Subramaniam SD, Ramachandran KI, Lakshmikanthan C, Krishna S, Sundaramoorthy SK. Smartphone-based fundus camera device (MII Ret Cam) and technique with ability to image peripheral retina. Eur J Ophthalmol 2016;26:142-4.

7. Rathi S, Tsui E, Mehta N, Zahid S, Schuman JS. The current state of teleophthalmology in the United States. Ophthalmology 2017;124:1729-34.

8. Kapoor R, Walters SP, Al-Aswad LA. The current state of artificial intelligence in ophthalmology. Surv Ophthalmol 2019;64:233-40.

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

10. Krause J, Gulshan V, Rahimy E, Karth P, Widner K, Corrado G, *et al.* Grader variability and the importance of reference standards for evaluating machine learning models for diabetic retinopathy. Ophthalmology 2018;125:1264-72.

11. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, *et al.* Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 2016;316:2402-10.

12. Gulshan V, Rajan R, Widner K, Wu D, Wubbels P, Rhodes T, *et al.* Performance of a deep-learning algorithm vs manual grading for detecting diabetic retinopathy in India. JAMA Ophthalmol 2019;137:987.

13. Abràmoff MD, Lou Y, Erginay A, Clarida W, Amelon R, Folk JC, *et al.* Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. Invest Ophthalmol Vis Sci 2016;57:5200-6.

14. 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 Digit Med 2018;1:39.

15. Tufail A, Kapetanakis VV, Salas-Vega S, Egan C, Rudisill C, Owen CG, *et al.* An observational study to assess if automated diabetic retinopathy image assessment software can replace one or more steps of manual imaging grading and to determine their cost-effectiveness. Health Technol Assess 2016;20:1-72.

16. Roy R, Lobo A, Lob A, Pal BP, Oliveira CM, Raman R, *et al.* Automated diabetic retinopathy imaging in Indian eyes: A pilot study. Indian J Ophthalmol 2014;62:1121-4.

17. Walton OB, Garoon RB, Weng CY, Gross J, Young AK, Camero KA, *et al.* Evaluation of automated teleretinal screening program for diabetic retinopathy. JAMA Ophthalmol 2016;134:204-9.

18. Rajalakshmi R, Subashini R, Anjana RM, Mohan V. Automated diabetic retinopathy detection in smartphone-based fundus photography using artificial intelligence. Eye 2018;32:1138-44.

19. Bhaskaranand M, Ramachandra C, Bhat S, Cuadros J, Nittala M, Sadda S, *et al.* The value of automated diabetic retinopathy screening with the Eye Art system: A study of more than 100,000 consecutive encounters from people with diabetes. Diabetes Technol Ther 2019;21:635-43.

20. Natarajan S, Jain A, Krishnan R, Rogye A, Sivaprasad S. Diagnostic accuracy of community-based diabetic retinopathy screening with an offline artificial intelligence system on a smartphone. JAMA Ophthalmol 2019. doi: 10.1001/JAMAOphthalmol. 2019.2923.