Effectiveness of a computer-assisted Algorithm for onsite screening of diabetic retinopathy from retinal photographs at diabetic outpatient clinics

Sheila John[1], Sangeetha Srinivasan[2], Keerthi Ram[3], Mohanasankar Sivaprakasam[3]

[1]*Sankara Nethralaya, Chennai, India*

[2]*Vision Research Foundation, Chennai, India*

[3] *Healthcare Technology Innovation Centre, IIT Madras, Chennai, India*

Purpose: To examine the effectiveness of a computer-assisted device (CAD) for onsite screening of diabetic retinopathy (DR) at diabetic outpatient clinics.

Methods: 1263 patients were examined over two and half years. The undilated fundus photographs were obtained from the clinic. An ophthalmologist and optometrist independently assessed the photographs. The assessment was done in a darkened room following a masked fashion and processed through CAD. Diabetic Retinopathy was defined according to the International Clinical Diabetic Retinopathy Disease Severity Scale. The severity of diabetic retinopathy was assessed based on the scale.

Results: 2526 eyes of 1263 patients were assessed. The algorithm successfully graded 2153 (85%) images with 63.04% sensitivity and 79.63% specificity in comparison to an ophthalmologist. The sensitivity and specificity were 60.87% and 79.05% respectively in comparison to an optometrist, The agreement between ophthalmologist and optometrist was kappa=0.835 for the presence of DR, 0.835 for severity of DR.

Conclusion: This algorithm may be utilized in a diabetic clinic for a quick screening with only the retinal photographs.

Keywords

algorithm, diabetes, retinopathy, sensitivity, specificity.

1. Introduction

The World Health Organization predicts that there will be a 150% increase in the diabetic population over the next decade; the diabetic epidemic presents considerable challenges to health care delivery in India. [1] Diabetic retinopathy (DR) is a significant cause of ocular morbidity. However, due to the dearth of resources in India. In type 2 diabetic patients, macular oedema [2] occurs most frequently. It is about 7.5 % and is the most common cause of moderate visual impairment amongst the working age adult group. However, less than one-third report for DR screening due to challenges in transportation. [3] Community based screening programs and innovative telemedicine strategies may improve compliance. [4-6] Telemedicine is cost effective and increasingly seen as a tool to reach the underserved population in the developing countries. It is effective for screening diabetic retinopathy in the rural areas. [5] In this study, we propose to utilize the established paradigms in diabetic retinopathy screening in the urban diabetic clinics based in Chennai.

An annual screening by physicians/general practitioners/diabetologists is recommended to prevent microvascular complications of diabetes. [4] This requires follow up with patients, a good network and understanding with ophthalmologists and physicians. Normally, those with diabetes do not visit an ophthalmologist unless they have visual loss. However, annual eye screening is imperative to avoid visual loss as per the guidelines of the International Diabetes Federation. In addition, patient compliance to regular eye exams is compromised due to multifactorial barriers such as lack of

ISBN: 978-91-89081-09-3
DOI: https://doi.org/10.15626/ishimr.2020.03

ophthalmologist(s), physical disability of patient, lack of awareness of visual loss and blindness due diabetic retinopathy and cost involved. [6] Nevertheless, tele screening [7,8] in combination with automated DR detection is a viable screening option that can overcome many of these barriers. It can also help screen a larger population. We previously reported the accuracy of a computer-assisted automated system in detecting DR from single-field fundus photographs acquired at the tele-screening site. [9] In this report, we examine the effectiveness of a computer-assisted device (CAD) for DR screening from retinal photographs. Patients with type 2 diabetes were screened for diabetic retinopathy at diabetic outpatient clinics by an ophthalmologist and optometrist using CAD.

2. Methods

The institutional ethics committee approved the study, and the research adhered to the tenets of the Declaration of Helsinki. Patients provided written informed consent before participation in the study.

Our DR image screening system was developed by Healthcare Technology Innovation Centre based on images from publicly available retinal image datasets (about 2000 images) from around the world and retrospective assessment of DR images from Sankara Nethralaya. The Algorithm assessed a sensitivity of 83% and specificity of 80% respectively.

Two diabetes clinics were chosen where the facility for such a screening by an ophthalmologist is lacking. The selection of diabetes clinics for the study was deemed suitable since physicians are usually the first point of contact for a person with diabetes mellitus.

Patients with already known or newly diagnosed type 2 diabetes who have not had fundus examination as reported by the patient or confirmed by available medical records were included. The fundus images were captured at the physician's or diabetologists facility/clinic at the time of a regular diabetic check-up.

1263 patients were examined from Jan 2015 to May 2017. The paramedical staff-fundus photographer and optometrists in diabetic clinics were trained to take fundus photographs without dilation using fundus cameras.

2.1 Fundus photography

The FORUS 3nethra Classic Fundus Camera (Forus Health Pvt Ltd, Bangalore, India) was utilized. Physiological mydriasis was attempted by having the patients wait in a darkened room for 5-10 minutes. Trained paramedical staff took single or multiple undilated 45˚field retinal photograph centred on the macula of each fundus. On each occasion, the right eye was photographed first and up to 3 minutes was allowed between each photograph to allow redilation of the left eye. Photography was repeated for images of suboptimal quality due to patient blinking, alignment, or poor fixation.

All images in a given session for each patient were uploaded to the web-based Telemedicine platform bundled with 3nethra fundus camera using Broadband connectivity at the diabetic clinic.

(Figure 1)

Figure 1 Image assessing platform for the graders

Ophthalmologist at the base hospital performed evaluation and grading of photographs. The reports generated were printed and handed over to the patient. The patient reports were assessed and sent to the ophthalmologists for further treatment.

The retinal photographs were stored as JPEG images and viewed in a darkened room on cathode ray tube screen. All retinal photographs were de-identified and coded with an identification number and uploaded to a secure database. The same retinal specialist, senior optometrist and algorithm at the base hospital using liquid crystal display computer monitor with the same resolution of 1280 x 800 reviewed all digital fundus images, nonmydriatic from the diabetic clinic.

Diagnostic criteria for diabetic retinopathy was defined according to the Proposed International Clinical Diabetic Retinopathy and diabetic macular oedema disease severity scales and severity of diabetic retinopathy was assessed accordingly. [10]

The grading of fundal features by the human graders were recorded using the same protocol and was based on retinal features alone, masked to other clinical information. Referable DR was defined as presence of DR in any one field of the fundus photograph for each eye separately. Incidental fundus photograph findings other than diabetic retinopathy were also documented. Fundus images were assessed whether they were gradable or not.

2.2 Data management and statistical analysis

Data were prepared in MS-Excel 2013 with de-identified patient's identification number. Information regarding age, sex, and duration of diabetes was recorded.
The design of DR analyser's software as a data-driven system provides specific task-related metrics for evaluation. Performance compared to human expert drives the algorithm refinement process.

Module evaluation: The lesion-level performance of DR analysers software detectors can be evaluated by comparing algorithm outputs C_2 against lesion annotations provided by clinicians. Two methods of evaluation are used:

- FROC analysis (TPR vs FPPI): for lesion detectors, and
- ROC analysis (TPR vs FPR): for normal anatomy detectors and DR referral analytics module (Figure 2)

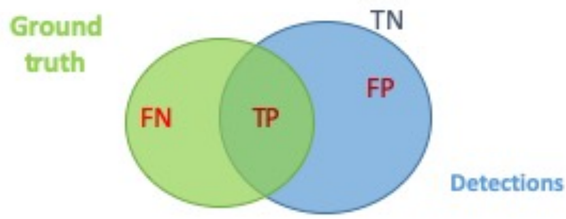

Figure 2 shows Module evaluation

Metrics used: AUC (area under ROC curve), sensitivity, specificity, precision, accuracy, confusion matrix

Lesion detection*:* computed per image

Patient demographics and clinical measures of the eye were summarized for the sample with descriptive statistics. Continuous variables were presented as mean + standard deviation or median with interquartile range (IQR) and categorical variables were presented as proportions. The algorithm processed the images fed into it using MATLAB software (MathWorks Inc., Natick, MA, USA) and provided numerical outputs for image gradeability (image gradeability score) and presence of DR (DR score). The image was considered gradable if the image gradeability score was >0.1 and DR was considered to be present if the DR score was >0.55. The cut offs were considered reasonable based on the beta testing during development and pilot testing before undertaking the study. The higher score reflected increased confidence in gradeability and presence of DR.

We determined disease status through nonmydriatic fundus examination by a retina specialist at the vitreoretinal outpatient service who will be/served as the reference standard. Wherever possible, we determine the presence of vision-threatening diabetic retinopathy (VTDR) defined as presence of diabetic macular edema and PDR as determined by the reference standard.

We estimated the primary outcome; the sensitivity of the algorithm to detect diabetic retinopathy from the diabetic clinic compared to the reference standard and included 95% Wilson confidence intervals (CI). The specificity, positive and negative predictive value (accuracy and precision respectively) was also estimated. The areas under the receiver operating characteristic curves were determined. For fundus imaging at the diabetic clinic, we calculated inter-observer agreement for the primary outcome (presence or absence of DR) as well as a secondary outcome (image gradeability) using a kappa statistic. All data was entered in Excel sheets (Microsoft Excel, Version 2010) and all statistical analysis were performed using STATA version 12.1, I/C (STATA Corp, Fort worth, Texas, USA). A p-values less than 0.05 was considered statistically significant.

2.3 Results

Demographics and DR status: We enrolled 2526 eyes of 1263 patients to test the accuracy of the algorithm to detect DR in a diabetic clinic setting. The mean age of participants was 54.5±10.6 years (median=55 years, IQR=46 – 62 years, range=34 – 83) and 66% were men. The mean duration of diabetes in this cohort was 8.5±7.3 years (median=7 years, IQR=3-12 years, range=0.5 – 30 years).

Algorithm Descriptive: The mean image gradeability score was 0.20±0.10 (median=0.20, IQR= 0.14 – 0.27). The gradeability score for gradable images was 0.23±0.08 in eyes with gradable images compared to 0.05+0.03 for those with ungradable images (p<0.001, Wilcoxon test).

The overall DR score was 0.29±0.24 (median=0.27, IQR=0.08-0.45). Eyes with DR had a mean score of 0.70±0.12 and those without DR had a DR score of 0.22±0.17 (p<0.001, Wilcoxon test).

2.3.1 Algorithm vs. Ophthalmologist

Table 1 shows the proportion of images that were gradable and ungradable by the algorithm and the ophthalmologist.

Table 1 Gradeability between the algorithm and the ophthalmologist

		Ophthalmologist		
		Ungradable	Gradable	Total
Algorithm	Ungradable	4	369	373 (15%)
	Gradable	24	2129	2153 (85%)
	Total	28	2498	2526 (100%)

Compared to the ophthalmologist (reference standard), 15% images were ungradable by the algorithm. Overall, the ophthalmologist found only 28 images to be ungradable compared to 373 images by the algorithm. There was only slight agreement in terms of image gradeability between the ophthalmologist and algorithm, Kappa= -0.001 (95%Ci= -0.019 - 0.018).

Table 2 shows the sensitivity and specificity of the algorithm in comparison to the ophthalmologist (reference standard).

Table 2 Sensitivity and Specificity of the algorithm to detect DR compared to Ophthalmologist.

Algorithm	**Ophthalmologist**				**Value**	**95% CI**
	Absent	Present	Total	Sens	63.04%	55.63%-70.03%
Absent	1841	69	1910 (76%)	Spec	79.63%	77.93%-81.25%
Present	471	117	588 (24%)	PPV	19.76%	16.61%-23.22%
Total	2312	186	2498 (100%)	NPV	96.44%	95.51%-97.22%

Sens, sensitivity; spec, specificity; PPV, positive predictive value; NPV, negative predictive value

The area under the receiver operator curve was 0.57 (95%CI=0.55 to 0.59). The internal validity of the algorithm was tested and is given below (Table 3).

2.3.2 Algorithm vs. Optometrist:

Table 3 shows the proportion of images that were gradable and ungradable by the algorithm and the optometrist.

Table 3 Gradeability of images by algorithm and optometrist.

		Optometrist		
		Ungradable	Gradable	Total
Algorithm	Ungradable	5	368	373 (15%)
	Gradable	21	2132	2153 (85%)
	Total	26	2500	2526 (100%)

Compared to the optometrist, 15% images were ungradable by the algorithm. Overall, the ophthalmologist found only 26 images to be ungradable compared to 373 images by the algorithm. There was only a slight agreement in terms of image gradeability between the ophthalmologist and algorithm, Kappa= 0.006 (95%CI = -0.014 - 0.026).

Tables 4 shows the validity in terms of sensitivity and specificity of the algorithm in comparison to that of the optometrist.

Table 4 Validity of the algorithm compared to that of the optometrist

Algorithm	Optometrist				Value	95% CI
	Absent	Present	Total	Sens	60.87%	52.88%-68.45%
Absent	1849	63	1912 (76%)	Spec	79.05%	77.34%-80.68%
Present	490	98	588 (24%)	PPV	16.67%	13.74%-19.93%
Total	2339	161	2500 (100%)	NPV	96.71%	95.80%-97.46%

Sens, sensitivity; spec, specificity; PPV, positive predictive value; NPV, negative predictive value

The area under the receiver operator curve was 0.56 (95%CI=0.55 to 0.58).

2.3.3 Optometrist vs. Ophthalmologist

Table 5 Agreement between optometrist and ophthalmologist in DR grading

Optometrist	Ophthalmologist			Variable				Value	95% CI
	Absent	Present	Total		Kappa	95%CI	Sens	81.52%	75.15% - 86.85%
Absent	2296	38	2334 (93.4%)	Gradeability	0.701	0.561-0.840	Spec	99.52%	99.15% - 99.76%
Present	11	153	164 (6.6%)	Presence of DR (Yes/No)	0.835	0.813-0.854	PPV	93.17%	88.10% - 96.54%
Total	2307	191	2498 (100%)	VTDR	0.835	0.812-0.843	NPV	98.54%	97.97% - 98.99%

Sens, sensitivity; spec, specificity; PPV, positive predictive value; NPV, negative predictive value

The sensitivity of the optometrist to detect DR was found to be 82% and specificity was 99% in detecting DR compared to ophthalmologist. The areas under the receiver operator curve was 0.95 (95%CI=0.93 to 0.97).

3. Discussion

In this study, we aimed to assess the performance of a computer-assisted device in screening for DR at diabetic clinic settings. We chose a diabetic clinic for the study because physicians are usually the first point of contact for a person with diabetes mellitus.

We observed that the algorithm demonstrated a sensitivity of 63% and specificity of 79% in the automated detection of DR, when compared to a reference standard (ophthalmologist grading). We observed that the algorithm has an acceptable specificity but at the expense of a lower sensitivity. One explanation could be that there were relatively lesser proportion of patient images with clinical DR. A higher specificity reduces the number of false-positive results that implies that unnecessary referrals are reduced. [11] High sensitivity issue concerns patient safety and high specificity concerns efficiency in a screening program.

In this study from retinal photographs, we identified cases of myelinated nerve fiber, choroidal sclerosis, asteroid hyalosis and chorioretinal scar at the macular area, drusen and other artefacts, which the algorithm identified as DR. (Figures 3-5)

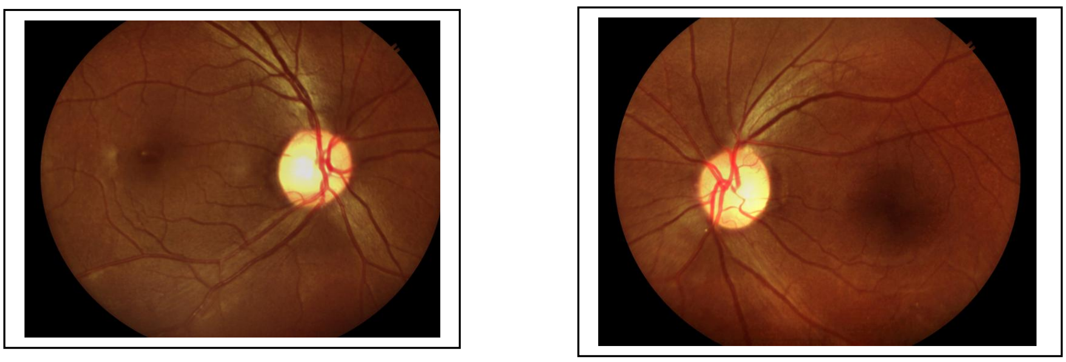

Figure 3. Diabetic Clinic – Algorithm diagnosed as absence of DR in diabetic Patient

Figure 4. Diabetic Clinic - Medullated nerve fibre right eye - Algorithm diagnosed as

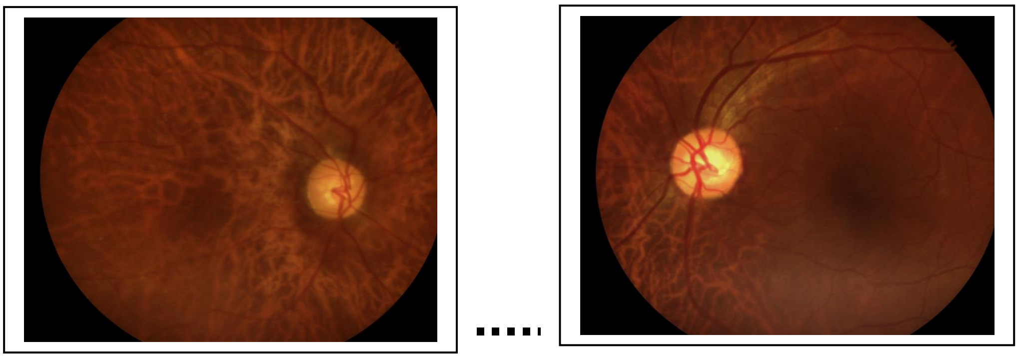

Figure 5. Diabetic Clinic - Choroidal sclerosis - Algorithm misdiagnosed as presence of DR

This could be also an explanation for the algorithm overestimating the presence of DR. This may be arguably considered a breakthrough. [12] The reason being that the algorithm is primarily trained only in the identification of DR. Nevertheless, it can detect other abnormalities such as drusen (although in our case, it identifies it as DR). This is advantageous given that other abnormal retinal findings must be thoroughly checked, by an ophthalmologist to rule out other retinal conditions.

Some patients had cataract, which attributed to the sub-optimal image gradeability. One explanation for this is that we utilized physiological dilatation, as against pharmacological dilatation. [13] [14] Indian eyes are known to have smaller pupils in scotopic conditions, thus limiting the image quality and thereby compromising the software's assessment capabilities. This may have been a contributing factor to some of the ungradable images. [15]

Gulshan et al [10] developed and validated an algorithm using 128,175 retinal images and reported a high sensitivity and specificity (both >93%) in identifying referable RDR. However, they utilized higher-end cameras. Nevertheless, we employed a locally made low-cost tabletop camera, which primarily served to obtain a large number of images at the diabetic clinic setting and is also suitable to screen large masses in the Indian population. In India, more than 2/3rds of the population reside in rural areas with one ophthalmologist available for every 100,000 people. [17] This is also encouraging because trained paramedical staff can represent a viable task force for screening DR [18] at remote places where it may be difficult to have access to an ophthalmologist or for an ophthalmologist to be physically present at the screening centre.

The strength of the algorithm was the accuracy in identifying a vast majority of those with vision-threatening retinopathy, which require urgent referral. Screening for diabetic retinopathy takes about 15 minutes in a diabetic clinic with physiological dilation, retinal fundus imaging and evaluation by computer-assisted algorithm. Annual dilated examination can be implemented for every diabetic patient at a reduced cost. This can be provided in a diabetic clinic and will eliminate a separate visit to the ophthalmologist. In addition, with the increasing prevalence of diabetes, the emergence of automated screening serves as a promising tool to address this public health issue especially in a country like India.

References

[1] Schmittdiel J A, Gopalan A, Lin M W, Banerjee S, Chau C V, Adams A S. *Population health management for diabetes: health care system-level approaches for improving quality and addressing disparities*. Curr Diab Rep. 2017; 17:31. doi:10.1007/s11892-017-0858-3.

[2] Corcóstegui B, Durán S, González-Albarrán M O, Hernández C, Ruiz-Moreno J M, Salvador J, et al. *Update on diagnosis and treatment of diabetic retinopathy: a consensus guideline of the working group of ocular health (Spanish Society of Diabetes and Spanish Vitreous and Retina Society)*. J Ophthalmol. 2017;2017. doi: 10.1155/2017/8234186.

[3] Raman R, Gella L, Srinivasan S, Sharma T. *Diabetic retinopathy: An epidemic at home and around the world*. Indian J Ophthalmol. 2016; 64: 69-75. doi: 10.4103/0301-4738.178150.

[4] Lee D J, Kumar N, Feuer W J, Chou C-F, Rosa P R, Schiffman J C, et al. *Dilated eye examination screening guideline compliance among patients with diabetes without a diabetic retinopathy diagnosis: the role of geographic access*. BMJ Open Diabetes Res Care [Internet]. 2014; 2(1): e000031. doi: 10.1136/bmjdrc-2014-000031.

[5] Combi C, Pozzani G, Pozzi G. *Telemedicine for developing countries. A survey and some design issues*. Appl Clin Inform. 2016;7:1025–50. doi:10.4338/ACI-2016-06-R-0089.

[6] Bresnick G H, Mukamel D B, Dickinson J C, Cole D R. *A screening approach to the surveillance of patients with diabetes for the presence of vision-threatening retinopathy*. Ophthalmology. 2000; 107: 19–24. doi:10.1016/s0161-6420(99)00010-x.

[7] John S, Sengupta S, Reddy S J, Prabhu P, Kirubanandan K, Badrinath S S. *The Sankara Nethralaya mobile teleophthalmology model for comprehensive eye care delivery in rural India*. Telemed J E Health [Internet]. 2012; 18: 382–387. Available from: http://online.liebertpub.com/doi/abs/10.1089/tmj.2011.0190

[8] Zimmer-Galler I E, Kimura A E, Gupta S. *Diabetic retinopathy screening and the use of telemedicine*. Curr Opin Ophthalmol. 2015; 26: 167–172.

[9] John S, Ram K, Sivaprakasam M, Raman R. *Assessment of computer-assisted screening technology for diabetic retinopathy screening in India - preliminary results and recommendations from a pilot study*. Stud Health Technol Inform. 2016; 231: 74-81.

[10] Wilkinson C P, Ferris F L 3rd, Klein R E, Lee P P, Agardh C D, Davis M, Dills D, Kampik A, Pararajasegaram R, Verdaguer J T; Global Diabetic Retinopathy Project Group. *Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales*. Ophthalmology. 2003 Sep; 110(9): 1677-1682. doi: 10.1016/S0161-6420(03)00475-5.

[11] Abràmoff M D, Folk J C, Han D P, Walker J D, Williams D F, Russell S R, Massin P, Cochener B, Gain P, Tang L, Lamard M, Moga D C, Quellec G, Niemeijer M. *Automated analysis of retinal images for detection of referable diabetic retinopathy*. JAMA Ophthalmol. 2013; 131: 351-7. doi: 10.1001/jamaophthalmol.2013.1743.

[12] Roach L. Artificial intelligence. EyeNet Magazine. 2017: 77-83.

[13] Horton MB, Brady C J, Cavallerano J, Abramoff M, Barker G, Blodi B, Chiang M, Crockett C, Garg S, Karth P, Liu Y, Maloney K, Newman C, Rathi S, Rosenson A, Sheth V, Silva P S, Stebbins K, Zimmer-Galler. Practice Guidelines for Ocular Telehealth-Diabetic Retinopathy. 2019 ,3rd edition , February 19, 2019.

[14] Murgatroyd H, Ellingford A, Cox A, et al. *Effect of mydriasis and different field strategies on digital image screening of diabetic eye disease*. Br J Ophthalmol. 2004; 88: 920-924. doi:10.1136/bjo.2003.026385.

[15] Scanlon P H, Foy C, Malhotra R, Aldington S J. *The influence of age, duration of diabetes, cataract, and pupil size on image quality in digital photographic retinal screening*. Diabetes Care. 2005; 28: 2448-2453. doi:10.2337/diacare.28.10.2448.

[16] 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: 2402-2410. doi:10.1001/jama.2016.17216.

[17] Lundquist B M, Sharma N, Kewalramani K. *Patient perceptions of eye disease and treatment in Bihar, India*. J Clinic Experiment Ophthalmol. 2012; 3:2. doi:10.4172/2155-9570.1000213.

[18] 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 Ophthalmology. 2019; 137: 1182-1188. doi: 10.1001/jamaophthalmol.2019.2923