



Accuracy of ultrawide-field fundus ophthalmoscopy-assisted deep learning for detecting treatment-naïve proliferative diabetic retinopathy

Toshihiko Nagasawa · Hitoshi Tabuchi · Hiroki Masumoto · Hiroki Enno · Masanori Niki · Zaigen Ohara · Yuki Yoshizumi · Hideharu Ohsugi · Yoshinori Mitamura

Received: 28 June 2018 / Accepted: 14 January 2019 / Published online: 23 February 2019
© The Author(s) 2019

Abstract

Purpose We investigated using ultrawide-field fundus images with a deep convolutional neural network (DCNN), which is a machine learning technology, to detect treatment-naïve proliferative diabetic retinopathy (PDR).

Methods We conducted training with the DCNN using 378 photographic images (132 PDR and 246 non-PDR) and constructed a deep learning model. The area under the curve (AUC), sensitivity, and specificity were examined.

Result The constructed deep learning model demonstrated a high sensitivity of 94.7% and a high specificity of 97.2%, with an AUC of 0.969.

Conclusion Our findings suggested that PDR could be diagnosed using wide-angle camera images and deep learning.

Keywords Ultrawide-field fundus ophthalmoscopy · Proliferative diabetic retinopathy · Deep learning · Deep convolutional neural network

Introduction

According to a World Health Organization report, the number of diabetic patients worldwide has increased from 108 million in 1980 to 422 million in 2014, and the prevalence of diabetes worldwide in adults (> 18 years of age) has increased from 4.7% in 1980 to 8.5% in 2014 [1]. Voigt et al. reported that 25.8% of diabetic patients have complications of retinopathy (nonproliferative 20.2%; proliferative 4.7%; unclassified 0.7%; blindness 0.1%) [2]. Early treatment, compared to deferral of photocoagulation, was associated with a small reduction in the incidence of severe visual loss [3]; however, undergoing fundus examination by an ophthalmologist is unrealistic and costly for diabetic patients. Furthermore, there is a large cost burden associated with diabetic retinopathy, and the financial impact may be even more severe for many patients with this complication who live in developing countries [4].

Recently, image processing technology using a deep learning application, which is a machine learning algorithm, has attracted attention because of its accuracy. Using this technology for medical imaging is being actively studied [5–7]. In fact, image

T. Nagasawa (✉) · H. Tabuchi · H. Masumoto · Z. Ohara · Y. Yoshizumi · H. Ohsugi
Department of Ophthalmology, Saneikai Tsukazaki Hospital, 68-1 Aboshi Waku, Himeji City, Hyogo Prefecture 671-1227, Japan
e-mail: t.nagasawa@tsukazaki-eye.net

H. Enno
Rist Inc., Tokyo, Japan

M. Niki · Y. Mitamura
Department of Ophthalmology, Institute of Biomedical Sciences, Tokushima University Graduate School, Tokushima, Japan

diagnosis has already been reported in ophthalmology [8–11]. In addition, the advent of wide-angle fundus cameras, such as the ultrawide-field scanning laser ophthalmoscope (Optos 200Tx; Optos plc, Dunfermline, UK) known as Optos, has made it possible to simply and noninvasively capture a wide range of fundus photographs [12–14]. In the present study, we assessed and determined the accuracy of ultrawide-field fundus images with deep learning to detect the presence of treatment-naïve proliferative diabetic retinopathy (PDR).

Methods

Dataset

The procedures in the present study conformed to the tenets of the Declaration of Helsinki. Informed consent was obtained from the subjects after they understood the study's nature and possible consequences.

The study dataset was comprised of 132 images and data from patients with treatment-naïve PDR; those without fundus diseases were extracted from April 1, 2011, to March 30, 2018, at the clinical database of the ophthalmology departments of Saneikai Tsukazaki Hospital and Tokushima University Hospital. These images were reviewed by three retinal specialists to assess the presence of PDR using mydriatic slit-lamp binocular indirect ophthalmoscopy and were registered in an analytical database. All patients underwent Optos and ultrawide-field fluorescein angiography (FA) (Fig. 1). The levels of diabetic retinopathy were defined from the retinal images using the Early Treatment Diabetic Retinopathy Study (ETDRS) severity scale [3]. Out of 378 fundus images, 132

images were from PDR patients, and 246 images were from normal subjects without PDR.

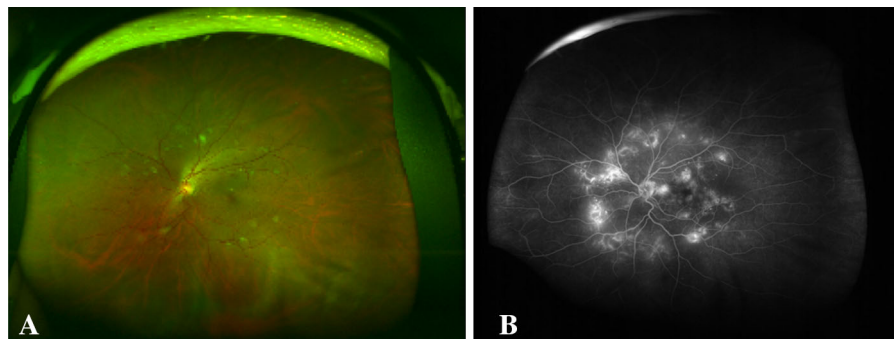
In this study, we used K-fold cross-validation [15]. Briefly, the image data were divided into K groups; K-1 groups were used as training data, and one group was used as validation data. This process was repeated until each group became a validation dataset. In the present study, we divided the data into nine groups. The images of the training dataset were augmented by adjusting for brightness, gamma correction, histogram equalization, noise addition, and inversion; augmenting the images increased the amount of learning data 18 times. The deep convolutional neural network (DCNN) model, as detailed below, was created and trained with preprocessed image data.

Deep learning model and training the model

We implemented a deep learning model that used a VGG-16 DCNN (Fig. 2), which is a type of DCNN that automatically learns the images' local features and generates a classification model [16–18]. The aspect ratio of the original Optos images was 3900×3072 pixels; for analysis, we resized the aspect ratio of all the input images to 256×192 pixels. The RGB image input had a range of 0–255; therefore, we normalized it to the range of 0–1 by dividing it by 255.

The VGG-16 comprised five blocks and three fully connected layers. Each block comprised some convolutional layers, followed by a max-pooling layer to decrease position sensitivity and improve generic recognition [19]. After flattening the output of block 5, there were two fully connected layers; the first removed the spatial information from the extracted features, and the last was a classification layer that used feature vectors of the target images acquired from

Fig. 1 A representative fundus image obtained by ultrawide-field scanning laser ophthalmoscopy. The presence of proliferative diabetic retinopathy (PDR) is seen on ultrawide-field fundus color image (A) and on fluorescein angiography (B)



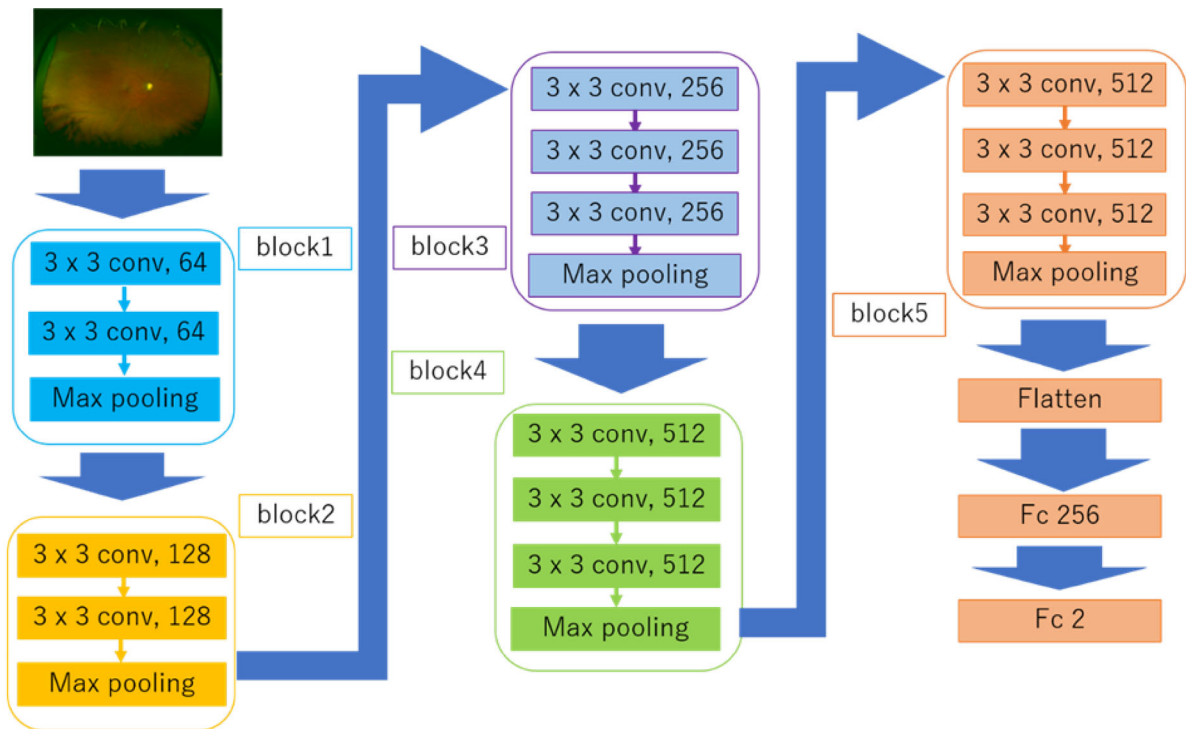


Fig. 2 Overall architecture of the VGG-16 model. The deep convolutional neural network (DCNN) used ImageNet parameters: the weights of blocks 1–4 are fixed. Block 5 and the fully connected layers were adjusted

the previous layers, together with the softmax function for binary classification. To improve the generalization performance, we conducted dropout processing; therefore, masking was performed with a probability of 25% for the first fully connected layer. Fine-tuning was used to increase the learning speed and optimize performance even with less data [20, 21]. We used parameters from ImageNet: blocks 1–4 were fixed, and block 5 and the fully connected layers were trained. The weights of block 5 and the fully connected layer that we were training were updated using the optimization momentum SGD algorithm (learning coefficient = 0.001, inertial term = 0.9), which is a stochastic gradient descent method [22, 23]. Of the 40 deep learning models obtained from 40 learning cycles, the one with the highest correct answer rate for the test data was selected as the DL model. To build and evaluate the model, Keras (<https://keras.io/ja/>) was run on TensorFlow (<https://www.tensorflow.org/>), which was written in Python.

Outcome

Receiver operating characteristic (ROC) curves were created based on the deep learning models' abilities to discriminate between PDR and non-DR images. These curves were evaluated using area under the curve (AUC), sensitivity, and specificity.

Statistical analysis

Student's *t* test was used to compare age, whereas Fisher's exact test was used to determine the ratios of men to women and right to left eye images. The 95% confidence intervals (CIs) of the AUCs were obtained, as follows. Images that were judged to exceed a threshold were defined as positive for PDR, and an ROC curve was created. We created nine models and nine ROC curves. For AUC, a 95% CI was obtained by assuming a normal distribution, using the means and standard deviations of the nine ROC curves. For sensitivity and specificity, we used the optimal cutoff values, which were the points at which both sensitivity and specificity were 100% in each ROC curve [24].

The ROC curve was calculated using scikit-learn, and the CIs for sensitivity and specificity were determined using SciPy. The other statistical analyses were performed using SPSS version 22 software (IBM, Armonk, New York, USA). A two-sided *P* value of < 0.05 was considered statistically significant.

Data availability

The Optos image datasets analyzed during the present study are available from the corresponding author upon request.

Heatmap creation

As illustrated in the heatmap, the DCNN focuses on which coordinate axes on the image were classified (Fig. 3). The heatmap was generated using gradient-weighted class activation mapping (Grad-CAM) [25]; a gradient layer using the first convolution layer of block 3 was designated. ReLU was specified as the backprop modifier.

Results

In total, 132 PDR images from 94 patients (mean age 55.3 ± 12.5 years; 90 men and 42 women; 69 left fundus images and 63 right fundus images) and 246 non-DR images from 199 patients (mean age 55.2 ± 13.9 years; 161 men and 85 women; 127 left fundus images and 119 right fundus images) were

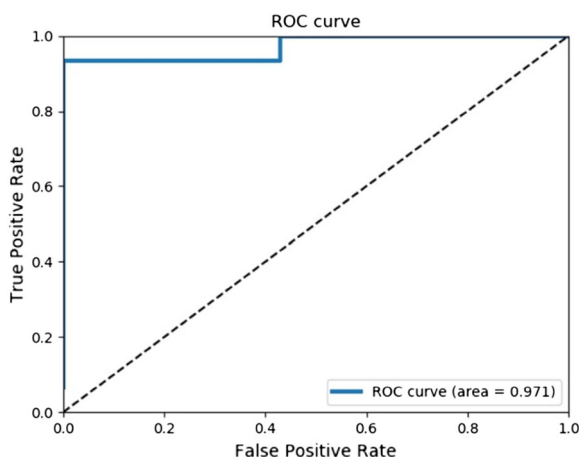


Fig. 3 Representative ROC curve of the deep learning model

analyzed. No significant differences were detected between these two groups in terms of age, sex, and left–right eye image ratio (Table 1).

Performance of the DCNN

For PDR diagnosis, the deep learning model had a sensitivity of 94.7% (95% CI 90.6–96.9%), specificity of 97.2% (95% CI 92.4–99.2%), and area under the curve (AUC) of 0.969 (95% CI 0.935–0.971) (Fig. 4).

Discussion

In this study, we investigated the deep learning method's efficacy in identifying referable treatment-naïve PDR based on 132 fundus photographs. The deep learning algorithm showed high sensitivity of 94.7%, high specificity of 97.2%, and AUC of 0.969 for the detection of treatment-naïve PDR. We focused on treatment-naïve PDR only, because it can need immediate treatment. Even when the diagnosis was based on color photographs only, the results were comparable to those made based on color fundus images and FA assessment by retinal specialists.

In the past, deep learning was examined at all stages of diabetic retinopathy, with good results were obtained [8, 26–29]; however, the fundus camera of the rear pole was used. In this study, we used a wide-angle ocular fundus camera, because diabetic retinopathy is an important disease that can affect both the posterior region and periphery of the retina. The ETDRS 7 [30] defined lesions predominantly around the standard field as predominantly peripheral lesions (PPLs). The extent of these PPLs is associated with retinopathy progression [31, 32]. Therefore, the type of camera used is important.

The drawback of the present study was that we did not examine diabetic maculopathy, which causes vision disturbances and can also be diagnosed using deep learning, as reported by Gulshan et al. [8]. The algorithms' ability to detect vision-threatening diabetic retinopathy is important to evaluate; the software's sensitivity is especially important to determine. Originally, deep learning required tens of thousands of data to investigate the presence or absence of a diagnosis; however, the number of treatment-naïve PDR cases had been limited. Therefore, further studies

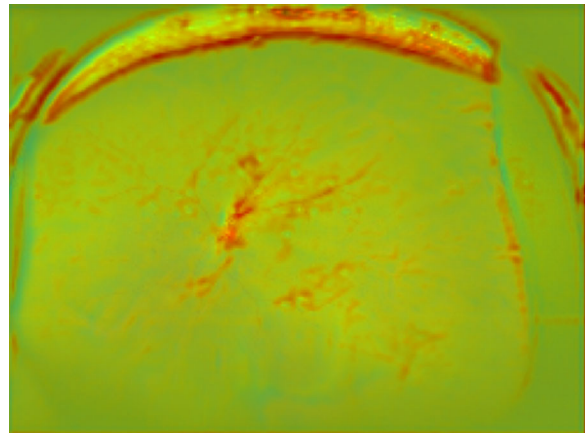
Table 1 Patient demographics

	PDR images	Normal images	<i>p</i> values	
Number of images (patients)	132 (94)	246 (199)		
Age	55.3 ± 12.5	55.2 ± 13.9	0.902	Student's
Sex (female)	42 (31.8%)	85 (34.6%)	0.648	Fisher's ex
Eye (left)	69 (52.3%)	127 (51.6%)	0.914	Fisher's ex

Data are presented as numbers (%), unless otherwise indicated

No statistically significant differences were observed between the groups

Fig. 4 Heatmap superimposed on the photograph. The red color represents the areas of deep neural network concentrations



are needed to assess whether diabetic retinopathy can be staged appropriately.

At present, making a final diagnosis based on images alone might not be accurate and prudent. We believe that image diagnosis should only be used to confirm a doctor's diagnosis. However, in developing countries where the number of physicians may also be limited, remote image diagnosis may be especially useful. Kanjee et al. reported remote diagnosis was cost-effective, noting further reduction in medical expenses when automatic diagnosis was available [33]. To address the most urgent medical problems in the world in an efficient, timely, and cost-effective manner, all available resources are needed. Therefore, introducing artificial intelligence in the medical field is timely, welcomed, and needed.

Although combining DCNN and Optos images can provide better results, it is not particularly superior to medical examination. Personal and actual examinations by ophthalmologists remain indispensable for definite diagnosis. Furthermore, both conventional angiography and optical coherence tomography

angiography, performed by a retinal specialist, are essential to confirm a qualitative diagnosis, assess treatment effects, and provide follow-up observations.

Conclusion

PDR could be diagnosed using an approach that involves wide-angle camera images and deep learning.

Acknowledgements We thank Masayuki Miki and the orthoptists at Tsukazaki Hospital for their support in data collection.

Compliance with ethical standards

Conflicts of interest The authors declare that they have no competing financial interests.

Human and animal rights Approval to perform this study was obtained from the institutional review boards of Saneikai Tsukazaki Hospital and Tokushima University Hospital.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits

unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- World Health Organization (2017) Media centre diabetes fact sheet. <http://www.who.int/mediacentre/factsheets/fs312/en/>. Accessed 30 Oct 2018
- Voigt M, Schmidt S, Lehmann T, Köhler B, Kloos C, Voigt U, Meller D, Wolf G, Müller UA, Müller N (2017) Prevalence and progression rate of diabetic retinopathy in type 2 diabetes patients in correlation with the duration of diabetes. *Exp Clin Endocrinol Diabetes* 126:570–576
- Early Treatment Diabetic Retinopathy Study Research Group (1991) Early photocoagulation for diabetic retinopathy. ETDRS report number 9. *Ophthalmology* 98:766–785
- Guariguata L, Whiting DR, Hambleton I, Beagley J, Linnenkamp U, Shaw JE (2014) Global estimates of diabetes prevalence for 2013 and projections for 2035. *Diabetes Res Clin Pract* 103:137–149
- LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521:436–444
- Liu S, Cai W, Che H, Pujol S, Kikinis R, Feng D, Fulham MJ (2015) Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer’s disease. *IEEE Trans Biomed Eng* 62:1132–1140
- Litjens G, Sánchez CI, Timofeeva N, Hermesen M, Nagtegaal I, Kovacs I, Hulsbergen-van de Kaa C, Bult P, van Ginneken B, van der Laak J (2016) Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Sci Rep* 6:26286
- Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nelson PC, Mega JL, Webster DR (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 316:2402–2410
- Asaoka R, Murata H, Iwase A, Araie M (2016) Detecting preperimetric glaucoma with standard automated perimetry using a deep learning classifier. *Ophthalmology* 123:1974–1980
- Fang L, Cunefare D, Wang C, Guymer RH, Li S, Farsiu S (2017) Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search. *Biomed Opt Express* 8:2732–2744
- Ohsugi H, Tabuchi H, Enno H, Ishitobi N (2017) Accuracy of deep learning, a machine-learning technology, using ultra-wide-field fundus ophthalmoscopy for detecting rhegmatogenous retinal detachment. *Sci Rep* 7:9425
- Prasad PS, Oliver SC, Coffee RE, Hubschman JP, Schwartz SD (2010) Ultra wide-field angiographic characteristics of branch retinal and hemicentral retinal vein occlusion. *Ophthalmology* 117:780–784
- Wessel MM, Aaker GD, Parlitsis G, Cho M, D’Amico DJ, Kiss S (2012) Ultra-wide-field angiography improves the detection and classification of diabetic retinopathy. *Retina* 32:785–791
- Ogura S, Yasukawa T, Kato A, Usui H, Hirano Y, Yoshida M, Ogura Y (2014) Wide-field fundus autofluorescence imaging to evaluate retinal function in patients with retinitis pigmentosa. *Am J Ophthalmol* 158:1093–1098
- Kohavi R (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proc Int Joint Conf Artif Intell* 2:1137–1145
- Deng J, Dong W, Socher R (2009) Imagenet: a large-scale hierarchical image database. *Comput Vis Pattern Recognit* 9:248–255
- Lee CY, Xie S, Gallagher P, Zhang Z, Tu Z (2015) Deeply-supervised nets. *AISTATS* 2:562–570
- Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, Berg AC, Fei-Fei L (2015) Imagenet large scale visual recognition challenge. *Int J Comput Vis* 115:211–252
- Scherer D, Müller A, Behnke S (2010) Evaluation of pooling operations in convolutional architectures for object recognition. In: *Proceedings of international conference on Artificial neural networks—ICANN*, pp 92–101
- Agrawal P, Girshick R, Malik J (2014) Analyzing the performance of multilayer neural networks for object recognition. In: *Proceedings of ECCV*, pp 329–344
- Redmon J, Divvala S, Girshick R, Farhadi F (2015) You only look once: unified real-time object detection. [arXiv:1506.02640](https://arxiv.org/abs/1506.02640)
- Nesterov Y (1983) A method for unconstrained convex minimization problem with the rate of convergence $O(1/k^2)$. *Doklady AN USSR* 269:543–547
- Qian N (1999) On the momentum term in gradient descent learning algorithms. *Neural Netw* 12:145–151
- Akobeng AK (2007) Understanding diagnostic tests 3: receiver operating characteristic curves. *Acta Paediatr* 96:644–647
- Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D (2016) Grad-CAM: visual explanations from deep networks via gradient-based localization. [arXiv:1610.02391v3](https://arxiv.org/abs/1610.02391v3)
- Abràmoff MD, Lou Y, Erginay A, Clarida W, Amelon R, Folk JC, Niemeijer M (2016) Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. *Invest Ophthalmol Vis Sci* 57:5200–5206
- Gargeya R, Leng T (2017) Automated identification of diabetic retinopathy using deep learning. *Ophthalmology* 124:962–969
- Takahashi H, Tampo H, Arai Y, Inoue Y, Kawashima H (2017) Applying artificial intelligence to disease staging: deep learning for improved staging of diabetic retinopathy. *PLoS ONE* 12:e0179790
- Ting DS, Cheung CY, Lim G, Tan GS, Quang ND, Gan A, Hamzah H, Garcia-Franco R, San Yeo IY, Lee SY, Wong EY, Sabanayagam C, Baskaran M, Ibrahim F, Tan NC, Finkelstein EA, Lamoureux EL, Wong IY, Bressler NM, Sivaprasad S, Varma R, Jonas JB, He MG, Cheng CY, Cheung GCM, Aung T, Hsu W, Lee ML, Wong TY (2017) Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal

- images from multiethnic populations with diabetes. *JAMA* 318:2211–2223
30. Early Treatment Diabetic Retinopathy Study Research Group (1991) Grading diabetic retinopathy from stereoscopic color fundus photographs—an extension of the modified Airlie House classification. ETDRS report number 10. *Ophthalmology* 98:786–806
 31. Silva PS, Cavallerano JD, Haddad NM, Kwak H, Dyer KH, Omar AF, Shikari H, Aiello LM, Sun JK, Aiello LP (2015) Peripheral lesions identified on ultrawide field imaging predict increased risk of diabetic retinopathy progression over 4 years. *Ophthalmology* 122:949–956
 32. Silva PS, Dela Cruz AJ, Ledesma MG, van Hemert J, Radwan A, Cavallerano JD, Aiello LM, Sun JK, Aiello LP (2015) Diabetic retinopathy severity and peripheral lesions are associated with nonperfusion on ultrawide field angiography. *Ophthalmology* 122:2465–2472
 33. Kanjee R, Dookeran RI, Mathen MK, Stockl FA, Leicht R (2016) Six-year prevalence and incidence of diabetic retinopathy and cost-effectiveness of tele-ophthalmology in Manitoba. *Can J Ophthalmol* 51:467–470