

Predicting 10-2 Visual Field From Optical Coherence Tomography in Glaucoma Using Deep Learning Corrected With 24-2/30-2 Visual Field

Yohei Hashimoto¹, Taichi Kiwaki², Hiroki Sugiura², Shotaro Asano¹, Hiroshi Murata^{1,10}, Yuri Fujino^{1,11}, Masato Matsuura¹, Atsuya Miki³, Kazuhiko Mori⁴, Yoko Ikeda^{4,5}, Takashi Kanamoto⁶, Junkichi Yamagami⁷, Kenji Inoue⁸, Masaki Tanito⁹, Kenji Yamanishi², and Ryo Asaoka^{1,12-15}

¹ Department of Ophthalmology, The University of Tokyo, Tokyo, Japan

² Graduate School of Information Science and Technology, The University of Tokyo, Tokyo, Japan

³ Department of Ophthalmology, Osaka University Graduate School of Medicine, Osaka, Japan

⁴ Department of Ophthalmology, Kyoto Prefectural University of Medicine, Kyoto, Japan

⁵ Oike-Ganka Ikeda Clinic, Kyoto, Japan

⁶ Hiroshima Memorial Hospital, Hiroshima, Japan

⁷ Department of Ophthalmology, JR General Hospital, Tokyo, Japan

⁸ Inouye Eye Hospital, Tokyo, Japan

⁹ Department of Ophthalmology, Shimane University Faculty of Medicine, Shimane, Japan

¹⁰ Department of Ophthalmology, National Center for Global Health and Medicine, Tokyo, Japan

¹¹ Department of Ophthalmology, Shimane University Faculty of Medicine, Izumo, Japan

¹² Department of Ophthalmology, Seirei Hamamatsu General Hospital, Shizuoka, Japan

¹³ Seirei Christopher University, Shizuoka, Japan

¹⁴ Nanovision Research Division, Research Institute of Electronics, Shizuoka University, Shizuoka, Japan

¹⁵ The Graduate School for the Creation of New Photonics Industries, Shizuoka, Japan

Correspondence: Ryo Asaoka, Department of Ophthalmology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan. e-mail: ryoasa0120@mac.com

Received: July 27, 2021

Accepted: September 28, 2021

Published: November 23, 2021

Keywords: deep learning; glaucoma; optical coherence tomography; visual field

Citation: Hashimoto Y, Kiwaki T, Sugiura H, Asano S, Murata H, Fujino Y, Matsuura M, Miki A, Mori K, Ikeda Y, Kanamoto T, Yamagami J, Inoue K, Tanito M, Yamanishi K, Asaoka R. Predicting 10-2 visual field from optical coherence tomography in glaucoma using deep learning corrected with 24-2/30-2 visual field. *Transl Vis Sci Technol.* 2021;10(13):28. <https://doi.org/10.1167/tvst.10.13.28>

Purpose: To investigate whether a correction based on a Humphrey field analyzer (HFA) 24-2/30-2 visual field (VF) can improve the prediction performance of a deep learning model to predict the HFA 10-2 VF test from macular optical coherence tomography (OCT) measurements.

Methods: This is a multicenter, cross-sectional study. The training dataset comprised 493 eyes of 285 subjects (407, open-angle glaucoma [OAG]; 86, normative) who underwent HFA 10-2 testing and macular OCT. The independent testing dataset comprised 104 OAG eyes of 82 subjects who had undergone HFA 10-2 test, HFA 24-2/30-2 test, and macular OCT. A convolutional neural network (CNN) DL model was trained to predict threshold sensitivity (TH) values in HFA 10-2 from retinal thickness measured by macular OCT. The predicted TH values was modified by pattern-based regularization (PBR) and corrected with HFA 24-2/30-2. Absolute error (AE) of mean TH values and mean absolute error (MAE) of TH values were compared between the CNN-PBR alone model and the CNN-PBR corrected with HFA 24-2/30-2.

Results: AE of mean TH values was lower in the CNN-PBR with HFA 24-2/30-2 correction than in the CNN-PBR alone (1.9dB vs. 2.6dB; $P = 0.006$). MAE of TH values was lower in the CNN-PBR with correction compared to the CNN-PBR alone (4.2dB vs. 5.3 dB; $P < 0.001$). The inferior temporal quadrant showed lower prediction errors compared with other quadrants.

Conclusions: The performance of a DL model to predict 10-2 VF from macular OCT was improved by the correction with HFA 24-2/30-2.

Translational Relevance: This model can reduce the burden of additional HFA 10-2 by making the best use of routinely performed HFA 24-2/30-2 and macular OCT.

Introduction

Glaucoma is characterized by progressive visual field (VF) damage and is the leading cause of irreversible blindness in the world.¹ Glaucomatous VF deterioration is accompanied by structural changes such as ganglion cell death and loss of axons.²⁻⁴ Structural damage can be detected by optical coherence tomography (OCT),⁵ and researchers have developed models to discriminate between glaucomatous and healthy OCT-imaged eyes using machine learning algorithms such as support vector machine (SVM),⁶ random forests,⁷ and deep learning (DL).^{8,9} Recent research has also demonstrated the potential for DL models to predict VF sensitivity from OCT images in patients with glaucoma.¹⁰⁻¹⁴

Patients suspected of having glaucoma are almost always tested with static automated perimetry using test points that are spaced 6° apart such as the Humphrey field analyzer (HFA) 24-2/30-2 test.^{15,16} This is because glaucomatous VF damage usually starts as Bjerrum scotoma or nasal step, which often appear within the central 30°. Recent studies have revealed that the measurement of the central 10° VF is, however, essential for the accurate assessment of glaucomatous damage.¹⁸ In particular, vision-related quality of life is more closely associated with the 10° VF than with the 24° or 30° VF. The problem is that it is costly and time consuming to perform 10° VF test in addition to a 24° or 30° VF test.¹⁹ This emphasizes the importance of OCT imaging, because it has the potential to reduce the number of VF measurements necessary to accurately monitor disease progression.

We recently reported two models to predict visual sensitivities of the HFA 10-2 test in a pointwise manner from spectral domain OCT (SD-OCT) in glaucoma patients: (1) a DL model using a convolutional neural network (CNN) with pattern-based regularization (PBR) [CNN-PBR],¹¹ showing good predictive performance of absolute error [AE] of the whole VF, 2.7 dB and pointwise mean absolute error [MAE], 5.5 dB; and (2) a DL model with correction by HFA 24-2/30-2 test results of the same eye, improving MAE from between 9.4 and 9.5 dB to 5.4 dB.²⁰ In the current study, we combined these two models and investigated whether it was beneficial to correct CNN-PBR-predicted 10-2 visual sensitivities using HFA 24-2/30-2 test results of the same eye.

Methods

This study was approved by the Research Ethics Committee of the Graduate School of Medicine and

Faculty of Medicine at the University of Tokyo, Inouye Eye Hospital, Kyoto Prefectural University of Medicine, Oike-Ganka Ikeda Clinic, JR Tokyo general hospital, Hiroshima Memorial Hospital, Osaka University Graduate School of Medicine, and Shimane University Faculty of Medicine. Informed consent for storing their data in the hospital database for the research purposes was obtained from all patients. This study was performed according to the tenets of the Declaration of Helsinki.

Training Dataset

Generally, a large amount of paired data of VF and OCT are required for training, but a small amount of paired data were available in the current study, because VF testing was performed more often than OCT scan in our clinical setting. We used the paired and no-paired datasets in the CNN-PBR model by first learning the patterns from nonpaired data (VF data alone) using an unsupervised method and then regularize (i.e., PBR) the prediction by the CNN referring to the obtained patterns. The details have been described elsewhere.¹²

The training dataset comprised 493 eyes of 285 subjects (407 eyes with open angle glaucoma [OAG] and 86 normative eyes). Subjects had undergone HFA 10-2 VF testing and OCT imaging. A second training dataset included 7715 HFA 10-2 VF tests that were not paired with SD-OCT images, which came from patients with glaucoma other than the paired dataset. All subjects underwent complete ophthalmic examinations, including biomicroscopy, gonioscopy, intraocular pressure measurement, funduscopy, refraction, best-corrected visual acuity measurement, and axial length measurements. Patients were enrolled during the period between April 2013 and August 2016 at the University of Tokyo Hospital, Inoue Eye Hospital, JR Tokyo General Hospital, and Hiroshima Memorial Hospital.

OAG was defined as follows: (1) presence of typical glaucomatous changes such as a disc rim notch and a retinal nerve fiber layer defect identified by ophthalmoscopy or fundus photography; (2) gonioscopically wide open angles of grade 3 or 4 based on the Shaffer classification; (3) visual acuity ≥ 0.5 LogMAR; (4) refractive error $< +3.0$ diopter; and (5) age 20 to 80 years old. Patients with ocular diseases that could affect the results of SD-OCT examinations and VF testing, such as diabetic retinopathy or age-related macular degeneration, were carefully excluded. Eyes with clinically significant senile cataract were also excluded.

Normative eyes were defined as follows: (1) no abnormal findings except for clinically insignificant senile cataract; (2) no history of ocular diseases that

could affect the results of SD-OCT and VF examinations; (3) normal VF test results according to the Anderson-Patella criteria; (4) refractive error $< +3.0$ diopter; and (5) age 20 to 80 years,

Testing Dataset

The testing dataset was independent of the training dataset. It comprised 104 OAG eyes of 82 subjects. Inclusion and exclusion criteria and VF and SD-OCT measurements were identical to those in the training dataset. In addition to HFA 10-2 test and OCT data, all eyes in the testing dataset had also undergone HFA 24-2/30-2 VF testing. The HFA 24-2/30-2 test data were used to correct the predicted threshold sensitivity (TH) values using CNN-PBR.

VF Testing

VF testing was performed with HFA 10-2 test using the Swedish Interactive Thresholding Algorithm standard strategy within three months from the measurement of SD-OCT. Near-refractive correction was used as necessary. All subjects had previously experienced the HFA test at least once. We excluded unreliable VFs with fixation loss $\geq 20\%$ or false-positive responses $\geq 15\%$, following the recommendation by the manufacturer. HFA 24-2/30-2 tests in the testing dataset were performed within three months of the HFA 10-2 test.

SD-OCT Measurement

RS 3000 (Nidek Co Ltd, Aichi, Japan) and OA-2000 (Tomey, Aichi, Japan) were used to obtain SD-OCT and axial length measurement data, respectively. All SD-OCT measurements were performed after pupil dilation with 1% tropicamide. We excluded data with apparent eye movement and involuntary blinking or saccade during the measurement, and imaging data with quality factor < 7 , as recommended by the manufacturer. A 9.0×9.0 mm image was centered on the fovea. The macular thicknesses of the three parts were exported as pixel images (512×128): macular retinal nerve fiber layer (RNFL), ganglion cell layer + inner plexiform layer, and outer segment + retinal pigment epithelium. We resized the images in each part to 224×224 pixels with a bicubic interpolation²¹ over 4×4 neighborhood and resampling using pixel area relation²² to inherit the parameters from ResNet, which is one of the most popular pretrained models for image classification.²³ Furthermore, data augmentation was performed via vertical flip. The full details are described in our previous study.¹²

Deep Learning (CNN-PBR)

We trained the CNN-PBR model to predict threshold (TH) values of the HFA 10-2 test from the three retinal thicknesses, using the parameters of ResNet. The details of CNN-PBR are described in our previous reports.^{11,12} In short, CNN-PBR has an advantage of making best use of paired (VF with OCT) data over CNN without PBR, by avoiding the problem of overfitting using VF data alone.

Correction With HFA 24-2/30-2 Test Results

To try to improve prediction performance, we further corrected the predicted TH values by using HFA 24-2/30-2 VF test results of the same eye, as detailed in Figure 1. First, the TH values of the 10-2 test at coordinates (X,Y): (3, 9), (9, 3), (3, -9), (9, -3), (-3, -9), (-9, -3), (-9, 3), and (-3, 9) were predicted using neighboring TH values, which were predicted with CNN-PBR, and the weights based on the distance to the test points. The predicted TH values of the HFA 10-2 test were averaged in each quadrant (superior nasal, inferior nasal, superior temporal, and inferior temporal). Second, the actual TH values of the HFA 24-2 test at $(\pm 3, \pm 3)$, $(\pm 3, \pm 9)$, $(\pm 9, \pm 3)$ were averaged in each quadrant. Finally, the difference between these averages was added to the predicted TH values of the 10-2 test in each quadrant.

Statistical Analysis

We compared the prediction performance between the two models: (1) CNN-PBR alone (CNN-PBR_{alone}) and (2) CNN-PBR with correction (CNN-PBR_{correction}). First, we compared absolute error (AE) of mean TH (mTH) values for the whole VF between the two models, using a linear mixed effects model whereby patients were regarded as the random effects. Second, we compared pointwise prediction performance through mean absolute error (MAE) between the two models, using a linear mixed effects model. MAE was calculated as:

$$MAE = \frac{\sum_{i=1}^{68} |\text{predicted visual sensitivity of the } i\text{th point} - \text{actual visual sensitivity of the } i\text{th point}|}{68}$$

where i = number of the 68 predicted test points.

We illustrated MAE at each testing point to investigate the trend of the prediction error caused by spacing position. Additionally, we illustrated the signed prediction error stratified by actual sensitivity of HFA 10-2 test using a boxplot. Statistical analysis was performed with Python (version 3.7.6; Python Software Foundation) and the statistical programming language

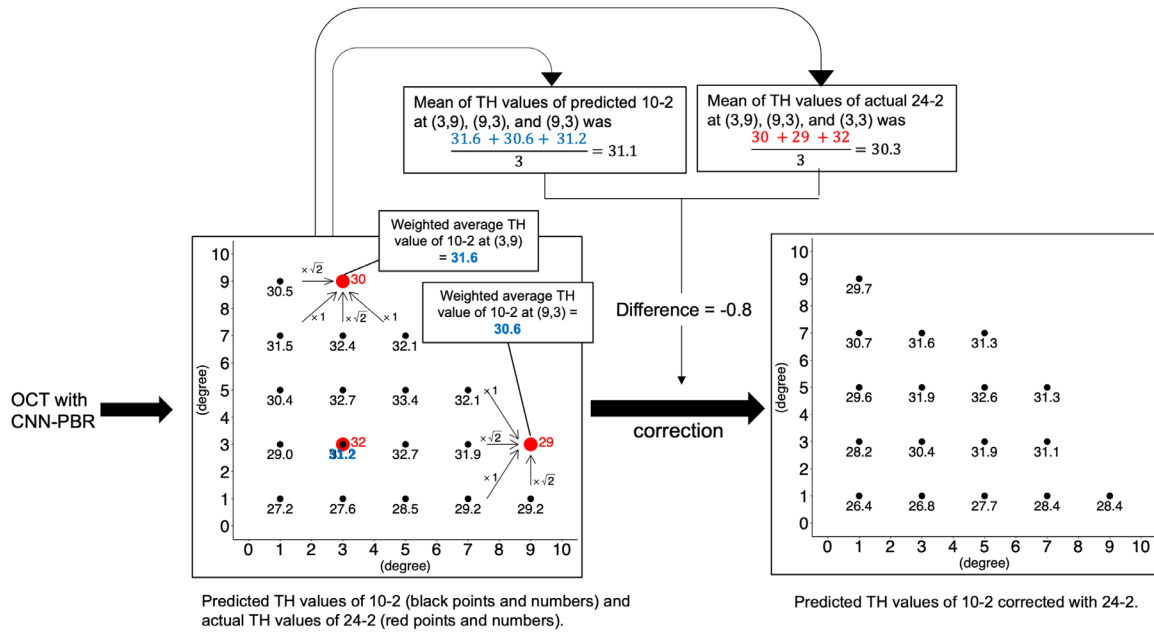


Figure 1. Outline of the HFA 24-2/30-2 correction method. The superior nasal VF in the left eye of a representative subject is shown. First, the TH values at (3,9) and (9,3) coordinates were predicted from the neighboring TH values, which were predicted with CNN-PBR, and the weights based on the distance to the test points. For example, distances between (1,9) and (3,9), between (1,7) and (3,9), between (3,7) and (3,9), and between (5,7) and (3,9) are 2, $2\sqrt{2}$, 2, $2\sqrt{2}$, and thus the weights become $\frac{1}{2} : \frac{1}{2\sqrt{2}} : \frac{1}{2} : \frac{1}{2\sqrt{2}} = \sqrt{2}:1:\sqrt{2} : 1$. Using these weights, the TH value of 10-2 at (3,9) was predicted as follows:

$$\frac{30.5 \times \sqrt{2} + 31.5 \times 1 + 32.4 \times \sqrt{2} + 32.1 \times 1}{\sqrt{2} + 1 + \sqrt{2} + 1} = 31.6$$

Likewise, the predicted TH value at (9,3) was 30.6. We averaged the TH values at (3,9), (9,3) and (3,3):

$$\frac{31.6 + 30.6 + 31.2}{3} = 31.1$$

Second, we averaged the TH value of 24-2 at (3,9), (9,3), and (9,9):

$$\frac{30 + 29 + 32}{3} = 30.3$$

Finally, the difference between these values ($30.3 - 31.1 = -0.8$) was added to the predicted TH values of 10-2. The corrections in the other quadrants were performed in a similar manner.

R language (version 3.6.3; R Foundation for Statistical Computing, Vienna, Austria).

Results

The Table shows the characteristics of the training and testing datasets. Actual TH values of the HFA 10-2 test are shown in Figure 2. The mTH value in the inferior temporal quadrant (26.3 dB) was higher than in the superior temporal (20.3 dB), superior nasal (17.8 dB), and inferior nasal quadrants (22.2 dB) (linear mixed model, all $P < 0.001$).

Significantly lower AE of mTH values were observed with CNN-PBR_{correction} than with the CNN-PBR_{alone} (1.9 dB vs. 2.6 dB; difference, -0.7 ; 95% confidence interval, -1.3 to -0.2 ; linear mixed model, $P = 0.006$) (Fig. 3). MAE of TH values was also significantly lower with CNN-PBR_{correction} compared to the CNN-PBR_{alone} model (4.2 dB vs. 5.3 dB; differ-

ence, -1.1 ; 95% confidence interval, -1.6 to -0.6 ; linear mixed model, $P < 0.001$) (Fig. 4). AE at each testing point are shown in Figure 5. In general, the AEs in the inferior temporal quadrant tended to be lower than those in other quadrants. Figure 6 shows the signed prediction error stratified by actual sensitivity of HFA 10-2. There was a trend toward more negative prediction error where actual sensitivity was high.

Discussion

In the present study, the HFA 10-2 test was predicted from SD-OCT imaging using a DL model (CNN-PBR) further corrected using HFA 24-2/30-2 test results of the same eye. Prediction performance was significantly improved using this correction method. Prediction errors were small: 1.9 dB (AE of mTH values) and 4.2 dB (MAE of TH values); we

Table. Characteristics of the Training and Testing Datasets

	Training	Testing
No. of subjects	285	82
Age (year)	53.7 ± 15.0	60.2 ± 12.1
Female	158 (55%)	38 (46%)
No. of eyes	493*	104†
Laterality (left)	241 (49%)	53 (51%)
Axial length (mm)	25.5 ± 2.0	25.5 ± 1.7‡
Mean threshold value of HFA 10-2 (dB)	24.5 ± 9.4	21.7 ± 8.5
Mean deviation of HFA 10-2 (dB)	-8.5 ± 9.5	-11.0 ± 8.9
Mean deviation of HFA 24-2/30-2 (dB)	NA	-10.9 ± 8.6
Macular RNFL (µm)	30.8 ± 9.2	27.7 ± 8.3
GCL + IPL (µm)	40.2 ± 9.0	36.8 ± 8.6
OS/RPE (µm)	67.3 ± 3.8	66.2 ± 3.6

GCL, ganglion cell layer; IPL, inner plexiform layer; OS, outer segment; RPE, retinal pigment epithelium.

Data are presented as n (%) or mean ± standard deviation.

* Consisted of 86 normative eyes and 407 eyes with open-angle glaucoma.

† All subjects were open-angle glaucoma (30 eyes, normal-tension glaucoma; 15 eyes, primary open glaucoma; 59 eyes, normal-tension glaucoma or primary open glaucoma was unknown).

‡ Axial length was unknown in two patients.

					14.8 (12.7)	15.8 (13.0)				
		15.6 (13.1)	15.5 (13.2)	14.9 (13.6)	15.8 (13.2)	16.6 (13.1)	16.2 (13.0)			
	16.4 (13.2)	16.7 (13.8)	17.3 (13.9)	16.7 (14.0)	16.7 (14.0)	17.0 (13.9)	17.1 (13.6)	16.4 (13.3)		
	19.9 (13.3)	21.4 (12.8)	21.6 (13.2)	19.3 (14.0)	18.5 (14.5)	17.5 (14.4)	17.4 (14.2)	16.5 (13.9)		
24.2 (10.6)	26.3 (9.9)	27.4 (9.7)	28.7 (9.0)	29.0 (7.7)	24.2 (11.4)	21.8 (12.7)	19.8 (13.2)	18.1 (13.9)	16.6 (13.5)	
27.1 (8.4)	28.1 (8.4)	29.5 (7.6)	30.0 (7.5)	29.4 (8.4)	28.8 (8.7)	26.6 (10.3)	23.0 (12.5)	20.6 (12.9)	19.4 (13.2)	
	27.7 (9.0)	28.1 (8.8)	28.5 (9.2)	28.6 (8.9)	27.0 (10.1)	24.9 (12.1)	22.0 (13.6)	20.0 (13.2)		
	24.9 (10.6)	26.2 (10.1)	26.1 (10.5)	25.2 (11.2)	23.6 (12.2)	22.5 (12.9)	20.8 (13.2)	18.9 (13.6)		
		22.5 (11.6)	22.7 (11.8)	22.0 (12.1)	20.8 (13.1)	20.3 (13.0)	18.9 (13.3)			
				20.6 (12.1)	18.8 (13.3)					

Figure 2. Actual threshold values of the 10-2 VF test (left eye). Mean (upper row) and standard deviation (lower row) values of all eyes at each test point are shown. Right eyes were mirror imaged.

believe these to be the smallest errors reported to date for this type of prediction model.

Many models have been reported to discriminate between glaucoma eyes and nonglaucoma eyes using OCT.^{6,7,24,25} Limited models have also been developed to predict VF measurements from OCT, but predicted measurements were usually mean VF sensitivity or sectoral VF sensitivity.^{10,13} In this research we

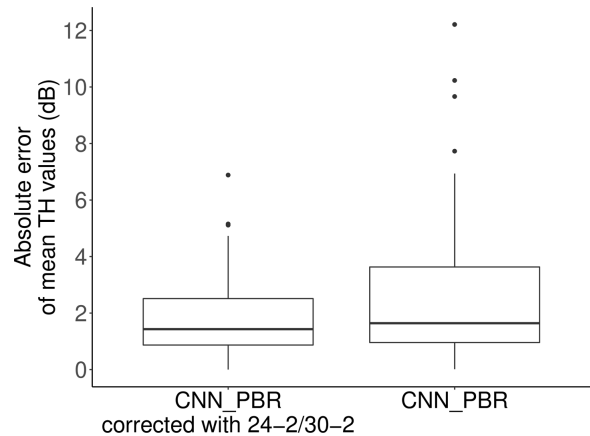


Figure 3. Absolute error of mean threshold values. The absolute error of CNN-PBR corrected with Humphrey field analyzer 24-2/30-2 test results was significantly lower than the same model without correction (1.9 dB vs. 2.6 dB; difference, -0.7; 95% confidence interval, -1.3 to -0.2; linear mixed model, $P = 0.006$).

predicted TH values in a pointwise manner; as shown in a recent paper, accurate pointwise predictions are more difficult than sectorial prediction.¹¹ The importance of pointwise prediction cannot be overstated when considering the application of such a model to real-world clinical settings. Given that test-retest VF variability lies between 1 and 2 dB in the central area and between 4 and 6 dB in the points at 27°,^{26,27} the current performance of pointwise prediction (MAE, 4.2 dB) within the central 10° area should be considered very good.

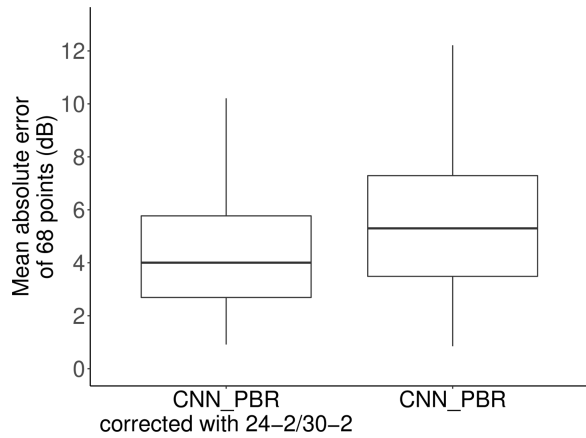


Figure 4. Mean absolute error of threshold values. The MAE of CNN-PBR corrected with Humphrey field analyzer 24-2/30-2 test results was significantly lower than the same model without correction (4.2 dB vs. 5.3 dB; difference, -1.1 ; 95% confidence interval, -1.6 to -0.6 ; linear mixed model, $P < 0.001$).

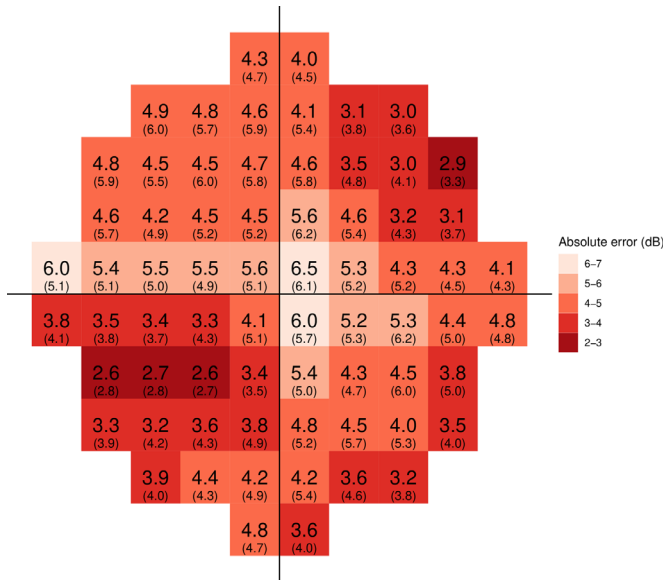


Figure 5. Pointwise absolute prediction error (left eye). Mean (*upper row*) and standard deviation (*lower row*) values of all eyes at each predicted 10-2 VF test point are shown. Right eyes were mirror imaged.

As widely acknowledged, the association between retinal structure and function is nonlinear.^{4,28} Structural damage precedes functional damage in glaucoma, that is, visual sensitivity has not deteriorated until RNFL thickness reaches a critical level.²⁹ DL models are helpful for the current prediction task whereby HFA 10-2 test sensitivity (response variable) and SD-OCT measurements (explanatory variable) are considered to be nonlinearly associated, because DL does not require any assumption, such as a linear relation between response and explanatory variables, which is

generally an assumption for conventional multivariable models.

In the current study, the AE in the inferior temporal area tended to be lower than the AE observed in other quadrants. This area corresponds to the preserved “central isle” of the VF seen in patients with advanced glaucoma,³⁰ and the TH values observed in this region remained relatively high. This is a possible explanation for the greater prediction performance in the inferior temporal area, because OCT is more useful for predicting VF in early-to-moderate glaucoma than in advanced glaucoma.^{29,31} Another possible reason may be the smaller standard deviation of actual TH values in the inferior temporal area. In addition, conventional VF testing, including the Swedish Interactive Thresholding Algorithm standard, decides VF sensitivity using the bracketing method. However, this method is relatively inaccurate compared to the thresholding method via frequency-of-seeing curves. This finding is much more obvious where VF sensitivity is low. This implies there may be a limit of the prediction performance where VF sensitivity is very low.

There are several limitations in the current study. First, predicted VF sensitivity was confined to the 10-2 test pattern rather than 24-2/30-2 test patterns because of the limited macular area captured by SD-OCT. Future wide-field OCT may solve the problem. Second, the model developed in the current study is not directly ready to be used in the clinical setting, because it has not been implemented in any medical support tools. However, it may be possible to integrate the current model into software in future. Third, the generalizability of the current findings may be limited. We used the same patients of the previous study¹¹ because the principal aim of the current study was to compare the prediction performance between the previous model (using only OCT data) and the current model (using both OCT and HFA 24-2/30-2 data). Future studies using other external large datasets are needed to validate the current findings. Fourth, there was a possibility that myopic eyes could bias the current results. Thus we iterated the analyses separating the eyes in the testing dataset to those with axial length of <26.5 mm and >26.5 mm.³² The resulting mean of MAE was 4.1 dB and 4.5 dB, respectively, which were almost the same values as that of 4.2 dB calculated in the overall eyes. Therefore we consider that bias caused by myopic eyes would not be very large. Fifth, similar to the fourth limitation, the existence of normal-tension glaucoma could bias the current results. The proportions of normal tension glaucoma and primary open-angle glaucoma were 29% (N = 30) and 14% (N = 15), respectively, in the testing dataset. The fine

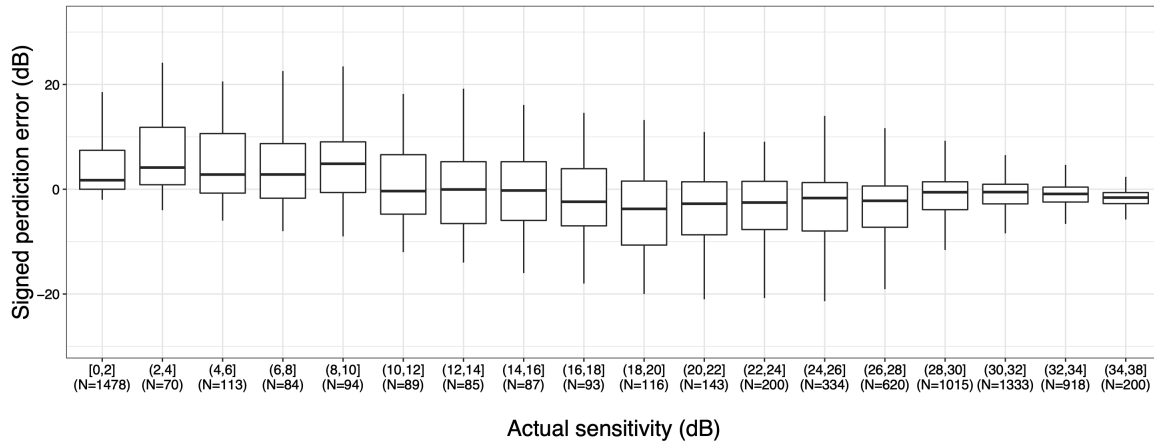


Figure 6. Pointwise signed prediction error (left eye). Mean (*upper row*) and standard deviation (*lower row*) values of all eyes at each predicted 10-2 VF test point are shown. Right eyes were mirror imaged.

classification (primary open-angle glaucoma or normal-tension glaucoma (NTG)) was unknown in the remaining eyes, because these patients were already under treatment when referred to our hospitals. The mean of MAE was 4.47 and 4.52 dB in eyes with normal-tension glaucoma and primary open angle glaucoma, respectively; these values were almost identical to that in the overall eyes (4.2 dB). However, future studies with larger sample size would be needed shedding light on this issue, because VF damages of NTG have been reported to be different from those in primary open-angle glaucoma. Sixth, by our definition, HFA 24-2/30-2 could have been done more than three months (maximum of six months) apart from the OCT measurement. However, all of the examined eyes had an interval of ≤ 3 months from OCT to HFA 24-2/30-2, except for one eye, so this would have only a negligible effect on the obtained results.³³

In conclusion, the current DL model (CNN-PBR), with correction based on 24-2/30-2 test results, demonstrated better prediction performance than the CNN-PBR model alone. Software/support tools equipped with this methodology, which we need to develop in the future, would be beneficial in the clinical setting.

Acknowledgments

Supported by the Ministry of Education, Culture, Sports, Science and Technology of Japan grants 18KK0253, 19H01114, 17K11418, 25861618, 20768254, and 00768351; the Translational Research program; Strategic Promotion for practical application of Innovative medical Technology (TR-SPRINT) from

Japan Agency for Medical Research and Development (AMED); JST-AIP JPMJCR19U4.

Disclosure: **Y. Hashimoto**, None; **T. Kiwaki**, None; **H. Sugiura**, None; **S. Asano**, None; **H. Murata**, None; **Y. Fujino**, None; **M. Matsuura**, None; **A. Miki**, None; **K. Mori**, None; **Y. Ikeda**, None; **T. Kanamoto**, None; **J. Yamagami**, None; **K. Inoue**, None; **M. Tanito**, None; **K. Yamanishi**, None; **R. Asaoka**, None

References

- Jonas JB, Aung T, Bourne RR, et al. Glaucoma. *Lancet*. 2017;390:2183–2193.
- Harwerth RS, Carter-Dawson L, Smith EL, et al. Neural Losses Correlated with Visual Losses in Clinical Perimetry. *Invest Ophthalmol Vis Sci*. 2004;45:3152.
- Kerrigan LA, Quigley HA, Pease ME, et al. Number of ganglion cells in glaucoma eyes compared with threshold visual field tests in the same persons. *Invest Ophthalmol Vis Sci*. 2000;41:8.
- Leite MT, Zangwill LM, Weinreb RN, et al. Structure-function relationships using the cirrus spectral domain optical coherence tomograph and standard automated perimetry. *J Glaucoma*. 2012;21:49–54.
- Hood DC. Improving our understanding, and detection, of glaucomatous damage: an approach based upon optical coherence tomography (OCT). *Prog Retin Eye Res*. 2017;57:46–75.
- Burgansky-Eliash Z, Wollstein G, Chu T, et al. Optical coherence tomography machine learning

- classifiers for glaucoma detection: a preliminary study. *Invest Ophthalmol Vis Sci.* 2005;46:4147.
7. Asaoka R, Hirasawa K, Iwase A, et al. Validating the usefulness of the “random forests” classifier to diagnose early glaucoma with optical coherence tomography. *Am J Ophthalmol.* 2017;174:95–103.
 8. Phene S, Dunn RC, Hammel N, et al. Deep learning and glaucoma specialists: The relative importance of optic disc features to predict glaucoma referral in fundus photographs. *Ophthalmology.* 2019;126:1627–1639.
 9. Asaoka R, Murata H, Hirasawa K, et al. Using deep learning and transfer learning to accurately diagnose early-onset glaucoma from macular optical coherence tomography images. *Am J Ophthalmol.* 2019;198:136–145.
 10. Park K, Kim J, Kim S, Shin J. Prediction of visual field from swept-source optical coherence tomography using deep learning algorithms. *Graefes Arch Clin Exp Ophthalmol.* 2020;258:2489–2499.
 11. Hashimoto Y, Asaoka R, Kiwaki T, et al. Deep learning model to predict visual field in central 10° from optical coherence tomography measurement in glaucoma. *Br J Ophthalmol.* 2020;105:507–513.
 12. Sugiura H, Kiwaki T, Yousefi S, et al. Estimating Glaucomatous Visual Sensitivity from Retinal Thickness with Pattern-Based Regularization and Visualization. In: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining - KDD '18.* London: ACM Press; 2018:783–792.
 13. Christopher M, Bowd C, Belghith A, et al. Deep learning approaches predict glaucomatous visual field damage from optical coherence tomography optic nerve head enface images and retinal nerve fiber layer thickness maps. *Ophthalmology.* 2020;127:346–356.
 14. Hemelings R, Elen B, Breda JB, et al. Pointwise visual field estimation from optical coherence tomography in glaucoma: a structure-function analysis using deep learning. *ArXiv210603793 Cs Eess.* 2021, <http://arxiv.org/abs/2106.03793>; Accessed September 12, 2021.
 15. Hood DC, Raza AS, de Moraes CGV, et al. Glaucomatous damage of the macula. *Prog Retin Eye Res.* 2013;32:1–21.
 16. Khoury JM, Donahue SP, Lavin PJ, Tsai JC. Comparison of 24-2 and 30-2 perimetry in glaucomatous and nonglaucomatous optic neuropathies. *J Neuroophthalmol.* 1999;19:100–108.
 17. Drance SM. The glaucomatous visual field. *Br J Ophthalmol.* 1972;56:186–200.
 18. Rao HL, Babu JG, Addepalli UK, et al. Retinal nerve fiber layer and macular inner retina measurements by spectral domain optical coherence tomograph in Indian eyes with early glaucoma. *Eye.* 2012;26:133–139.
 19. Crabb DP, Russell RA, Malik R, et al. Frequency of visual field testing when monitoring patients newly diagnosed with glaucoma: mixed methods and modelling. *Health Serv Deliv Res.* 2014;2:1–102.
 20. Asano S, Asaoka R, Murata H, et al. Predicting the central 10 degrees visual field in glaucoma by applying a deep learning algorithm to optical coherence tomography images. *Sci Rep.* 2021;11:2214.
 21. Das V. A novel diagnostic information based framework for super-resolution of retinal fundus images. *Comput Med Imaging Graph.* 2019;72:22–33.
 22. Thévenaz P, Blu T, Unser M Image interpolation and resampling. *Handbook of medical imaging, processing and analysis.* 2000;1:393–3420.
 23. He K, Zhang X, Ren S, Sun J Deep Residual Learning for Image Recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* Las Vegas: IEEE; 2016:770–778.
 24. Mwanza J-C, Warren JL, Budenz DL Combining spectral domain optical coherence tomography structural parameters for the diagnosis of glaucoma with early visual field loss. *Invest Ophthalmol Vis Sci.* 2013;54:8393.
 25. Baskaran M, Ong E-L, Li J-L, et al. Classification algorithms enhance the discrimination of glaucoma from normal eyes using high-definition optical coherence tomography. *Invest Ophthalmol Vis Sci.* 2012;53:2314.
 26. Parrish RK, II, Schiffman J, Anderson DR Static and kinetic visual field testing: reproducibility in normal volunteers. *Arch Ophthalmol.* 1984;102:1497–1502.
 27. Heijl A, Lindgren G, Olsson J Normal variability of static perimetric threshold values across the central visual field. *Arch Ophthalmol.* 1987;105:1544–1549.
 28. Altangerel U, Spaeth GL, Rhee DJ Visual function, disability, and psychological impact of glaucoma. *Curr Opin Ophthalmol.* 2003;14:100–105.
 29. Hood DC, Kardon RH. A framework for comparing structural and functional measures of glaucomatous damage. *Prog Retin Eye Res.* 2007;26:688–710.
 30. Weber J, Schultze T, Ulrich H The visual field in advanced glaucoma. *Int Ophthalmol.* 1989;13:47–50.
 31. Swanson WH, Felius J, Pan F Perimetric defects and ganglion cell damage: Interpreting linear

- relations using a two-stage neural model. *Invest Ophthalmol Vis Sci.* 2004;45:466.
32. Ohno-Matsui K, Lai TYY, Lai C-C, Cheung CMG Updates of pathologic myopia. *Prog Retin Eye Res.* 2016;52:156–187.
33. Thonginnetra O, Greenstein VC, Chu D, et al. Normal versus high tension glaucoma: a comparison of functional and structural defects. *J Glaucoma.* 2010;19:151–157.