AMERICAN ACADEMY
OF OPHTHALMOLOGY®

Explainable Artificial Intelligence–Assisted Exploration of Clinically Significant Diabetic Retinal Neurodegeneration on OCT Images

Miyo Yoshida, MD, Tomoaki Murakami, MD, PhD, Kenji Ishihara, MD, PhD, Yuki Mori, MD, PhD, Akitaka Tsujikawa, MD, PhD

Purpose: To explore clinically significant diabetic retinal neurodegeneration in OCT images using explainable artificial intelligence (XAI) and subsequent evaluation by retinal specialists.

Design: A single-center, retrospective, consecutive case series.

Participants: Three hundred ninety-seven eyes from 397 diabetic retinopathy patients for XAI-based screening and 244 fellow eyes for subjective human evaluation.

Methods: We acquired 30° horizontal OCT images centered on the fovea. An artificial intelligence (AI) model was developed to infer visual acuity (VA) reduction using fine-tuned RETFound-OCT. Attention maps highlighting regions contributing to VA inference were generated using layer-wise relevance propagation. Retinal specialists assessed OCT findings based on salient regions indicated by XAI. Two newly described findings, a needle-like appearance of the ganglion cell layer (GCL)/inner plexiform layer (IPL) ("ice-pick sign") and dot-like alterations in the outer nuclear layer (ONL) ("salt-and-pepper sign"), were evaluated alongside 2 established findings: EZ disruption and choroidal hypertransmission.

Main Outcome Measures: Identification of clinically significant OCT findings associated with diabetic retinal neurodegeneration.

Results: The AI model effectively discriminated eyes with poor vision (decimal VA $\leq$0.5) from those with good vision (VA $\geq$1.0) (area under the receiver operating characteristic curve of 0.947). Explainable artificial intelligence–based attention maps highlighted salient regions in the GCL/IPL (65.2% or 70.0%), ONL (52.2% or 28.3%), EZ (39.1% or 21.7%), and choroid (26.1% or 5.00%) in eyes with poor or good vision, respectively. Subjective evaluation by retinal specialists revealed the frequencies of these 4 findings as follows: ice-pick sign (32.4%), EZ disruption (25.0%), salt-and-pepper sign (16.0%), and choroidal hypertransmission (13.5%). Eyes with decimal VA $\leq$0.9 had these findings more frequently than those with VA $\geq$1.0 ($P < 0.001$ for all comparisons). Salt-and-pepper sign and choroidal hypertransmission exhibited high specificity for identifying eyes with poor vision. Statistical analyses demonstrated more significant associations between EZ disruption, salt-and-pepper sign, and hypertransmission compared with their relationships with the ice-pick sign.

Conclusions: Artificial intelligence–assisted exploration of OCT findings identified 2 established lesions and 2 novel OCT biomarkers indicative of clinically significant diabetic retinal neurodegeneration.

Financial Disclosure(s): The author(s) have no proprietary or commercial interest in any materials discussed in this article. *Ophthalmology Science 2025;5:100804 © 2025 by the American Academy of Ophthalmology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).*

Supplemental material available at www.ophthalmologyscience.org.

Diabetic retinopathy (DR) is a leading cause of blindness in the working-age population, characterized by retinal damage through various pathological mechanisms.[1] In addition to angiogenesis and increased vascular permeability in proliferative DR and diabetic macular edema (DME), inflammatory responses and neuronal lesions exacerbate visual impairment. Disruption of the neurovascular unit has prompted the introduction of the updated staging system, diabetic retinal disease.[2] Retinal function assessments, such as retinal sensitivity and electroretinogram, indicate impairment before clinical DR development, supported by evidence of retinal neuron loss, including ganglion cells.[3–6] These findings have led to the recognition of diabetic retinal neurodegeneration.[7]

Structural OCT enables detailed visualization of neuronal components in both healthy and diseased retinas, including diabetic retinal disease. The ganglion cell layer/inner plexiform layer (GCL/IPL) is thinner in DR patients compared with healthy individuals, potentially reflecting histological ganglion cell loss.[8–11] Nonperfused areas exhibit disrupted lamellar structure in inner retinas on OCT, and disorganization of retinal inner layers and ellipsoid zone (EZ)

https://doi.org/10.1016/j.xops.2025.100804
ISSN 2666-9145/25

disruption are associated with poor vision in DME.[12−16] Moreover, lesions in the retinal pigment epithelium (RPE) and choriocapillaris have been detected using fundus autofluorescence and OCT angiography, respectively.[15,17]

Explainable artificial intelligence (XAI) has gained significant attention for enhancing the transparency and interpretability of black-box artificial intelligence (AI) models.[18] In classification and regression tasks, tools like SHapley Additive exPlanations illustrate input−output dependencies. For image analyses, backpropagation-based methods such as layer-wise relevance propagation (RELPROP) and gradient-weighted class activation mapping generate saliency maps that highlight regions influencing inference.[19] Saliency maps have been successfully applied to fundus photographs and OCT images to infer DR and DME.[20,21]

Recently, the RETFound foundation model for fundus imaging demonstrated versatility in diagnosing ocular diseases and inferring systemic conditions through fine-tuning on large pretrained datasets.[22] Its capabilities also extend to generating task-specific attention maps. However, the potential of XAI to assist clinical researchers in defining novel clinical entities remains unclear.

In this study, we utilized XAI to identify salient OCT regions relevant to visual acuity (VA) inference in DR and conducted an expert-driven exploration of known and novel lesions through subjective retinal assessment.

Methods

Participants

This retrospective study consecutively enrolled patients with DR without macular edema whose spectral-domain OCT (SD-OCT) images of sufficient quality were acquired at the Department of Ophthalmology, Kyoto University Hospital. The study adhered to the Declaration of Helsinki and was approved by the Ethics Committee of the Kyoto University Graduate School and Faculty of Medicine. Inclusion criteria are as follows: presence of DR, included confirmed absence of center-involving DME (central subfield thickness >305 μm or 320 μm for women and men, respectively), and written informed consent. Exclusion criteria were (1) axial length <22 mm or >26 mm, (2) severe media opacities, (3) other ocular diseases affecting VA, including glaucoma, epiretinal membrane, or vitreomacular traction, (4) ocular treatments within 6 months before image acquisition, such as intraocular surgery, anti-VEGF injections, ocular steroids, or photocoagulation, (5) prior vitrectomy combined with removal of the epiretinal membrane, vitreomacular traction, or internal limiting membrane, and (6) poor image quality. If both eyes met the inclusion criteria, 1 eye was randomly selected for AI-based screening, while the fellow eye was used for subjective evaluation by retinal specialists.

Imaging

After best-corrected decimal VA was measured, and subsequently, comprehensive ophthalmic examinations were performed, and DR severity was classified according to the International Clinical Diabetic Retinopathy Severity Scales.[23] Axial length was measured using partial coherence interferometry (IOL Master, Carl Zeiss Meditec, Inc). Spectral-domain OCT images were acquired using the Spectralis OCT (Heidelberg Engineering) in raster scan (49 sections of 20° B-scans, each composed of 512 A-scans) and 30° cross-hair modes (composed of 1536 A-scans). Based on preliminary evaluation (data not shown), we selected 20° horizontal retinal sections cropped from 30° cross-hair B-scan images to create AI models for VA inference and to assess OCT findings subjectively. After constructing three-dimensional images based on raster scan data, central subfield thickness (within a 1 mm diameter) was measured using the ETDRS grid.

Explainable Artificial Intelligence

A fine-tuned VA prediction model based on RETFound-OCT was prepared using a previously described method with modifications.[22] A stepwise procedure included (1) VA labeling for the OCT image, (2) fine-tuning RETFound-OCT, (3) validation of the prediction model, (4) receiver operating characteristic curve analysis, and (5) attention map creation.

We classified and labeled eyes into good VA (decimal VA ≥1.0) and 4 thresholds of poor VA (decimal VA ≤0.3, 0.5, 0.7, or 0.9 [Snellen equivalent ≤20/63, 20/40, 20/29, or 20/22]). Eyes were randomly allocated into training, validation, and test datasets in a 4:1:2 ratio. Cross-entropy loss for training and validation datasets was automatically calculated by RETFound at each epoch to evaluate model accuracy and monitor potential overfitting, respectively. By observing loss value trends as the number of epochs increased, the optical epoch count was determined. Model weights corresponding to this selected epoch were saved as a checkpoint for evaluation on the validation set. Model performance was assessed using accuracy, F1-score, recall, and precision calculated from the validation dataset. Each saved checkpoint was subsequently applied to the test dataset to generate receiver operating characteristic curves, enabling evaluation of the model's clinical use.

Based on receiver operating characteristic analysis results and sample size considerations, the VA prediction model comparing eyes with decimal VA ≥1.0 (Snellen VA ≥20/20) and VA ≤0.5 (Snellen VA ≤20/40) was selected for further analysis. Attention maps were generated using RELPROP, an explanation tool specifically designed for transformer-based networks.[24] Test data were input into the prediction model, and backpropagation across network layers was conducted to calculate each pixel's contribution within the raw OCT images. Pixels with high contributions were visualized using warm colors, while low-contribution pixels were represented by cool colors, facilitating a visual interpretation of salient image regions related to VA inference.

Human Subjective Assessment

The salient regions highlighted in red on the attention maps were subjectively evaluated. Retinal and choroidal layers identified on sectional SD-OCT images included the nerve fiber layer, GCL/IPL, inner nuclear layer, outer plexiform layer, outer nuclear layer (ONL), ellipsoid zone (EZ), outer segments, RPE, and choroid. The retinal layers corresponding to the centroids of the salient regions were identified by retinal specialists.

The results of XAI encouraged us to evaluate morphological features in the GCL/IPL, ONL, EZ, and choroid on sectional and horizontal OCT images centered on the fovea. Thinning of the GCL/IPL around the fovea was characterized by a needle-like appearance, distinct from the typical scalpel-shaped GCL/IPL observed in healthy or glaucomatous eyes. In other words, the inner border of the GCL/IPL showed a straight contour in some eyes with DR, whereas it appeared convex in healthy subjects. This novel OCT finding was referred to as the "ice-pick sign." Its presence was further assessed in the nasal and temporal regions within the central 3 mm of the fovea.

Additional qualitative findings in the outer retinal layers and choroid were examined within the central 1 mm and the nasal and temporal subfields of the parafovea (1−3 mm diameter). Moderately reflective, small, oval, or round-shaped dots were diffusely delineated within the ONL in some eyes with DR, which we defined as the "salt-and-pepper sign." This finding was distinguished from hyperreflective foci, which vary in size and are observed as discrete lesions. Two previously recognized OCT findings, EZ disruption and choroidal hypertransmission, were also assessed.

Statistics

All data were expressed as the median (interquartile range), with statistical significance set at $P < 0.05$. Normality was evaluated using the Shapiro−Wilk test. Continuous variables were analyzed using the Mann−Whitney *U*-test or Kruskal−Wallis test with Bonferroni correction, while categorical variables were assessed using Fisher exact test or chi-square test. Spearman rank correlation coefficient was employed to evaluate statistical associations. All statistical analyses were conducted using SPSS software (version 24; IBM). Receiver operating characteristic curves were generated using Matplotlib (version 3.8.3) and Scikit-learn (version 1.4.1) libraries. Intergrader agreement for qualitative findings was evaluated using Cohen kappa coefficient.

Results

XAI-Assisted Screening of Clinically Relevant Regions

In this retrospective study, we analyzed 397 eyes from 397 DR patients without macular edema. The patients' characteristics are presented in Table S1 (available at www.ophthalmologyscience.org). Among these eyes, 207 exhibited good vision (decimal VA ≥1.0 [Snellen VA ≥20/20]), while 190, 125, 86, and 50 eyes had decimal VA ≤0.9, 0.7, 0.5, and 0.3, respectively (corresponding to Snellen VA ≤20/22, 20/29, 20/40, and 20/63). Using sectional OCT images, we developed VA inference models based on the fine-tuned RETFound-OCT framework. Performance metrics demonstrated higher accuracy and area under the receiver operating characteristic curve in inferring decimal VA ≤0.3 and 0.5 ([Snellen VA ≤20/63 and 20/40]; Table S2, available at www.ophthalmologyscience.org). These results were consistently observed in the test dataset (Fig. 1). Receiver operating characteristic analyses yielded similar results in cases without prior vitrectomy (area under the receiver operating characteristic curve = 0.951).

To identify retinal layers containing clinically relevant lesions, we applied RELPROP, an XAI method, to generate attention maps highlighting salient regions. These maps were created by comparing eyes with VA ≥1.0 (Snellen VA ≥20/20) to those with VA ≤0.5 (Snellen VA ≤20/40) (Fig. 2). Salient regions particularly associated with VA differences were frequently observed in the GCL/IPL, inner nuclear layer, ONL, EZ, RPE, and choroid, each demonstrating unique frequencies (Table 3). The EZ and ONL were more frequently highlighted in the nasal and central regions, whereas the choroid was more commonly emphasized in the temporal and central regions. Additionally, the EZ, ONL, and choroid were delineated as salient regions more frequently in eyes with poor vision than in those with good vision. Similar frequencies were observed in subgroup analyses of eyes with nonproliferative DR and proliferative DR and those without prior vitrectomy (Tables S4−S6, available at www.ophthalmologyscience.org).

Human Subjective Assessment of OCT Findings

Two retinal specialists independently evaluated the OCT findings. The patients' characteristics are presented in Table S7 (available at www.ophthalmologyscience.org). Four distinct findings were assessed on horizontal OCT images: the ice-pick sign in the GCL/IPL ($\kappa = 0.88$ and 0.89 in the nasal and temporal regions, respectively), the salt-and-pepper sign in the ONL ($\kappa = 0.96$, 0.95, and 0.94 in the nasal, central, and temporal regions), EZ disruption ($\kappa = 0.94$, 0.97, and 0.91), and choroidal hypertransmission ($\kappa = 0.98$, 0.95, and 0.95). Representative examples of each finding were shown in Figures 3 and S4 (available at www.ophthalmologyscience.org). The inner nuclear layer and RPE occasionally exhibited thinning or thickening,

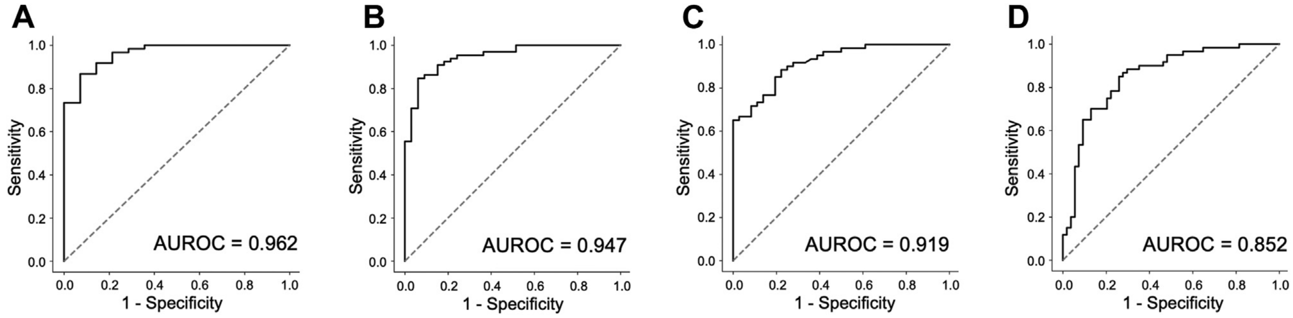


Figure 1. ROC curves for VA inference using fine-tuned RETFound-OCT models. ROC curves were generated from test data to infer poor VA (decimal VA thresholds of **[A]** ≤0.3 [Snellen equivalent = 20/63; 14 eyes], **[B]** ≤0.5 [20/40; 23 eyes], **[C]** ≤0.7 [20/29; 36 eyes], and **[D]** ≤0.9 [20/22; 54 eyes]), compared with VA ≥1.0 (20/20; 60 eyes). AUROC = area under the ROC curve; ROC = receiver operating characteristic; VA = visual acuity.

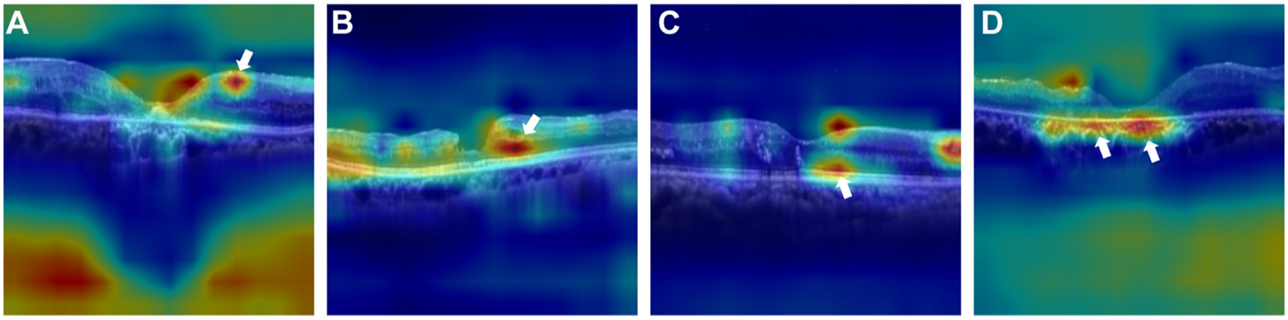


Figure 2. Salient regions identified by explainable artificial intelligence for VA comparison. Attention maps highlight clinically relevant layers in eyes with VA ≥1.0 (Snellen equivalent = 20/20) compared with those with VA ≤0.5 (20/40). Arrows indicate specific regions associated with VA differences, including **(A)** the ganglion cell layer/inner plexiform layer, **(B)** the outer nuclear layer, **(C)** the ellipsoid zone, and **(D)** the choroid (arrow = salient region). Red regions represent areas with high contributions to VA inference by the AI model. AI = artificial intelligence; VA = visual acuity.

either focally or diffusely; however, retinal specialists could not identify definitive qualitative findings specific to eyes with poor vision.

The frequencies of each finding across different VA levels were determined for the central, nasal, and temporal regions (Figs. 5 and S6–S9, available at www.ophthalmologyscience.org). In eyes with VA ≥0.7 (Snellen VA ≥20/29), the ice-pick sign in the GCL/IPL was observed more frequently than the other OCT findings. Across all VA levels, the salt-and-pepper sign and choroidal hypertransmission were less frequently delineated than EZ disruption. Eyes with these OCT findings had poorer VA than those without them (Table S8, available at www.ophthalmologyscience.org). These 4 findings demonstrated high specificity in distinguishing eyes with decimal VA ≥1.0 (Snellen VA ≥20/20) from those with VA ≤0.9 (Snellen VA ≤20/22) (Table S9, available at www.ophthalmologyscience.org). Subgroup analyses of eyes without prior vitrectomy showed similar results (Table S10, available at www.ophthalmologyscience.org). These OCT findings were more commonly observed in eyes with moderate nonproliferative DR or proliferative DR (Table S11, available at www.ophthalmologyscience.org).

All 4 findings were statistically associated with one another. A correlation matrix revealed that the ice-pick sign exhibited low correlation with the other 3 findings (Fig. 10). In contrast, EZ disruption, the salt-and-pepper sign, and choroidal hypertransmission coincided more frequently, particularly within the same anatomical regions. Additionally, the relative frequencies of each subjective OCT finding supported these statistical associations (Table S12, available at www.ophthalmologyscience.org). Ellipsoid zone disruption, the salt-and-pepper sign, and choroidal hypertransmission were more frequently delineated in the central and temporal regions compared with other combinations. In some cases, these OCT findings were observed exclusively in the central or temporal regions.

Discussion

This study investigated the morphological characteristics of the neuroretina associated with VA in DR patients without

Table 3. Salient Regions in Attention Maps Comparing 23 Eyes with Decimal VA ≤0.5 (Snellen VA ≤20/40) and 60 Eyes with VA ≥1.0 (20/20)

	Nasal		Central		Temporal	
Decimal VA (Snellen)	*VA ≤0.5 (20/40)*	*VA ≥1.0 (20/20)*	*VA ≤0.5 (20/40)*	*VA ≥1.0 (20/20)*	*VA ≤0.5 (20/40)*	*VA ≥1.0 (20/20)*
NFL	1 (4.35%)	3 (5.00%)	–	–	0 (0.00%)	0 (0.00%)
GCL/IPL	10 (43.5%)	35 (58.3%)	–	–	9 (39.1%)	25 (41.7%)
INL	5 (21.7%)	17 (28.3%)	–	–	5 (21.7%)	10 (16.7%)
OPL	0 (0.00%)	3 (5.00%)	–	–	0 (0.00%)	6 (10.0%)
ONL	6 (26.1%)	9 (15.0%)	8 (34.8%)	10 (16.7%)	6 (26.1%)	4 (6.67%)
EZ	8 (34.8%)	4 (6.67%)	5 (21.7%)	10 (16.7%)	1 (4.35%)	2 (3.33%)
OS	1 (4.35%)	0 (0.00%)	0 (0.00%)	1 (1.67%)	0 (0.00%)	0 (0.00%)
RPE	4 (17.4%)	4 (6.67%)	2 (8.70%)	16 (26.7%)	3 (13.0%)	5 (8.33%)
Choroid	1 (4.35%)	1 (1.67%)	5 (21.7%)	2 (3.33%)	6 (26.1%)	2 (3.33%)

EZ = ellipsoid zone; GCL/IPL = ganglion cell layer/inner plexiform layer; INL = inner nuclear layer; NFL = nerve fiber layer; ONL = outer nuclear layer; OPL = outer plexiform layer; OS = photoreceptor outer segment; RPE = retinal pigment epithelium; VA = visual acuity.

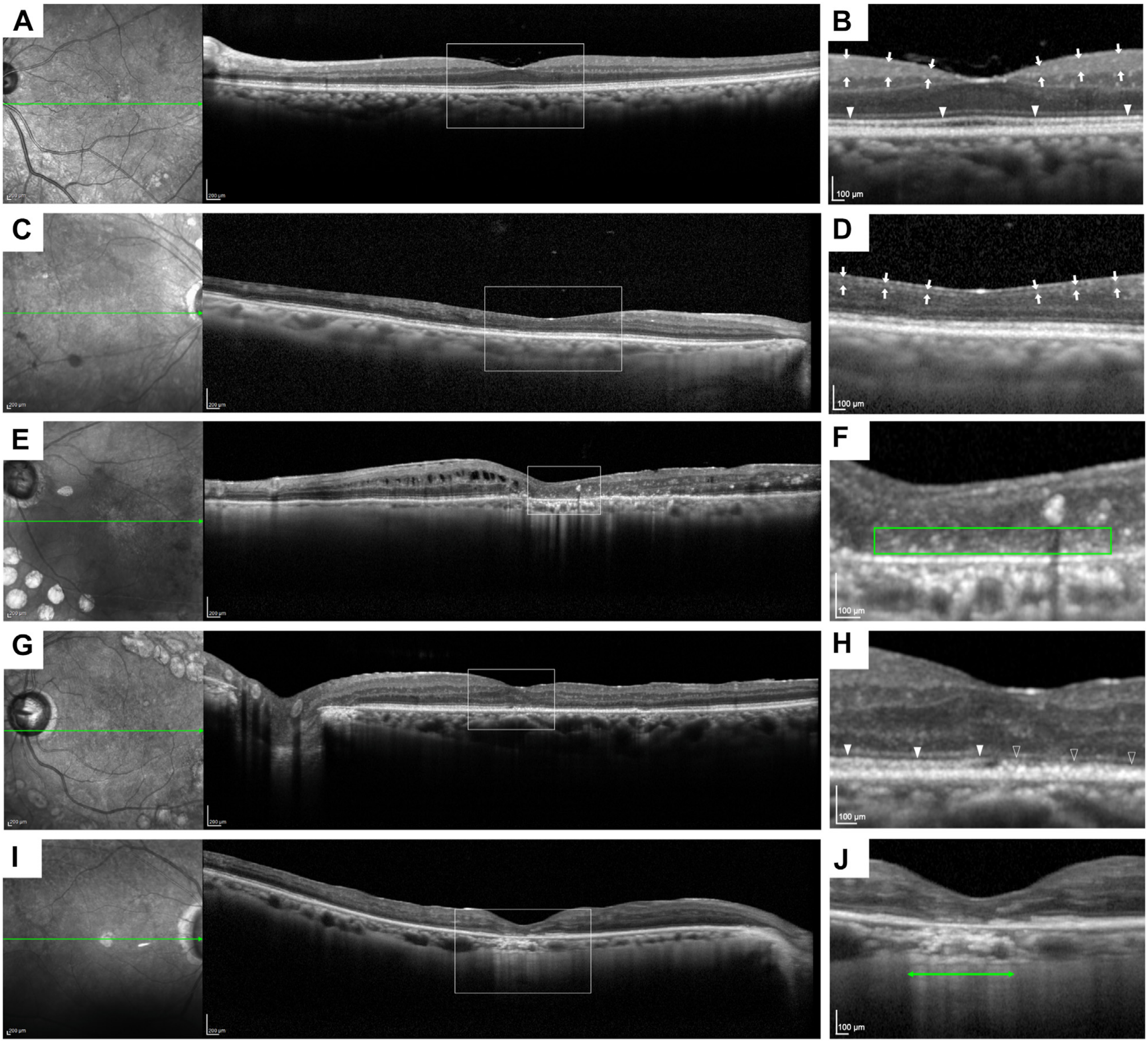


Figure 3. Representative OCT images from 5 patients with diabetic retinopathy. (B, D, F, H, J) Magnified views of the white rectangles in the original OCT images (A, C, E, G, I). **A and B,** A horizontal OCT image without the 4 specific findings. White arrows = GCL/IPL; white arrowheads = EZ. **C and D,** Ice-pick sign in the GCL/IPL (white arrows). (E, F) Salt-and-pepper sign in the outer nuclear layer (ONL), filled with moderately reflective dots (green rectangle). **G and H,** EZ disruption in the temporal areas (hollow arrowheads) compared with the intact EZ in the nasal areas (white arrowheads). **I and J**, Choroidal hypertransmission (green double-headed arrow). EZ = ellipsoid zone; GCL/IPL = ganglion cell layer/inner plexiform layer.

DME using OCT images. Artificial intelligence–based models for VA inference were constructed by fine-tuning the clinical foundation model RETFound-OCT. Through XAI-based analyses, attention maps were generated, highlighting salient regions on sectional OCT images that were potentially associated with VA differences. Notably, the GCL/IPL and photoreceptor/RPE layers were frequently emphasized. As a novel approach in human–computer interaction, retinal specialists interpreted both known and unknown OCT findings while referencing these saliency regions. This process confirmed 2 established findings: EZ disruption, commonly observed in DME, and choroidal hypertransmission, typically associated with age-related macular degeneration. Additionally, 2 novel findings were identified: the ice-pick sign in the GCL/IPL and the salt-and-pepper sign in the ONL. Statistical analyses demonstrated that the salt-and-pepper sign and choroidal hypertransmission were highly specific biomarkers for VA reduction.

We selected the fine-tuned RETFound-OCT model due to the limited number of available cases.[22] Fine-tuning pretrained foundation models requires fewer cases compared with traditional deep learning approaches. While models comparing decimal VA $\geq$1.0 versus VA $\leq$0.5 or 0.3 (Snellen VA $\geq$20/20 vs. VA $\leq$20/40 or 20/63) achieved high accuracy and area under the receiver operating characteristic curve, discrimination between VA $\leq$0.9 and VA

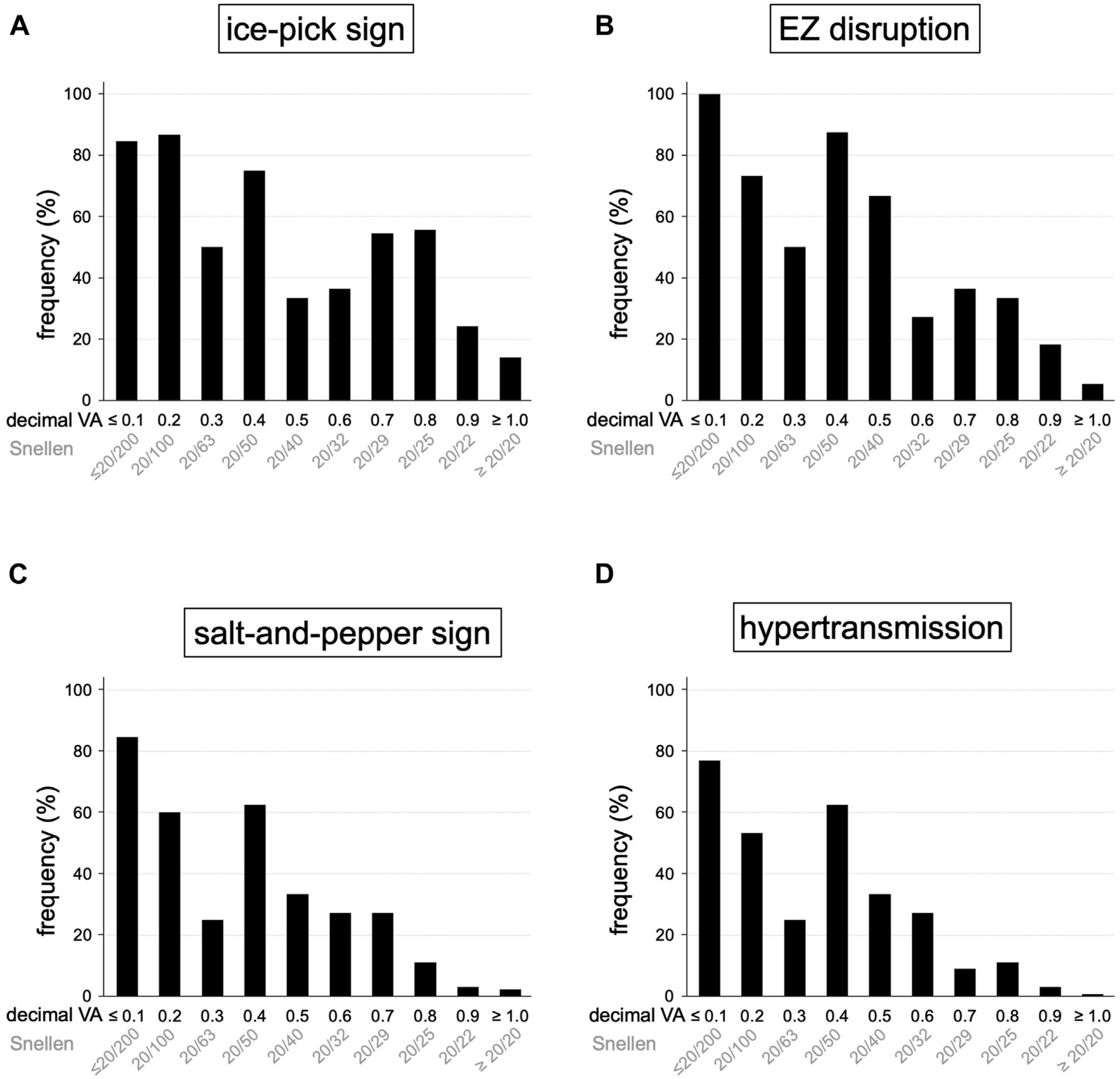
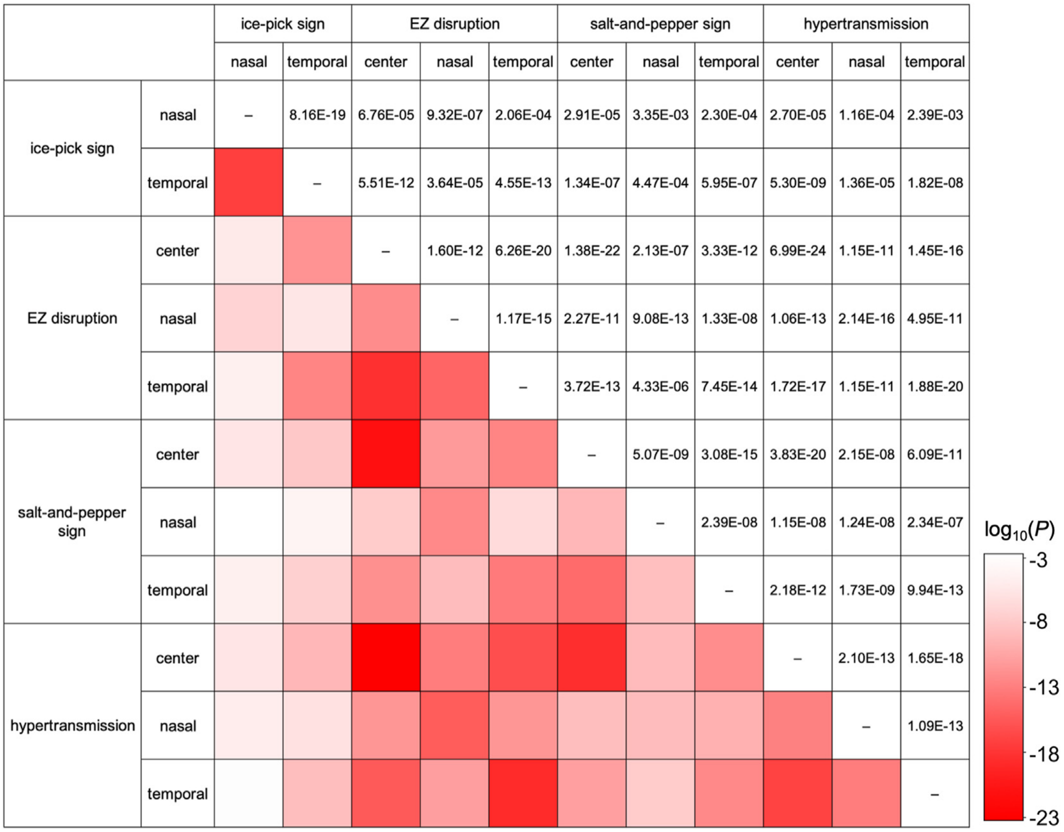


Figure 5. Frequencies of subjective OCT findings at different decimal VA levels. **A,** Ice-pick sign, **(B)** EZ disruption, **(C)** salt-and-pepper sign, and **(D)** choroidal hypertransmission. EZ = ellipsoid zone; VA = visual acuity.

≥1.0 (Snellen VA ≤20/22 and VA ≥20/20) was less precise. This suggests that VA prediction depends on structural features in OCT images, while other factors such as diabetic macular ischemia and transient neuronal dysfunction may also contribute.[25,26] Future multimodal approaches integrating structural and functional data may enhance model performance.

Attention maps created using RELPROP highlighted specific regions for further evaluation. However, the frequencies of salient regions indicated by XAI and the OCT findings identified by human assessment did not always align, underscoring the need for cautious interpretation. We hypothesized that OCT findings could be categorized into 3 types: newly developed lesions, structural loss, and structural deformation. The ice-pick sign may be classified as a form of retinal structural deformation. The AI model might detect this OCT finding based solely on its morphological features, regardless of the eye's pathological or healthy status. Consequently, the GCL/IPL was highlighted in the attention maps for eyes with both good and poor vision. Conversely, the EZ, ONL, and choroid were more frequently emphasized in eyes with poor vision, possibly due to their association with pathological lesions.

Consistent with previous research, GCL/IPL thinning was observed in eyes with poor vision, likely reflecting retinal ganglion cell dendrite and cell body involvement



		ice-pick sign		EZ disruption			salt-and-pepper sign			hypertransmission		
		nasal	temporal	center	nasal	temporal	center	nasal	temporal	center	nasal	temporal
ice-pick sign	nasal	–	8.16E-19	6.76E-05	9.32E-07	2.06E-04	2.91E-05	3.35E-03	2.30E-04	2.70E-05	1.16E-04	2.39E-03
	temporal		–	5.51E-12	3.64E-05	4.55E-13	1.34E-07	4.47E-04	5.95E-07	5.30E-09	1.36E-05	1.82E-08
EZ disruption	center			–	1.60E-12	6.26E-20	1.38E-22	2.13E-07	3.33E-12	6.99E-24	1.15E-11	1.45E-16
	nasal				–	1.17E-15	2.27E-11	9.08E-13	1.33E-08	1.06E-13	2.14E-16	4.95E-11
	temporal					–	3.72E-13	4.33E-06	7.45E-14	1.72E-17	1.15E-11	1.88E-20
salt-and-pepper sign	center						–	5.07E-09	3.08E-15	3.83E-20	2.15E-08	6.09E-11
	nasal							–	2.39E-08	1.15E-08	1.24E-08	2.34E-07
	temporal								–	2.18E-12	1.73E-09	9.94E-13
hypertransmission	center									–	2.10E-13	1.65E-18
	nasal										–	1.09E-13
	temporal											–

Figure 10. Correlation matrix for OCT findings. A pseudocolor map displaying statistical associations between 4 OCT findings (ice-pick sign, EZ disruption, salt-and-pepper sign, and choroidal hypertransmission) in the nasal, central, and temporal regions. Red indicates strong associations, while white represents weak associations. EZ = ellipsoid zone.

in diabetic eyes.[27,28] The ice-pick sign was qualitatively defined as needle-like thinning of the GCL/IPL, distinct from the scalpel-like thinning observed in glaucoma and the layered appearance seen in ischemic retinal vein occlusion.[29] Future studies should develop AI-driven models to automatically distinguish GCL/IPL thinning patterns. Furthermore, while hyperglycemia directly impairs retinal neurons, the relationship between GCL/IPL thinning and diabetic macular ischemia warrants further investigation.

Ellipsoid zone disruption was associated with poor VA in DR eyes without macular edema, consistent with previous findings in DME.[12,30–33] Statistical analyses suggested that EZ disruption initially occurs in the central or temporal subfields before extending to the nasal subfield, a pattern distinct from the center-involving damage typical of DME. The frequent highlighting of the nasal subfield in attention maps may indicate late-stage photoreceptor damage and subsequent VA reduction.

Statistical analyses indicated that choroidal hypertransmission partially overlapped with areas of EZ disruption in DR. In age-related macular degeneration, this finding indicates synchronized photoreceptor and RPE atrophy.[34] It prompted us to hypothesize that the affected areas correspond to concurrent photoreceptor and RPE damage. This agrees with the previously proposed concept of diabetic retinal pigment epitheliopathy, supported by recent fundus autofluorescence studies.[15,35] Future longitudinal research should elucidate the clinical progression from EZ disruption to RPE damage and identify potential therapeutic targets.

The salt-and-pepper sign, a novel OCT finding, was characterized by moderately reflective dots diffusely located in the ONL, sometimes extended to other layers. This pattern differs from hyperreflective foci, which appear as highly reflective and clustered deposits in the outer plexiform layer and subretinal spaces. Hyperreflective foci are considered to correspond to lipid-laden macrophages.[36] We speculated that the salt-and-pepper sign reflects photoreceptor cell death and a corresponding reduction in ONL cell density.[37] Histological studies documenting liquefaction necrosis in diabetic retinas suggest that swollen cell bodies may also contribute to reduced cellular density.[38] Future clinicopathological investigations should explore the histological correlates of this OCT finding.

Deep learning is a widely accepted AI technique to assist image-based diagnosis, although its decision-making process often lacks transparency. Explainable artificial intelligence methods have been developed to visualize salient regions of input images that contribute to model output. However, various XAI approaches generate saliency maps at different stages of convolutional neural networks, and each may yield different results. Future studies are needed to establish a gold standard for generating and interpreting XAI saliency maps in ophthalmic images.

In the era of advanced AI, novel human−computer interaction frameworks offer promising avenues for addressing medical challenges.[39,40] Human-in-the-loop models, in which human annotators refine ambiguous data points, have been widely adopted to improve AI performance. We employed a complementary approach involving XAI-based screening followed by human assessment. Explainable artificial intelligence enhances model transparency and trust by highlighting known legions while also revealing unexpected areas for evaluation.[18] This bidirectional validation process could further confirm clinical relevance and facilitate the discovery of novel findings. Further studies should assess the efficacy of various collaborative models in clinical practice.

This preliminary study has several limitations. Although we excluded eyes that underwent vitrectomy combined with manipulation of the vitreomacular interface, we should consider prior vitrectomy as a confounding factor. It was conducted at a single center, which may limit generalizability. We evaluated only horizontal OCT sections, which were obtained using a specific OCT machine, and used specific AI models to infer VA and to create saliency maps. Future studies should incorporate other imaging modalities and validation datasets to ensure reproducibility. Additionally, while 4 qualitative OCT findings were proposed, their automatic detection through AI remains to be established. Comprehensive multimodal analyses integrating functional, histological, and OCT data are needed to advance the clinical understanding of diabetic retinal neurodegeneration.

In conclusion, we employed AI-assisted analysis to explore clinically relevant OCT findings indicative of neuroretinal damage in DR. This novel human−computer interaction scheme not only validated known lesions but also identified new OCT findings, advancing the clinical definition of diabetic retinal neurodegeneration and suggesting potential biomarkers for future therapeutic strategies.

Footnotes and Disclosures

Originally received: January 28, 2025.
Final revision: April 4, 2025.
Accepted: April 18, 2025.
Available online: April 24, 2025. Manuscript no. XOPS-D-25-00059.

Department of Ophthalmology and Visual Sciences, Kyoto University Graduate School of Medicine, Kyoto, Japan.

Disclosure(s):

All authors have completed and submitted the ICMJE disclosures form.

The authors have no proprietary or commercial interest in any materials discussed in this article.

This study was funded by a Grant-in-Aid for Scientific Research of the Japan Society for the Promotion of Science (grant number: 23K09004). The funding organization had no role in the design or conduct of this research.

HUMAN SUBJECTS: Human subjects were included in this study. The study adhered to the Declaration of Helsinki and was approved by the Ethics Committee of the Kyoto University Graduate School and Faculty of Medicine. Inclusion criteria are as follows: presence of diabetic retinopathy, included confirmed absence of center-involving diabetic macular edema (central subfield thickness >305 μm or 320 μm for women and men, respectively), and written informed consent.

No animal subjects were used in this study.

Author Contributions:

Conception and design: Yoshida, Murakami

Data collection: Yoshida, Murakami, Ishihara, Mori, Tsujikawa

Analysis and interpretation: Yoshida, Murakami, Tsujikawa

Obtained funding: Murakami

Overall responsibility: Murakami

Abbreviations and Acronyms:
AI = artificial intelligence; **DME** = diabetic macular edema; **DR** = diabetic retinopathy; **GCL** = ganglion cell layer; **IPL** = inner plexiform layer; **ONL** = outer nuclear layer; **RELPROP** = layer-wise relevance propagation; **RPE** = retinal pigment epithelium; **VA** = visual acuity; **XAI** = explainable artificial intelligence.

Keywords:
Diabetic retinal neurodegeneration, Explainable artificial intelligence, Foundation model, OCT, Visual acuity.

Correspondence:
Tomoaki Murakami, MD, PhD, Department of Ophthalmology and Visual Sciences, Kyoto University Graduate School of Medicine, 54 Shougoin Kawahara-cho, Sakyo-ku, Kyoto 606-8507, Japan. E-mail: mutomo@kuhp.kyoto-u.ac.jp.

References

1. Yau JWY, Rogers SL, Kawasaki R, et al. Global prevalence and major risk factors of diabetic retinopathy. *Diabetes Care*. 2012;35:556−564.
2. Sun JK, Aiello LP, Abràmoff MD, et al. Updating the staging system for diabetic retinal disease. *Ophthalmology*. 2021;128:490−493.
3. Bresnick GH, Korth K, Groo A, Palta M. Electroretinographic oscillatory potentials predict progression of diabetic retinopathy. *Arch Ophthalmol*. 1984;102:1307.
4. Barber AJ, Lieth E, Khin SA, et al. Neural apoptosis in the retina during experimental and human diabetes. Early onset and effect of insulin. *J Clin Invest*. 1998;102:783−791.
5. Sohn EH, van Dijk HW, Jiao C, et al. Retinal neurodegeneration may precede microvascular changes characteristic of diabetic retinopathy in diabetes mellitus. *Proc Natl Acad Sci U S A*. 2016;113:E2655−E2664.
6. Montesano G, Gervasoni A, Ferri P, et al. Structure−function relationship in early diabetic retinopathy: a spatial correlation

analysis with OCT and microperimetry. *Eye*. 2017;31:931−939.

7. Abramoff MD, Fort PE, Han IC, et al. Approach for a clinically useful comprehensive classification of vascular and neural aspects of diabetic retinal disease. *Invest Ophthalmol Vis Sci*. 2018;59:519.
8. Ng DS, Chiang PP, Tan G, et al. Retinal ganglion cell neuronal damage in diabetes and diabetic retinopathy. *Clin Exp Ophthalmol*. 2016;44:243−250.
9. Simó R, Stitt AW, Gardner TW. Neurodegeneration in diabetic retinopathy: does it really matter? *Diabetologia*. 2018;61:1902−1912.
10. Lee MW, Lee WH, Ryu CK, et al. Effects of prolonged type 2 diabetes on the inner retinal layer and macular microvasculature: an optical coherence tomography angiography study. *J Clin Med*. 2020;9:1849.
11. Lim HB, Shin YI, Lee MW, et al. Longitudinal changes in the peripapillary retinal nerve fiber layer thickness of patients with type 2 diabetes. *JAMA Ophthalmol*. 2019;137:1125.
12. Murakami T, Nishijima K, Akagi T, et al. Optical coherence tomographic reflectivity of photoreceptors beneath cystoid spaces in diabetic macular edema. *Invest Ophthalmol Vis Sci*. 2012;53:1506−1511.
13. Chung H, Park B, Shin HJ, Kim HC. Correlation of fundus autofluorescence with spectral-domain optical coherence tomography and vision in diabetic macular edema. *Ophthalmology*. 2012;119:1056−1065.
14. Sun JK, Lin MM, Lammer J, et al. Disorganization of the retinal inner layers as a predictor of visual acuity in eyes with center-involved diabetic macular edema. *JAMA Ophthalmol*. 2014;132:1309−1316.
15. Yoshitake S, Murakami T, Horii T, et al. Qualitative and quantitative characteristics of near-infrared autofluorescence in diabetic macular edema. *Ophthalmology*. 2014;121:1036−1044.
16. Dodo Y, Murakami T, Uji A, et al. Disorganized retinal lamellar structures in nonperfused areas of diabetic retinopathy. *Invest Ophthalmol Vis Sci*. 2015;56:2012−2020.
17. Dodo Y, Suzuma K, Ishihara K, et al. Clinical relevance of reduced decorrelation signals in the diabetic inner choroid on optical coherence tomography angiography. *Sci Rep*. 2017;7:5227.
18. Barredo Arrieta A, Díaz-Rodríguez N, Del Ser J, et al. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf Fusion*. 2020;58:82−115.
19. Binder A, Montavon G, Lapuschkin S, et al. *Layer-wise relevance propagation for neural networks with local renormalization layers*. Barcelona, Spain: ICANN; 2016:63−71.
20. Gargeya R, Leng T. Automated identification of diabetic retinopathy using deep learning. *Ophthalmology*. 2017;124:962−969.
21. Liu Y, Zhang F, Gao X, et al. Lesion-aware attention network for diabetic nephropathy diagnosis with optical coherence tomography images. *Front Med (Lausanne)*. 2023;10:1259478.
22. Zhou Y, Chia MA, Wagner SK, et al. A foundation model for generalizable disease detection from retinal images. *Nature*. 2023;622:156−163.
23. Wilkinson CP, Ferris FL, Klein RE, et al. Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. *Ophthalmology*. 2003;110:1677−1682.
24. Chefer H, Gur S, Wolf L. *Transformer interpretability beyond attention visualization. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. New York, NY: The Computer Vision Foundation; 2021:782−791.
25. Cheung CMG, Fawzi A, Teo KY, et al. Diabetic macular ischaemia- a new therapeutic target? *Prog Retin Eye Res*. 2022;89:101033.
26. Terada N, Murakami T, Ishihara K, et al. Clinical relevance of parafoveal intercapillary spaces and foveal avascular zone in diabetic retinopathy without macular edema. *Invest Ophthalmol Vis Sci*. 2022;63:4.
27. El-Fayoumi D, Badr Eldine NM, Esmael AF, et al. Retinal nerve fiber layer and ganglion cell complex thicknesses are reduced in children with type 1 diabetes with no evidence of vascular retinopathy. *Invest Opthalmol Vis Sci*. 2016;57:5355.
28. Tang Z, Chan MY, Leung WY, et al. Assessment of retinal neurodegeneration with spectral-domain optical coherence tomography: a systematic review and meta-analysis. *Eye*. 2021;35:1317−1325.
29. Beykin G, Norcia AM, Srinivasan VJ, et al. Discovery and clinical translation of novel glaucoma biomarkers. *Prog Retin Eye Res*. 2021;80:100875.
30. Alasil T, Keane PA, Updike JF, et al. Relationship between optical coherence tomography retinal parameters and visual acuity in diabetic macular edema. *Ophthalmology*. 2010;117:2379−2386.
31. Spaide RF, Curcio CA. Anatomical correlates to the bands seen in the outer retina by optical coherence tomography. *Retina*. 2011;31:1609−1619.
32. Forooghian F, Stetson PF, Meyer SA, et al. Relationship between photoreceptor outer segment length and visual acuity in diabetic macular edema. *Retina*. 2010;30:63−70.
33. Scarinci F, Jampol LM, Linsenmeier RA, Fawzi AA. Association of diabetic macular nonperfusion with outer retinal disruption on optical coherence tomography. *JAMA Ophthalmol*. 2015;133:1036−1044.
34. Sadda SR, Guymer R, Holz FG, et al. Consensus definition for atrophy associated with age-related macular degeneration on OCT. *Ophthalmology*. 2018;125:537−548.
35. Kang EC, Seo Y, Byeon SH. Diabetic retinal pigment epitheliopathy: fundus autofluorescence and spectral-domain optical coherence tomography findings. *Graefes Arch Clin Exp Ophthalmol*. 2016;254:1931−1940.
36. Bolz M, Schmidt-Erfurth U, Deak G, et al. Optical coherence tomographic hyperreflective foci. *Ophthalmology*. 2009;116:914−920.
37. Park SH, Park JW, Park SJ, et al. Apoptotic death of photoreceptors in the streptozotocin-induced diabetic rat retina. *Diabetologia*. 2003;46:1260−1268.
38. Yanoff M, Fine BS, Brucker AJ, Eagle RC. Pathology of human cystoid macular edema. *Surv Ophthalmol*. 1984;28:505−511.
39. Krishnan R, Rajpurkar P, Topol EJ. Self-supervised learning in medicine and healthcare. *Nat Biomed Eng*. 2022;6:1346−1352.
40. Li J, Guan Z, Wang J, et al. Integrated image-based deep learning and language models for primary diabetes care. *Nat Med*. 2024;30:2886−2896.