Combining Structural and Vascular Parameters to Discriminate Among Glaucoma Patients, Glaucoma Suspects, and Healthy Subjects

Alessandro Rabiolo1–3,*, Federico Fantaguzzi2,3, Riccardo Sacconi2,3, Francesco Gelormini2,3, Enrico Borrelli2,3, Giacinto Triolo4, Paolo Bettin2,3, Andrew I. McNaught1,5, Joseph Caprioli6, Giuseppe Querques2,3, and Francesco Bandello2,3

1 Department of Ophthalmology, Gloucestershire Hospitals NHS Foundation Trust, Cheltenham, UK
2 School of Medicine, Vita-Salute San Raffaele University, Milan, Italy
3 Division of Head and Neck, Ophthalmology Unit, IRCCS Ospedale San Raffaele, Milan, Italy
4 Ophthalmology Department, Fatebenefratelli and Ophthalmic Hospital, ASST-Fatebenefratelli-Sacco, Milan, Italy
5 School of Health Professions (Faculty of Health), University of Plymouth, Plymouth, UK
6 Stein Eye Institute, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA

Correspondence: Alessandro Rabiolo, Department of Ophthalmology, Gloucestershire Hospitals NHS Foundation Trust, Sandford Road, Cheltenham GL53 7AN, UK. e-mail: rabiolo.alessandro@gmail.com

Received: August 9, 2021
Accepted: November 14, 2021
Published: December 20, 2021

Keywords: peripapillary retinal nerve fiber layer; macular ganglion cell–inner plexiform layer; perfusion density; glaucoma diagnosis; machine learning

Purpose: Compare the ability of peripapillary and macular structural parameters, vascular parameters, and their integration to discriminate among glaucoma, suspected glaucoma (GS), and healthy controls (HCs).

Methods: In this study, 196 eyes of 119 patients with glaucoma (n = 81), patients with GS (n = 48), and HCs (n = 67) underwent optical coherence tomography (OCT) and OCT angiography to measure peripapillary retinal nerve fiber layer (pRNFL), macular ganglion cell–inner plexiform layer (mGCIPL) thicknesses, radial peripapillary capillary perfusion density (RPC-PD), and macular GCIPL perfusion density (GCIPL-PD). Parameters were integrated regionally with logistic regression and globally with machine learning algorithms. Diagnostic performances were evaluated with area under the receiver operating characteristic (AUROC) curves.

Results: Patients with glaucoma had mild to moderate damage (median, −3.3 dB; interquartile range, −6.5 to −1.4). In discriminating between patients with glaucoma and the HCs, pRNFL thickness had higher AUROC curve values than RPC-PD for average (0.87 vs. 0.62; P < 0.001), superior (0.86 vs. 0.54; P < 0.001), inferior (0.90 vs. 0.71; P < 0.001), and temporal (0.65 vs. 0.51; P = 0.02) quadrants. mGCIPL thickness had higher AUROC curve values than GCIPL-PD for average (0.84 vs. 0.68; P < 0.001), superotemporal (0.76 vs. 0.65; P = 0.016), superior (0.72 vs. 0.57; P = 0.004), and inferonasal (0.78 vs. 0.65, P = 0.012) sectors. All structural multisector indices had higher diagnostic ability than vascular ones (P < 0.001). Combined structural–vascular indices did not outperform structural indices. Similar results were found to discriminate glaucoma from GS.

Conclusions: Combining structural and vascular parameters in a structural-vascular index does not improve diagnostic ability over structural parameters alone.

Translational Relevance: OCT angiography does not add additional benefit to structural OCT in early to moderate glaucoma diagnosis.
Introduction

Glaucoma is a chronic, progressive optic neuropathy with typical alterations of the optic nerve head and retinal nerve fiber layer. Although stereoscopic optic disc photography has traditionally been used to evaluate structural abnormalities, peripapillary optical coherence tomography (OCT) and, more recently, macular OCT have gained increasing popularity among clinicians. There is strong evidence that structural parameters perform well in identifying patients with glaucoma, and this imaging modality represents a valuable tool for glaucoma diagnosis, including early glaucoma.1–4

Optical coherence tomography angiography (OCTA) is a recent technological advancement in OCT imaging that allows the assessment of retinal vasculature in a fast, non-invasive, and depth-resolved manner.5 Previous studies have shown that patients with glaucoma have decreased capillary perfusion density at both peripapillary and macular regions.6–11 A recent Ophthalmology Technology Assessment Report by the American Academy of Ophthalmology reviewing the published literature about OCTA in glaucoma has suggested that peripapillary and macular vascular parameters may be useful in the glaucoma diagnosis, providing complementary information to functional and structural parameters.12 However, it is still unclear whether vascular parameters provide any additional information over structural parameters, and their role in routine glaucoma clinical practice is uncertain.

In this study, we compare the ability of peripapillary and macular structural and vascular parameters to distinguish between patients with glaucoma and those with suspected glaucoma and healthy subjects and to test whether the combination of structural and vascular information in a single, composite index improves the diagnostic accuracy of either single modality used alone.

Methods

Study Population

Patients with glaucoma, suspected glaucoma, and healthy subjects were prospectively enrolled in this cross-sectional study between September 2016 and August 2019 at the Glaucoma Unit and General Ophthalmology Unit of the Department of Ophthalmology, San Raffaele Scientific Institute, Milan, Italy. The study was approved by the San Raffaele Hospital Ethics Committee and conformed to the tenets of the Declaration of Helsinki. All enrolled subjects provided written consent to participate in observational studies.

Patients with a clinical diagnosis of primary open-angle glaucoma or glaucoma suspect made by a glaucoma consultant were prospectively identified and invited to enter the study if they met the study criteria. Patients with glaucoma were defined by the presence of optic nerve head (ONH) glaucomatous changes (e.g., neuroretinal rim thinning or notching, characteristic retinal nerve fiber layer [RNFL] defects indicative of glaucoma) with or without visual field (VF) damage, as determined by the glaucoma consultant treating the patient. Patients with glaucoma were later confirmed to have evidence of glaucomatous VF loss in at least one eye. Patients with suspected glaucoma were defined by either suspicious-looking ONH (e.g., large cup, inter-eye cup asymmetry, cup vertical elongation, optic disc hemorrhage with no rim loss) as determined by the glaucoma consultant or ocular hypertension, defined as untreated intraocular pressure above 21 mmHg on two or more visits at least 6 months apart. Patients with suspected glaucoma were required to have normal VF with no evidence of characteristic glaucomatous damage, Glaucoma Hemifield Test within normal limits, and pattern standard deviation within 95% confidence limits. Healthy subjects were recruited in the ophthalmology department among staff members, patients’ spouses or friends (not relatives), and patients with no ocular disease who came for refraction to the General Ophthalmology unit. Healthy subjects had no evidence of glaucomatous damage and were required to have (1) normal appearing ONH at the dilated fundus examination with intact rim, intact RNFL, and inter-eye cup-to-disc ratio asymmetry < 0.2; (2) normal VF with Glaucoma Hemifield Test within normal limits and pattern standard deviation within 95% confidence limits; and (3) intraocular pressure ≤ 21 mmHg. Inclusion criteria common to all groups included age ≥ 18 years; open angles at the gonioscopy, defined as trabecular meshwork visible for >180°; and spherical equivalent between −6 and +3 diopters. Common exclusion criteria included the presence of any other ocular (other than glaucoma) or systemic disease potentially affecting OCT and OCTA results; significant media opacity or poor fixation preventing adequate image acquisition; and previous intraocular surgery except uncomplicated cataract extraction >6 months prior enrollment. The diagnostic category of each participant was assigned based on the worst eye; both eyes
Figure 1. Quantification of capillary perfusion density for peripapillary (A–D) and macular (E–F) angiograms. Peripapillary en face angiograms were binarized with the Zeiss nerve fiber layer microvasculature density (v0.9) algorithm. The algorithm applies a ring-shaped ROI to the image centered on the ONH with an inner and outer annulus of 2- and 6-mm diameter, excludes large retinal vessels, and quantifies the capillary perfusion density within the ROI area after applying a mask to exclude large retinal vessels. The superimposed green circular dotted line represents the RNFL circle scan used to quantify structural RNFL thickness. Similarly, macular en face angiograms were binarized with the Zeiss superficial and GCIPL analysis (v0.3) algorithm. The algorithm applies a ROI analogous to the GCIPL grid with a radius of 3.0 mm; excludes a central elliptical area (0.5-mm vertical radius and 0.6-mm horizontal radius) corresponding to the foveola; and calculates the capillary perfusion density within the ROI area.

of the patients were included in the analysis if they met the inclusion/exclusion criteria. Glaucoma, glaucoma suspect, and healthy control diagnoses were not dependent on OCT imaging. Goldmann applanation tonometry was used to measure intraocular pressure, and the consultant was not masked to tonometry values during measurements.

Each patient underwent comprehensive ophthalmic examinations, including Goldmann applanation tonometry, dynamic gonioscopy, ultrasound pachymetry, and dilated fundus examination; Humphrey VF with the Swedish Interactive Thresholding Algorithm (SITA)-standard algorithm, 30-2 grid, and stimulus III; spectral-domain OCT (Cirrus HD-OCT 5000; Carl Zeiss Meditec, Jena, Germany), and swept-source OCTA (PLEX Elite 9000; Carl Zeiss Meditec). OCT and OCTA imaging were acquired after pupil dilation. All of the OCT and OCTA images were reviewed by two investigators (AR, FF), and images with segmentation errors, poor centration of the peripapillary or macular ring, or known artifacts (e.g., motion, blinking) were excluded.

PlexElite 9000 Swept-Source OCT and OCTA

The PLEX Elite 9000 device (Carl Zeiss Meditec) was used to acquire peripapillary and macular 6 × 6-mm angiocubes centered on the ONH and fovea, respectively. Technical specifications and algorithms at the foundation of this imaging device have been detailed elsewhere.13,14 Anonymized raw data were exported and uploaded in the Advanced Retina Imaging (ARI) Network Hub (https://arinetworkhub.com), a cloud collaboration platform providing its members access to a wide range of prototype algorithms in development for research analysis of OCTA images acquired with the PLEX Elite 9000 device. Peripapillary and macular cubes were binarized with the nerve fiber layer microvasculature density algorithm (v0.9) and the superficial and ganglion cell–inner plexiform layer (GCIPL) analysis algorithm (v0.3), respectively, which are early prototype, proprietary algorithms developed by Zeiss (Figure 1).13,15 Because these are prototype algorithms, all output images and values were reviewed carefully.
to verify the accuracy of the analysis. With regard to structural parameters, average and sector-wise peripapillary RNFL and macular GCIPL thickness values were extracted. As a comparison, Cirrus and PlexElite structural measurements were compared, and excellent agreement was found (data not shown). For the peripapillary cube, a fully automated segmentation algorithm was applied to isolate the radial peripapillary capillary plexus (inner limiting membrane to the outer boundary of RNFL); a ring-shaped region of interest (ROI) centered on the ONH with an inner and outer annulus of 2- and 6-mm diameter was applied; and analysis was restricted to the area inside the ROI. After the exclusion of large retinal vessels, average and quadrant-wise peripapillary capillary perfusion densities (defined as the proportion of white pixels over the total ROI area) were calculated. For the macular cube, the algorithm output two angiography slabs: GCIPL and superficial capillary plexus (inner limiting membrane to the outer boundary of inner plexiform layer). The former slab was used in this study to ensure better matching with structural measurements. Average and sector-wise capillary perfusion densities were calculated for ROIs analogous to those of the GCIPL sector division.

Statistical Analysis

Statistical analysis was performed with the open-source software R (R Foundation for Statistical Computing, Vienna, Austria). All tests were two tailed, and \( P < 0.05 \) was considered statistically significant. The distribution (normal vs. non-normal) of all continuous variables was assessed through the inspection of frequency histograms and quantile–quantile plots. Continuous normal and non-normal variables are reported as mean \( \pm \) SD and median (interquartile range [IQR]), respectively; discrete variables are reported as frequencies and proportions.

Differences in the patient-related variables among the three groups were evaluated with analysis of variance (ANOVA) with Tukey post hoc test and \( \chi^2 \) test for continuous and categorical variables, respectively. Differences between demographic and eye-related clinical variables were evaluated with a linear mixed model, where the patient identification (ID) was the random effect variable to account for within-subject correlations due to the inclusion of two eyes of the same patient. Differences in structural and vascular parameters among groups were tested with linear mixed models, where the fixed factors were diagnostic status, age, and signal strength to account for age and signal strength differences among groups, and the patient ID was the random effect variable to account for within-subject correlations due to the inclusion of two eyes of the same patient.

Diagnostic performance was evaluated with area under the receiver operating characteristic (AUROC) curves and sensitivity at 90% specificity. All ROC curves were clustered for the patient ID (to account for the inclusion of both eyes of the same participant) to estimate, unbiased AUROC 95% confidence intervals and test pairwise differences, as proposed by Obuchowski.\(^{16}\) For this section of the study about diagnostic performance, ROC curves were built on age- and signal strength (SS)-adjusted structural and vascular parameters. Age- and SS-adjusted values were calculated with a method similar to that reported in previous articles.\(^{17,18}\) We calculated normative equations to obtain normal values for age, SS, and sector using a linear mixed model on healthy eyes. As we did not have a separate dataset of healthy eyes to use as an external normative dataset, model fitting and parameter estimation were performed using the same pool of healthy patients. To avoid using the same eye for both model fitting and parameter estimations, we used a twofold cross-validation design. Specifically, the dataset of healthy subjects was randomly divided in two partitions: the first half was used to estimate normative equations for the second half and vice versa. For patients with glaucoma and patients with suspected glaucoma, the normative models were calculated from all healthy controls. For each eye and each sector, age and SS normative values were predicted, and the differences between observed and normative values were calculated for all of the measurements of interest in the cohort of patients and used to build the ROC curves. This approach allowed us to account for age and SS without using these variables for discrimination among groups.

The integration of structural OCT and vascular OCTA parameters was performed with two methods: region-by-region and multisector. In the region-by-region method, corresponding structural and vascular regions (e.g., superior peripapillary quadrant) were integrated into a single index with the method proposed by Pepe.\(^{19}\) For each pair of corresponding structural and vascular regions, we fit a logistic model where the dependent variable was the binary diagnostic status (glaucoma vs. healthy or glaucoma vs. glaucoma suspects), and independent variables were the corresponding structural and vascular parameters. The fitted values extracted from the logistic model were used to generate the ROC curves.

Multisector indices were global multivariable indices and were generated with a variety of different algorithms, including elastic net regression and three
supervised machine learning classifiers (random forest, support vector machine [SVM], naïve Bayes). For each of the multisector methods, we built three distinct classifiers: (1) structural classifier, which combined all age- and SS-adjusted structural parameters; (2) vascular classifier, which combined all age- and SS-adjusted vascular parameters; and (3) combined structural–vascular classifier, which included all age- and SS-adjusted structural and vascular parameters. As these classifiers are based on a large number of variables and have a high degree of complexity, they may overfit the data. Overfitted models lead to biased overly optimistic diagnostic capabilities in the dataset used for generating them, and they do not perform well in new, unknown datasets. To avoid overfitting, we used a fivefold cross-validation to estimate the unbiased performance of the multisector models. In fivefold cross-validation, the dataset is split into five partitions of approximately the same size. The model is estimated on every subset but one and tested on the left-out subset, which is unknown to the model as it is not used to generate the model. The process is repeated until all of the partitions are used to train and test the model. The probabilities of having glaucoma estimated on the test samples were used to build the ROC curves.

Two sets of analyses were performed to distinguish between (1) patients with glaucoma and healthy controls and (2) patients with glaucoma and those with suspected glaucoma. We also performed a subanalysis dividing patients in early (mean deviation [MD] \(> -6\) dB) and moderate to advanced glaucoma (MD \(\leq -6\) dB) to investigate the diagnostic performance of the structural and vascular parameters as a function of the disease severity.

This study included 196 of 119 patients with glaucoma, glaucoma suspects, and healthy subjects, after excluding 27 eyes of 23 patients because of poor image quality, centration, and segmentation failure. Demographic and main clinical characteristics are listed in Table 1. Patients with glaucoma (62.8 ± 13.5 years) were significantly older than glaucoma suspect (50.2 ± 15.3 years; \(P < 0.001\)) and healthy controls (50.3 ± 14.4 years; \(P < 0.001\)), whereas there was no difference between glaucoma suspects and healthy subjects (\(P = 1\)). Overall, most patients with glaucoma had mild to moderate VF damage with a median mean deviation of \(-3.3\) dB (IQR, \(-6.5\) to \(-1.4\)). Fifty-four, 22, and five eyes had MD better than \(-6\) dB, between \(-6\) and \(-12\) dB, and worse than \(-12\) dB, respectively. SS did not significantly differ among groups for macular scans (\(P = 0.13\)). Healthy subjects had significantly higher SS values than both patients with glaucoma (\(P = 0.026\)) and glaucoma suspects (\(P = 0.005\)) at the peripapillary scan, whereas patients with glaucoma and patients with suspected glaucoma did not significantly differ (\(P = 0.55\)).

Peripapillary RNFL thickness and capillary perfusion density values are illustrated in Figure 2 and detailed in Supplementary Table S1. Peripapillary RNFL thickness values were statistically different among the various groups. For structural parameters, peripapillary RNFL differences between glaucoma and

### Table 1. Demographic and Main Clinical Data of Patient Cohort

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Glaucoma</th>
<th>GS</th>
<th>HC</th>
<th>(P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients/eyes, (n)</td>
<td>52/81</td>
<td>29/48</td>
<td>38/67</td>
<td>—</td>
</tr>
<tr>
<td>Age (y), mean ± SD</td>
<td>62.8 ± 13.5</td>
<td>50.2 ± 15.3</td>
<td>50.3 ± 14.4</td>
<td>(&lt;0.001)</td>
</tr>
<tr>
<td>Caucasian, (n)</td>
<td>52</td>
<td>29</td>
<td>38</td>
<td>—</td>
</tr>
<tr>
<td>Male/female, (n)</td>
<td>20/32</td>
<td>16/13</td>
<td>16/22</td>
<td>0.34</td>
</tr>
<tr>
<td>Right eye/left eye, (n)</td>
<td>40/41</td>
<td>23/25</td>
<td>34/33</td>
<td>—</td>
</tr>
<tr>
<td>SS nerve, median (IQR)</td>
<td>8.0 (7.0–9.0)</td>
<td>8.0 (7.0–8.25)</td>
<td>9.0 (8.0–9.0)</td>
<td>0.004</td>
</tr>
<tr>
<td>SS macula, median (IQR)</td>
<td>8.0 (8.0–9.0)</td>
<td>8.0 (8.0–9.0)</td>
<td>9.0 (8.0–9.5)</td>
<td>0.13</td>
</tr>
<tr>
<td>VF MD (dB), median (IQR)</td>
<td>(-3.3) ((-6.5) to (-1.4))</td>
<td>(-1.2) ((-2.5) to (-0.1))</td>
<td>0.0 ((-3.0) to (0.9))</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Disc area (mm(^2)), mean ± SD</td>
<td>1.9 ± 0.4</td>
<td>1.9 ± 0.4</td>
<td>1.9 ± 0.3</td>
<td>0.72</td>
</tr>
<tr>
<td>Rim area (mm(^2)), mean ± SD</td>
<td>0.9 ± 0.3</td>
<td>1.1 ± 0.2</td>
<td>1.4 ± 0.3</td>
<td>(&lt;0.001)</td>
</tr>
</tbody>
</table>

GS, glaucoma suspect; HC, healthy control. Bold values indicate a statistical significance at \(P < 0.05\).
Figure 2. Peripapillary RNFL thickness (top left), peripapillary capillary perfusion density (top right), macular GCIPL thickness (bottom left), and macular capillary perfusion density (bottom right) values among patients with glaucoma, patients with suspected glaucoma, and healthy subjects.

healthy controls were more pronounced than between glaucoma and glaucoma suspects; the peripapillary RNFL difference between glaucoma suspects and healthy subjects was small and significant only for the average \( (P = 0.014) \) and inferior quadrant \( (P = 0.031) \) values. In multivariable linear mixed models (Supplementary Table S2), age was not significantly associated with peripapillary RNFL values for any parameter of interest \( (P = 0.06 \text{ or above}) \), whereas SS was significantly associated with the peripapillary RNFL thickness of the temporal quadrant only \( (P = 0.016) \). For vascular parameters, patients with glaucoma had lower peripapillary capillary perfusion density than healthy controls only for the inferior quadrant \( (P = 0.001) \) and lower peripapillary capillary perfusion density than glaucoma suspects for the average \( (P = 0.002) \) and inferior \( (P < 0.001) \) quadrants. Peripapillary capillary perfusion density values did not differ significantly between glaucoma suspects and healthy subjects. As shown in Supplementary Table S2, older age was significantly associated with lower peripapillary perfusion density values at all quadrants \( (P = 0.03 \text{ or below}) \), except for the inferior \( (P = 0.81) \) and temporal \( (P = 0.07) \) quadrants; SS was not significantly associated with peripapillary perfusion density at any quadrant \( (P = 0.08 \text{ or above}) \). Macular structural GCIPL thickness and capillary perfusion density values are illustrated in Figure 2 and detailed in Supplementary Table S3. Patients with glaucoma had significantly lower macular GCIPL values for every measure of interest when compared with healthy controls \( (P < 0.001) \) and glaucoma suspects \( (P = 0.039 \text{ or below}) \). On the other hand, macular perfusion density in patients with glaucoma was lower for the average \( (P = 0.049 \text{ vs. glaucoma suspects}) \), inferior \( (P = 0.020 \text{ and } P = 0.001 \text{ for patients with glaucoma vs. controls and glaucoma suspects, respectively}) \), and inferotemporal sectors \( (P < 0.001 \text{ and } P < 0.001 \text{ for patients with glaucoma vs. controls and glaucoma suspects, respectively}) \), whereas all other sectors did not differ among groups. Both macular GCIPL and macular capillary perfusion density values were not significantly different between
Figure 3. ROC curves of peripapillary structural, vascular, and combined structural–vascular parameters to distinguish between patients with glaucoma and either healthy controls (top row) or glaucoma suspects (bottom row). Vertical dotted lines indicate a value of 1-specificity of 0.1. GS, glaucoma suspect.

glaucoma suspects and healthy controls. As shown in Supplementary Table S4, older age was significantly associated with thinner macular GCIPL for average ($P = 0.037$), superior ($P = 0.035$), and inferotemporal ($P = 0.049$) quadrants; older age was also associated with lower capillary perfusion density ($P = 0.006$ or below) for every parameter of interest. SS was significantly associated with macular perfusion density only for superior ($P = 0.034$) and superotemporal sectors ($P = 0.043$).

Diagnostic abilities of peripapillary structural RNFL, vascular peripapillary radial capillary perfusion density, and their regional integrated index are illustrated in Figure 3 and Supplementary Table S5. Peripapillary RNFL had a significantly higher ability than peripapillary perfusion density to discriminate between patients with glaucoma and healthy controls for the average value (AUROC 0.872 vs. 0.616; $P < 0.001$), inferior quadrant (AUROC 0.904 vs. 0.712; $P < 0.001$), superior quadrant (AUROC 0.855 vs. 0.540; $P < 0.001$), and temporal quadrant (AUROC 0.647 vs. 0.512; $P = 0.02$). There was no difference between the two imaging modalities with respect to the nasal quadrant. The structural–vascular integrated index was not superior to structural RNFL thickness alone in any sector. Similar results were obtained in the discrimination between patients with glaucoma and those with suspected glaucoma, with the difference that the diagnostic ability of the temporal quadrant was similar between structural and vascular parameters.

Diagnostic abilities of macular GCIPL thickness, capillary perfusion density, and their regionally integrated indices are illustrated in Figure 4 and Supplementary Table S5. For the distinction between patients with glaucoma and healthy subjects, macular GCIPL thickness outperformed capillary perfusion density for the average value (AUROC 0.842 vs. 0.675; $P < 0.001$), superotemporal sector (AUROC 0.756 vs. 0.648; $P = 0.016$), superior sector (AUROC 0.717 vs. 0.569; $P = 0.004$), superonasal sector (AUROC 0.701 vs. 0.557; $P = 0.01$), inferotemporal sector (AUROC 0.898 vs. 0.717; $P < 0.001$), inferior sector (AUROC 0.866 vs. 0.687; $P < 0.001$), and inferonasal sector (AUROC 0.783 vs. 0.650; $P = 0.012$). The integrated regional structural–vascular index was not superior to structural parameters for any macular comparison. Similar results were obtained in the distinction between patients with glaucoma and glaucoma suspects, with
the exception of the superonasal sector, which had similar discriminating ability between structural and vascular parameters.

Figure 5 and Supplementary Table S6 show the diagnostic properties of the multisector indices (structural, vascular, and structural–vascular combined). Regardless of the algorithm used, structural multisector indices had higher diagnostic abilities than the vascular ones to distinguish patients with glaucoma from healthy subjects ($P < 0.001$) and glaucoma suspects ($P \leq 0.004$). Among the tested algorithms, the structural SVM classifier had the best ability to discriminate between patients with glaucoma and healthy controls (AUROC 0.922). Combined structural–vascular classifiers did not outperform structural classifiers in the discrimination of patients with glaucoma from healthy subjects ($P \geq 0.17$) or glaucoma suspects ($P \geq 0.28$), except for the SVM between patients with glaucoma and glaucoma suspects ($P = 0.008$).

Table 2 illustrates sensitivity values at 90% specificity of the various imaging parameters. Overall, multisector classifiers had higher sensitivity at 90% specificity than the individual parameters. At matched specificity, sensitivities of vascular variables were consistently lower than structural ones, and the structural–vascular combined index (regardless of the combination method) performed similarly to structural parameters alone.

Analyses were repeated with stratification of eyes into mild and moderate to advanced glaucoma. As shown in Supplementary Figure S1 and Supplementary Tables S7 and S8, differences among groups for both structural and vascular parameters were greater in eyes with moderate to advanced disease than those with early glaucoma. In early glaucoma, all structural variables of interest, except peripapillary RNFL at the nasal and temporal sectors, were significantly different when compared with controls; on the other hand, no vascular variable differed significantly between the two groups. In moderate to advanced glaucoma, all structural thickness values were significantly thinner than for the controls, except for the peripapillary temporal quadrant. A significant difference between moderate to advanced glaucoma and healthy controls was seen for many vascular parameters, both in the peripapillary (average and inferior quadrants) and macular areas (average, superotemporal, inferotemporal, inferior, and inferonasal sectors). Similar results were found for the discriminative analysis (Supplementary Figs. S2 and S3, Supplementary Table S9). ROC curves for discriminating moderate to severe glaucoma from healthy controls had higher AUROC values than those for early glaucoma vs. controls. Structural parameters performed better than vascular parameters, and regional structural-vasculature integration did not significantly outperform structural data alone.
**Discussion**

In this study, we compared the ability of peripapillary and macular structural, vascular, and combined structural–vascular indices to discriminate between patients with glaucoma, patients with suspected glaucoma, and healthy subjects. We found that structural parameters had good diagnostic properties, which were significantly higher than vascular ones in both the peripapillary and macular areas. The combination of structural and vascular information into a unique region-by-region or multiparametric index obtained with supervised machine learning did not outperform structural OCT measurements.

Peripapillary and macular OCT are established imaging modalities in glaucoma. Our findings are in agreement with the multitude of studies showing the good ability of structural OCT to distinguish patients with glaucoma from glaucoma suspects and healthy subjects. In accordance with previous studies, we also found that average and inferior peripapillary RNFL thickness values were thinner in patients with suspected glaucoma than healthy subjects, whereas none of the macular GCIPL or OCTA parameters differed between these two groups. The explanation for this finding is speculative. In a longitudinal study, Lalezary and colleagues evaluated predictors of development of glaucomatous damage in a cohort of 114 patients with suspected glaucoma, and they found that thinner baseline peripapillary RNFL was independently predictive for glaucoma development. Patients with suspected glaucoma may have subclinical disease not detected by clinical evaluation, which might manifest itself later on; however, it remains unclear why differences were found only with the peripapillary, and not macular, structural OCT measures. Because patients with suspected glaucoma were defined by suspicious ONH features in the clinical examination, another possible explanation is that we selected patients differing from healthy subjects for structural peripapillary abnormalities that were evident on the peripapillary OCT.

*Figure 5.* ROC curves of structural, vascular, and combined structural–vascular multisector indices to distinguish between patients with glaucoma and either healthy controls (top row) or glaucoma suspects (bottom row). *Vertical dotted* lines indicate a value of 1-specificity of 0.1.
Table 2. Sensitivity at 90% Specificity

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Glaucoma vs. HC (%)</th>
<th>Glaucoma vs. GS (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Structure Vasculature SV Integration</td>
<td>Structure Vasculature SV Integration</td>
</tr>
<tr>
<td>Peripapillary</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>74.1</td>
<td>29.6</td>
</tr>
<tr>
<td>Superior quadrant</td>
<td>74.1</td>
<td>23.5</td>
</tr>
<tr>
<td>Nasal quadrant</td>
<td>17.3</td>
<td>14.8</td>
</tr>
<tr>
<td>Inferior quadrant</td>
<td>80.2</td>
<td>42.0</td>
</tr>
<tr>
<td>Temporal quadrant</td>
<td>32.1</td>
<td>9.9</td>
</tr>
<tr>
<td>Macula</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>58.0</td>
<td>29.6</td>
</tr>
<tr>
<td>Superotemporal</td>
<td>50.6</td>
<td>30.9</td>
</tr>
<tr>
<td>Superior</td>
<td>43.2</td>
<td>23.5</td>
</tr>
<tr>
<td>Superonasal</td>
<td>46.9</td>
<td>19.8</td>
</tr>
<tr>
<td>Inferotemporal</td>
<td>71.6</td>
<td>44.4</td>
</tr>
<tr>
<td>Inferior</td>
<td>64.2</td>
<td>39.5</td>
</tr>
<tr>
<td>Inferonasal</td>
<td>53.1</td>
<td>22.2</td>
</tr>
<tr>
<td>Multisector</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elasticnet regression</td>
<td>82.7</td>
<td>56.8</td>
</tr>
<tr>
<td>Random forest</td>
<td>80.2</td>
<td>49.4</td>
</tr>
<tr>
<td>SVM</td>
<td>84.0</td>
<td>43.2</td>
</tr>
<tr>
<td>Naïve Bayes</td>
<td>85.2</td>
<td>45.7</td>
</tr>
</tbody>
</table>

SV, structural–vascular.

Despite having fair diagnostic ability, vascular parameters performed worse than structural parameters in most of the selected sectors in both the peripapillary and macular regions. Comparing our results with those of other authors is not straightforward, as various studies have reported a wide range of AUROC values and are highly heterogeneous in terms of study population, glaucoma severity and subtype, OCTA device, angiocube size, and post-processing technique. Overall, AUROC values were in agreement with previous studies.\textsuperscript{5–7,10} In our study, peripapillary and macular vascular parameters had similar discriminative abilities. This finding is in agreement with a study by Chen and colleagues,\textsuperscript{7} which evaluated the ability of peripapillary and macular vascular parameters to discriminate between 26 patients with glaucoma and 27 healthy controls. In their study, the authors found equivalent and extremely high diagnostic performances of peripapillary and macular vascular parameters with AUROC curve values of 0.93 and 0.94, respectively. The exclusion of patients with preperimetric glaucoma and the inclusion of patients with more advanced glaucoma (average MD, −9 dB) is likely responsible for such overly optimistic AUROC curve values, which are the highest reported in the literature. On the other hand, other authors have reported different results. Both Rao et al.\textsuperscript{29} and Triolo et al.\textsuperscript{6} found that peripapillary perfusion density had better diagnostic performance (0.83 and 0.88, respectively) than macular perfusion density (0.63 and 0.71, respectively). Both of these studies quantified macular perfusion density on a small macular region, and it has been argued that wider angiocubes are required to detect glaucomatous damage, as it usually affects peripheral macular areas first.\textsuperscript{30} This, however, is not likely to be the reason for such a discrepancy. Richter and colleagues\textsuperscript{10} compared the diagnostic performance of peripapillary and macular vascular parameters with large angiocubes (6 × 6 mm) and found that macular vascular parameters had lower diagnostic accuracy than peripapillary ones. We also compared the diagnostic ability of a smaller region of interest, matching the GCIPL grid, and a larger one covering the entire 6 × 6-mm angiocube, and we did not observe any significant difference (data not shown).

OCTA is an important component of multimodal retinal imaging, and most recent OCT devices have an integrated angiographic module; this imaging modality will be widely available on devices already used in glaucoma practice. Although vascular parameters had lower diagnostic abilities than structural parameters, we investigated whether integrating these two
imaging modalities might improve the diagnosis of glaucoma. If structural and vascular parameters really provide complementary information, one would expect an increase in diagnostic ability with combined indices; on the other hand, if one measure is the carbon copy of the other, then the information will be redundant, and integrated indices will not likely add any extra value. We integrated structural and vascular parameters both sector by sector on a regional basis and with various statistics to obtain a multiparametric index. We found that, regardless of the method used, combined structural–vascular indices were not superior to structural parameters. There is limited, and inconsistent, evidence as to whether OCTA provides any additional advantage over structural OCT. Kwon and colleagues\textsuperscript{31} evaluated the diagnostic properties of peripapillary structural parameters, vascular parameters, and their combination in a cohort of Korean patients. They calculated sensitivity and specificity to diagnose glaucoma of structural, vascular, structural and vascular parameters combined with either AND or OR-Boolean logic, and found that structural-vascular combinations lead to increased sensitivity and specificity. This study has shortcomings: the combination of structural and vascular parameters relied on a rudimentary method, which transformed continuous parameters into binary classifiers based on the selection of optimal ROC cutoffs with Youden’s index. This approach has several disadvantages, including transforming a continuous variable into a binary one with significant loss of information, and the selection of cutoff values based on different parts of the ROC curve with different levels of sensitivity and specificity. Kim and associates\textsuperscript{32} investigated the diagnostic properties of macular GCIP\textsubscript{L}, superficial capillary perfusion density, and their combination sector by sector in a cohort of patients with normal-tension glaucoma and suspected glaucoma. They found that the inferotemporal macular sector had the best diagnostic ability for both structural and vascular parameters, and the integration of the two imaging modalities did not further improve glaucoma detection. The structural–vascular index calculated by Kim and colleagues\textsuperscript{32} was similar to the regional one used in our study and was based on the integration of corresponding structural and vascular sectors with logistic regression. Although the study by Kim et al.\textsuperscript{32} had limitations, including the small sample size, a specific subset of glaucoma, no information on ethnicities other than Korean, lack of a proper control group, and no information about peripapillary OCTA, its results are in accordance with our findings.

In a recent study, Bowd and colleagues\textsuperscript{33} combined peripapillary and macular structural and vascular parameters with a gradient boosting classifier, which is a supervised machine learning algorithm sharing many similarities with the random forest classifier used in our study. They found a significantly better performance of the classifier combining all structural and vascular parameters, although the difference was small and of limited clinical relevance. Andrade De Jesus and colleagues\textsuperscript{34} recently investigated the performance of multisector and multilayer indices combining all OCTA parameters in glaucoma diagnosis, and they found that such indices did not outperform peripapillary RNFL thickness. In our study, multiparametric classifiers combining all of the structural parameters had the best discriminative ability and the highest specificity at 90\% specificity than all of the other indices. This finding is not unexpected, as various studies have shown that the combination of various peripapillary and macular OCT parameters may improve the diagnostic ability of individual parameters.\textsuperscript{35,36}

Multiparameter indices were developed with the idea of maximizing OCT diagnostic properties and simplifying the clinician’s job by integrating multiple and often discordant parameters into a single classifier. Although these indices may be helpful summary metrics, we warn the reader against the uncritical and unconsidered use of such integrated indices in clinical practice, including those used in this study. Glaucoma has specific patterns of RNFL and GCIP\textsubscript{L} damage that differ from other optic neuropathies and are associated with specific patterns of visual field damage.\textsuperscript{37–39} In the attempt to provide a unique, simplified index, combined metrics do not retain spatial information. Also, some of the most recent combination algorithms, such as the ones used in this study, have a “black box” approach with very little control on how precisely the model behaves across the full range of clinical scenarios. Although having one best-performing structural metric can look appealing to many, skilled clinicians should be able to combine in their mind the multitude of structural, functional, clinical, and demographic information to diagnose glaucoma. However, combined indices may have a complementary role, similar to visual field global indices. The additional inclusion of vascular parameters in the machine learning classifiers did not significantly improve the diagnostic ability.

This study was not designed to answer the ultimate question of whether structural damage precedes vascular damage or vice versa. Differences in structural parameters were more pronounced than vascular ones; also, patients with early glaucoma demonstrated
significant structural damage with less evident vascular damage as measured with these techniques. This finding is in agreement with a recent cross-sectional study by Hirasawa and colleagues, who found that only one-third of early glaucoma had more vascular damage than structural damage, with lower SS being at least partially responsible for more pronounced vascular damage. Hence, one may conclude that the reduced perfusion density is secondary to the ganglion cell degeneration, with vascular damage being a direct consequence of structural damage. Glaucoma, however, is a complex and heterogeneous group of diseases, and the relative contribution of structural versus vascular abnormalities may vary in different patients and at different stages of the disease. Also, OCTA is a fairly recent technology, and, at the current stage of technological development, its resolution might be insufficient to identify early and subtle vascular changes. In contrast, OCT imaging was developed in the early 1990s and so has undergone several hardware and software improvements. Whether structural damage precedes vascular damage cannot be determined with cross-sectional studies but only with careful, long-term longitudinal studies.

A limitation of this study is that the different subgroups were not matched for age and signal strength. Age differences among groups in case-control glaucoma diagnostic studies are not uncommon, as patients with glaucoma are often older than healthy subjects and glaucoma suspects. Because older age and lower SS are associated with lower structural and vascular quantitative values, unadjusted values may lead to biased results and overestimate the true diagnostic ability of structural and vascular parameters. Previous studies tried to address this issue by integrating unbalanced covariates (e.g., age) into a single index with multivariable logistic regression and using regression fitted values to build ROC curves, as proposed by Pepe. However, we believe that the use of this statistical technique in this regard is flawed and may lead to biased high AUROC curve values. Suppose that age and SS values are added into a multivariable classifier along with other parameters of interest. In that case, the model will not compensate for differences in these confounders but instead will use them to discriminate among groups. We used a more sophisticated statistical technique to compensate for age and SS differences without involuntarily using these variables as discriminators. This method is similar to that used by previous studies. Briefly, we estimated cross-validated normative equations to obtain normal values for age, SS, and sector, and we predicted sector-wise age and SS normative values for each study eye. Then, we calculated the difference between observed and normative values, which indicates how much a certain value deviates from that expected in a cohort of healthy patients of the same age and with the same SS value. The interpretation of this differential value is analogous to visual field total deviation. Differences between observed and normative values were finally used to build the ROC curves. Supplementary Figures S4 and S5 compare ROC curves for unadjusted capillary perfusion density values; capillary perfusion density values adjusted for age and SS with the method proposed by Pepe; capillary perfusion density values adjusted for age and SS with the method used in this study based on normative equations; and empty models with age and SS as covariates. It is evident that age and SS have a moderate discriminative ability by themselves, and their introduction into a multivariable model inflates the diagnostic accuracy of the parameter of interest. On the other hand, adjustment with normative equations produced ROC curves slightly lower than those obtained with unadjusted values, as one would expect when accounting for age and SS imbalance.

Other limitations of this study should be acknowledged. Differences in axial length may cause differences in the magnification of retinal images and inaccuracy in uncorrected capillary perfusion density calculations. As axial length was not measured in our cohort of patients, we did not apply any correction for image magnification. Because both structural and vascular parameters are prone to image magnification secondary to axial length differences, this limitation is likely to affect both measurements. Also, spherical equivalents, which are a surrogate measure for axial length, did not differ significantly among the study groups. All the subjects included in this study are of European descent, and results might not be generalizable to patients of different ethnicities. A previous study has highlighted ethnicity-related differences in macular capillary perfusion density in healthy subjects; however, clinical implications of this finding are still uncertain. A large number of statistical tests were performed in this study, and this increases the rate of type I error. We refrained, however, from performing any adjustments for multiple comparisons because of the exploratory nature of the study.

Vascular quantitative metrics are dependent on the OCTA instrument, angiocube size, and quantification algorithm; similarly, OCT-derived thickness values vary as a function of the device used. Although different instruments, scanning patterns, and post-processing and quantification algorithms may lead to slightly different absolute values, we believe that such small differences would not change the conclusions of
this study. Recent studies\textsuperscript{50} have suggested that
three-dimensional perfusion density analysis of angiocubes
might improve the diagnostic abilities of OCTA in
retinal diseases, but this is not commercially avail-
able and has never been applied to glaucoma; further
studies may investigate whether it improves glaucoma
discrimination. The definition of glaucoma in this
study mandated the presence of clinically detectable
structural damage, and this could have inflated the
true diagnostic abilities of structural and vascular
measurements. Because of its cross-sectional design
with predefined separation of cases and controls, this
study may have overestimated the diagnostic accuracy
of imaging modalities. For all of the reasons mentioned
above, we invite the readers to evaluate the absolute
AUROC curve values carefully. Finally, the diagno-
sis of glaucoma was based on the subjective evalu-
ation of the optic nerve appearance by experienced
glaucoma specialists. This resembles what happens in
clinical practice, where the treating physician makes
a diagnosis of glaucoma based on the whole clinical
picture rather than applying predefined criteria. Previ-
ous studies have shown substantial variability in the
interpretation of optic disc changes, even among expert
clinicians.\textsuperscript{51}

In conclusion, structural parameters discriminate
patients with glaucoma from those with suspected
glaucoma and healthy subjects better than vascular
parameters with both peripapillary and macular scans.
Combining structural and vascular parameters in a
structural–vascular index does not improve diagnostic
ability over structural parameters alone. Multiparamet-
ric indices based on a combination of structural param-
eters perform better than individual metrics, especially
at high specificity values. With current technology,
OCTA does not add additional benefit to structural
OCT for the diagnosis of early to moderate glaucoma.

\section*{Acknowledgments}

The authors thank Stephanie Magazzeni, Deborah
Cosette, Luis de Sisternes, Sophie Kubach, and Mary
Durbin for their support in the image analysis and
quantification.

Part of this study was presented as a poster at the
ARVO virtual annual meeting, 2020.

Disclosure: A. Rabiolo, None; F. Fantaguzzi,
None; R. Sacconi, None; F. Gelormini, None; E.
Borrelli, None; G. Triolo, None; P. Bettin, None;
A.I. McNaught, Allergan (F), Easyscan (F), Novartis
(F), Pfizer (F), Visufarma (F), Zeiss (F); J. Caprioli,
Aerie (F), Alcon (F), Allergan (F), Glokos (F), New
World Medical (F); G. Querques, Alimera Sciences (F),
Allergan (F), Amgen (F), Bayer Schering Pharma (F),
Heidelberg Engineering (F), KBH (F), LEH Pharma
(F), Lumbithera (F), Novartis (F), Sandoz (F), Sif (F),
Soof-Fidea (F), Zeiss (C, F); F. Bandello, Alcon (F),
Alimera Sciences (F), Allergan (F), Farmila Théa (F),
Bayer Schering Pharma (F), Bausch + Lomb (F),
Genentech (F), Hoffmann-La Roche (F), Novagali
Pharma (F), Novartis (F), Sanofi (F), Thrombogenics
(F), Zeiss (C, F)

\section*{References}

1. Nouri-Mahdavi K, Nowroozizadeh S, Nassiri N,
et al. Macular ganglion cell/inner plexiform layer
measurements by spectral domain optical coherence
tomography for detection of early glau-
coma and comparison to retinal nerve fiber layer
1307.e2.

Macular versus retinal nerve fiber layer parameters
for diagnosing manifest glaucoma: a system-
atic review of diagnostic accuracy studies. \textit{Ophthal-

Optic nerve head and fibre layer imaging for diag-
nosing glaucoma. \textit{Cochrane Database Syst Rev.}
2015;2015:CD008803.

4. Triolo G, Rabiolo A. Optical coherence tomogra-
phy and optical coherence tomography angiogra-
phy in glaucoma: diagnosis, progression, and cor-
relation with functional tests. \textit{Ther Adv Ophthal-

5. Rabiolo A, Carnevali A, Bandello F, Querques
G. Optical coherence tomography angiography:
evolution or revolution? \textit{Exp Rev Ophthalmol.}

Optical coherence tomography angiography mac-
ular and peripapillary vessel perfusion density in
healthy subjects, glaucoma suspects, and glaucoma
5722.

7. Chen HS, Liu CH, Wu WC, Tseng HJ, Lee YS.
Optical coherence tomography angiography of the
superficial microvasculature in the macular and
peripapillary areas in glaucomatous and healthy
32. Kim JS, Kim YK, Baek SU, et al. Topographic correlation between macular superficial microvascular density and ganglion cell-inner plexiform layer
Structural-Vascular Indices to Diagnose Glaucoma


