

New Method of Early RRMS Diagnosis Using OCT-Assessed Structural Retinal Data and Explainable Artificial Intelligence

Miguel Ortiz¹, Ana Pueyo^{2,3}, Francisco J. Dongil⁴, Luciano Boquete⁴, Eva M. Sánchez-Morla⁵⁻⁷, Rafael Barea⁴, Juan M. Miguel-Jimenez⁴, Almudena López-Dorado⁴, Elisa Vilades^{2,3}, María J. Rodrigo^{2,3}, Beatriz Cordon^{2,3}, and Elena Garcia-Martin^{2,3}

¹ School of Physics, University of Melbourne, Victoria, Australia

² Department of Ophthalmology, Miguel Servet University Hospital, Zaragoza, Spain

³ Aragon Institute for Health Research (IIS Aragon), Miguel Servet Ophthalmology Innovation and Research Group (GIMSO), University of Zaragoza, Zaragoza, Spain

⁴ Biomedical Engineering Group, Department of Electronics, University of Alcalá, Alcalá de Henares, Spain

⁵ Institute of Psychiatry and Mental Health, Gregorio Marañón University Hospital, IISGM, Madrid, Spain

⁶ Faculty of Medicine, Complutense University of Madrid, Madrid, Spain

⁷ CIBERSAM: Biomedical Research Networking Centre in Mental Health, Madrid, Spain

Correspondence: Elena Garcia-Martin, C/ Padre Arrupe. Servicio de Oftalmología, Edificio de Consultas Externas, Planta 1, Zaragoza 50009, Spain. e-mail: egmvivax@yahoo.com

Received: October 15, 2024

Accepted: January 3, 2025

Published: February 10, 2025

Keywords: optical coherence tomography (OCT); relapsing-remitting multiple sclerosis (RRMS); multiple sclerosis (MS); early diagnosis; recursive feature extraction; Shapley additive explanations (SHAP); posterior pole protocol

Citation: Ortiz M, Pueyo A, Dongil FJ, Boquete L, Sánchez-Morla EM, Barea R, Miguel-Jimenez JM, López-Dorado A, Vilades E, Rodrigo MJ, Cordon B, Garcia-Martin E. New method of early RRMS diagnosis using OCT-assessed structural retinal data and explainable artificial intelligence. *Transl Vis Sci Technol*. 2025;14(2):14, <https://doi.org/10.1167/tvst.14.2.14>

Purpose: The purpose of this study was to provide the development of a method to classify optical coherence tomography (OCT)-assessed retinal data in the context of automatic diagnosis of early-stage multiple sclerosis (MS) with decision explanation.

Methods: The database used contains recordings from 79 patients with recently diagnosed relapsing-remitting multiple sclerosis (RRMS) and no history of optic neuritis and from 69 age-matched healthy control subjects. Analysis was performed on the thicknesses (average right and left eye value and inter-eye difference) of the macular retinal nerve fiber layer (mRNFL), macular ganglion cell layer (mGCL), macular inner plexiform layer (mIPL), and macular inner retinal complex layer (mIRL), dividing the macular area into six analysis zones. Recursive feature extraction (RFE) and Shapley additive explanations (SHAP) are combined to rank relevant features and select the subset that maximizes classifier (support vector machine [SVM]) performance.

Results: Of the initial 48 features, 20 were identified as maximizing classifier accuracy ($n = 0.9257$). The SHAP values indicate that average thickness has greater relevance than inter-eye difference, that the mGCL and mRNFL are the most influential structures, and that the peripheral papillomacular bundle and the supero-temporal quadrant are the zones most affected.

Conclusions: This approach improves the success rate of automatic diagnosis of early-stage RRMS and enhances clinical decision making transparency.

Translational Relevance: Retinal structure assessment using OCT data could constitute a noninvasive means of diagnosing early-stage MS. This new high-accuracy and high-explainability method of analysis can be implemented in most hospitals and healthcare centers.



Introduction

Alterations in the retinal layer structure in patients with multiple sclerosis (MS) were assessed using histological analysis¹ and, noninvasively, using retinal optical coherence tomography (OCT) and angiography OCT (angio-OCT).² Systematic literature reviews and meta-analyses demonstrate that OCT detects thinning of the retinal nerve fiber layer (RNFL), the ganglion cell layer (GCL), and the inner retinal complex layer (IPL), even in the early stages of MS.³⁻⁵

The current version of the McDonald criteria⁶ does not include use of retinal OCT scans to confirm dissemination in space, even though OCT procedures are used in MS diagnosis in 38% of cases worldwide.⁷ Recent research indicates the potential diagnostic advantages of adding OCT assessment of the optic nerve to current diagnostic criteria as a fifth analysis region.^{8,9}

Previous papers have explored implementation of assisted MS diagnosis systems supported solely by artificial intelligence analysis of OCT recordings in the early stages of the disease.¹⁰⁻¹³

A key characteristic of OCT recordings is the high number of features they can obtain: depending on the acquisition equipment, it is possible to segment between 6 and 10 retinal structures. Moreover, for each structure, the technology can obtain information on either 40×60 points (Triton swept-source OCT) or 8×8 points (Spectralis OCT Fourier domain technology), among other options. Although literature review initially identifies the structures and zones most affected, it is of key importance to perform feature selection (FS) to optimize automatic classifier performance.

FS consists of selecting a subset of the initial features that maximizes the performance of a classifier. Recursive feature elimination (RFE) recursively discards the least relevant features. The initial features are trained and a ranking value is obtained based on their contribution to the classification. The least relevant features are then discarded and the procedure is repeated with the remaining ones until an optimal result is achieved. The first version of RFE, used to perform gene selection in cancer diagnosis, used a support vector machine (SVM) with a linear kernel¹⁴ as a classifier because it is possible to ascertain the relevance of each feature using the support vectors of this classifier.

Applied to diagnosis of early-stage relapsing-remitting multiple sclerosis (RRMS),¹³ it investigates the use of RFE with a linear-kernel SVM classifier to obtain the seven most relevant OCT features (3 relate to average thickness and 4 to inter-eye difference) and

confirms the correlation with the clinical status of the patients as assessed on the Expanded Disability Status Scale (EDSS). The classifier based on the top-ranked features achieved an accuracy of 0.88.

Another feature-ranking option is to use Shapley additive explanations (SHAP) to assign a relevance value for a particular prediction to each feature,¹⁵ thus distinguishing between relevant features and irrelevant ones. The SHAP method is based on the Shapley game theory values used to explain the output of any individual model. In medicine, SHAP values have been used to select features in references.^{16,17}

Given that there is room to improve the algorithms used in OCT-based diagnosis of early-stage RRMS, this paper aims to identify a subset of OCT features capable of effectively differentiating between healthy control subjects and patients. To enhance algorithm performance and generate insights into the input features' contribution to the prediction, it proposes ranking the features using recursive feature elimination and SHAP values.

Methods

This study was approved by the Clinical Research Ethics Committee of Aragon (Zaragoza, Spain) and conducted in accordance with the principles of the Declaration of Helsinki. All participants provided written informed consent after receiving a detailed explanation of the study. Data from these patients and controls were previously analyzed in references.^{12,13}

The eligibility criteria for the patient cohort included the following: (1) RRMS phenotype and disease duration < 26 months, because the aim is to improve the diagnosis of early-stage RRMS (the most common MS phenotype); and (2) no history of optic neuritis (ON) in either eye (to exclude easily diagnosed subjects). The healthy controls had no history of ocular or neurological disease and presented no signs or symptoms of them. For the study group, the criteria were ages 17 to -70 years and tolerance of high-quality OCT examination.

The exclusion criteria for both cohorts were as follows: (1) best-corrected visual acuity < 0.5 (Snellen charts); (2) refractive errors > 5 diopters of spherical equivalent refraction or 3 diopters of astigmatism; (3) intraocular pressure > 20 millimeters of mercury (mm Hg); (4) media opacification (nuclear color/opalescence or cortical or posterior subcapsular lens opacity < 1 according to the Lens Opacities Classification System III) to exclude dense cataracts that can affect OCT

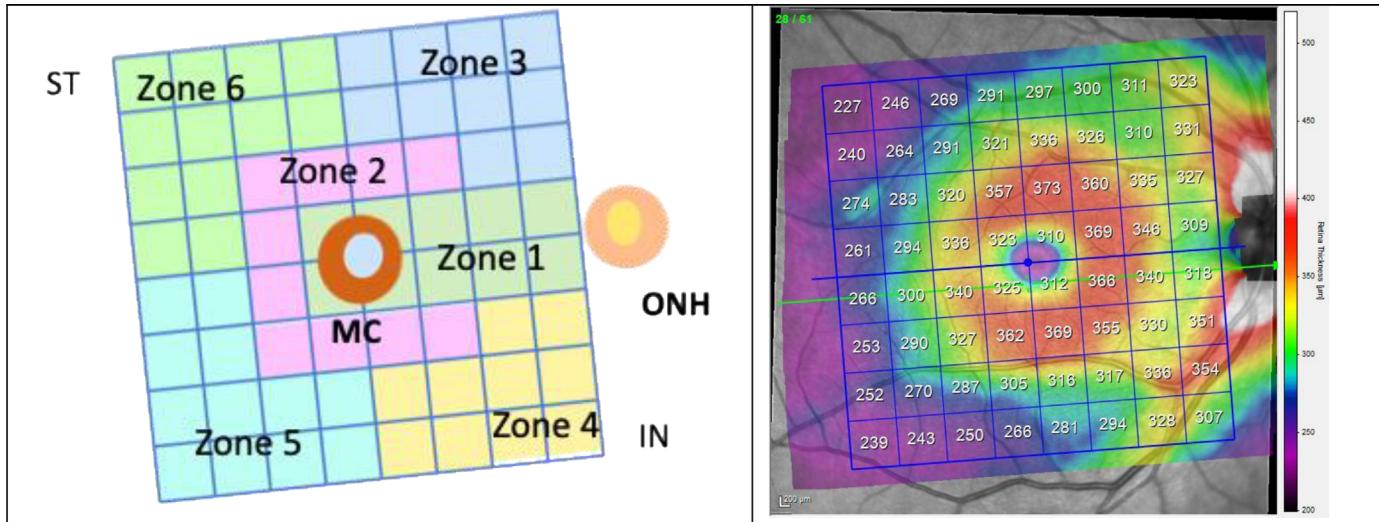


Figure 1. Left: Definition of the analysis zones on the posterior pole grid of the macular area of the right eye using the Spectralis optical coherence tomography (OCT) protocol. The posterior pole analysis protocol uses the Anatomic Positioning System (APS) to adjust the scans of the axis fovea to the center of Bruch's membrane. Right: Spectralis OCT recording showing macular thicknesses on the 8×8 grid that corresponds to the analysis zones defined in the figure on the left. ONH, optic nerve head; MC, macula center (fovea); IN, inferior nasal; ST, superior temporal.

thickness measurements (not common in patients with the mean age of this cohort); (5) concomitant ocular disease (including glaucoma or retinal pathology); and (6) other systemic conditions potentially affecting the visual system.

OCT Method

The OCT recordings were made as per the APOSTEL recommendations.¹⁸ OCT measurements of both eyes were taken with a spectral-domain (SD) OCT (Spectralis, Heidelberg Engineering, Heidelberg, Germany) device using the posterior pole asymmetry analysis protocol.¹⁹ The protocol obtained the thicknesses of nine retinal structures. This study uses the four retinal structures that exhibit the greatest capacity to discriminate between control eyes and those of patients with RRMS: mRNFL, mGCL, mIPL, and macular inner retinal layer (mIRL) complex (mIRL = mRNFL + mGCL + mIPL).

The 25×30 explored area centered on the foveola was represented as an 8×8 grid that reflects overall retinal thickness. Each layer was segmented using the inbuilt software (HRA version 6.0.7.0). Scan quality was analyzed according to the OSCAR-IB Consensus Criteria for Retinal Quality Assessment.²⁰ No manual correction was applied to the OCT output and only high-quality scans (25/40) were used for analysis.

The 8×8 grid was divided into 6 analysis zones: zone 1 = the central papillomacular bundle, zone 2 = the peripheral papillomacular bundle, zone 3 =

the supero-nasal quadrant, zone 4 = the infero-nasal quadrant, zone 5 = the infero-temporal quadrant, and zone 6 = the supero-temporal quadrant (see Fig. 1).

The average thickness values (AVG) of the two eyes of each subject and the difference in thickness between the right eye and the left eye (DIFF) were both analyzed. The initial number of features per subject was as follows: 6 zones \times 2 measurements \times 4 structures = 48 features.

Feature Selection and Classification

Automatic classifier performance deteriorates if the classifier uses a large number of inputs that do not contribute relevant information and may be correlated. There is no single broad solution to the feature selection problem. In this paper, RFE is used to create a SHAP ranking that assigns each feature a relevance value for a particular prediction¹⁵ (Fig. 2).

The details of this procedure are as follows:

- The initial number of features is $N_{RFE} = 48 = N$.
- The classifier is trained using this number of features and its confusion matrix is obtained using the leave-one-out (LOO) method and the degree of accuracy. The SHAP values, which assign a relevance value for this prediction to each input, are obtained.
- The least relevant feature ($N = N - 1$) is discarded and the process is repeated from step b onward until $N = 1$.

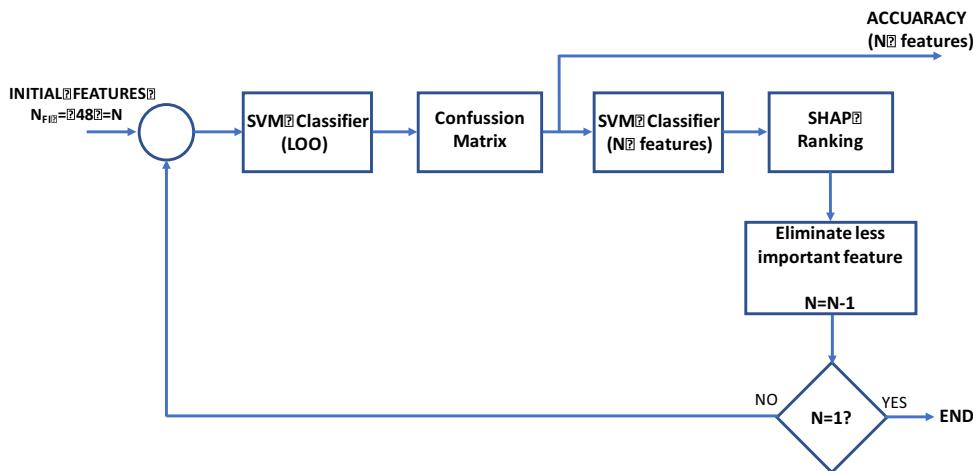


Figure 2. Diagram of the RFE-SHAP method. LOO, leave-one-out; SVM, support vector machine.

This process can be plotted (accuracy versus N) and the subset of features ($N_{RELEVANT}$) that obtains the greatest accuracy value is selected.

A linear-kernel SVM classifier ($C = 1.0$ [regularization parameter], penalty = L2 and loss function = squared hinge) is used.

Results

Study Cohort and OCT Protocol

Table 1 shows the demographic and clinical characteristics of the two cohorts. The RRMS group comprises 68 female subjects and 11 male subjects, whereas the healthy group comprises 46 female subjects and 23 male subjects.

Table 1. Demographic and Clinical Characteristics of the Patients With Multiple Sclerosis and Healthy Controls

	RRMS Patients (n = 79 Patients)	Healthy Controls (n = 69 Subjects)	P Value
Age, y, mean \pm SD	45.64 \pm 13.59	46.94 \pm 12.64	0.604
Visual acuity, Snellen	0.83 \pm 0.57	0.91 \pm 0.36	0.026
Disease duration since definitive diagnosis, y, mean \pm SD	1.42 \pm 0.72	—	—
EDSS score, median (range)	1.28 (1–3)	—	—
Treatment	Avonex: 3 Betaseron: 7 Rebif: 6 Capaxone: 6 Tecfidera: 11 Gilenya: 32 Mayzent: 3 Aubagio: 9 No treatment: 2	—	—

EDSS, Expanded Disability Status Scale; RRMS, relapsing-remitting multiple sclerosis; SD, standard deviation.

Figure 3 shows the results obtained using the RFE-SHAP method. Model performance peaked (accuracy = 0.9257) at 21 and 20 features. The analysis below is performed with the $N_{RELEVANT} = 20$ features that make the most significant contribution to the classification.

Figure 4 shows the ranked relevance of each classifier input variable. The SHAP values vary throughout the range (0.1201–0.0178) and decrease progressively, meaning that a threshold value that identifies a group of features that are clearly more relevant than others cannot be determined. Of the 20 features that influence the classifier, 12 correspond to average thickness values (aggregate SHAP values = 0.7753) and 8 to the inter-eye difference (aggregate value = 0.3838). In addition, seven entries in the top half of the ranking are average thickness values and three are inter-eye difference values. Therefore, average thickness has a

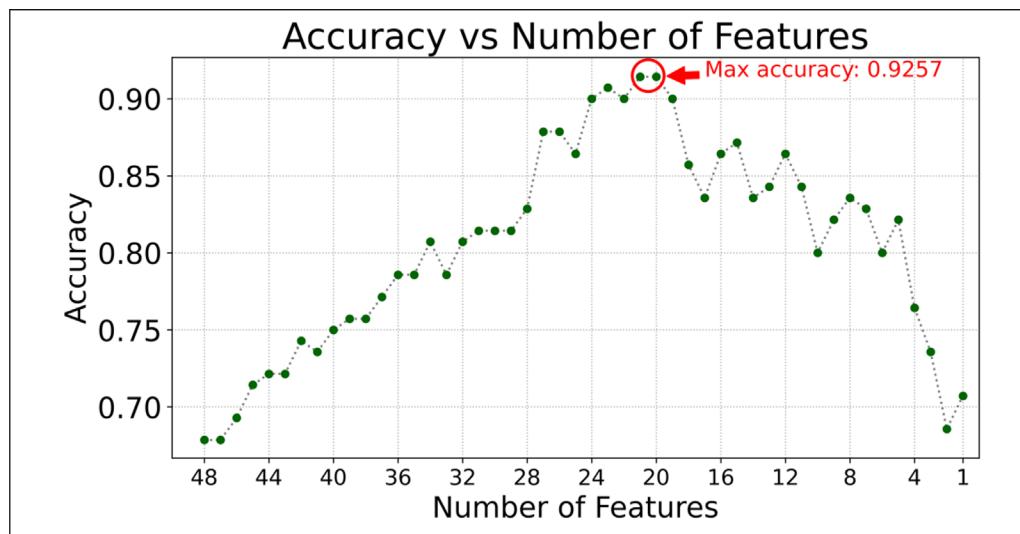


Figure 3. RFE–SHAP algorithm performance.

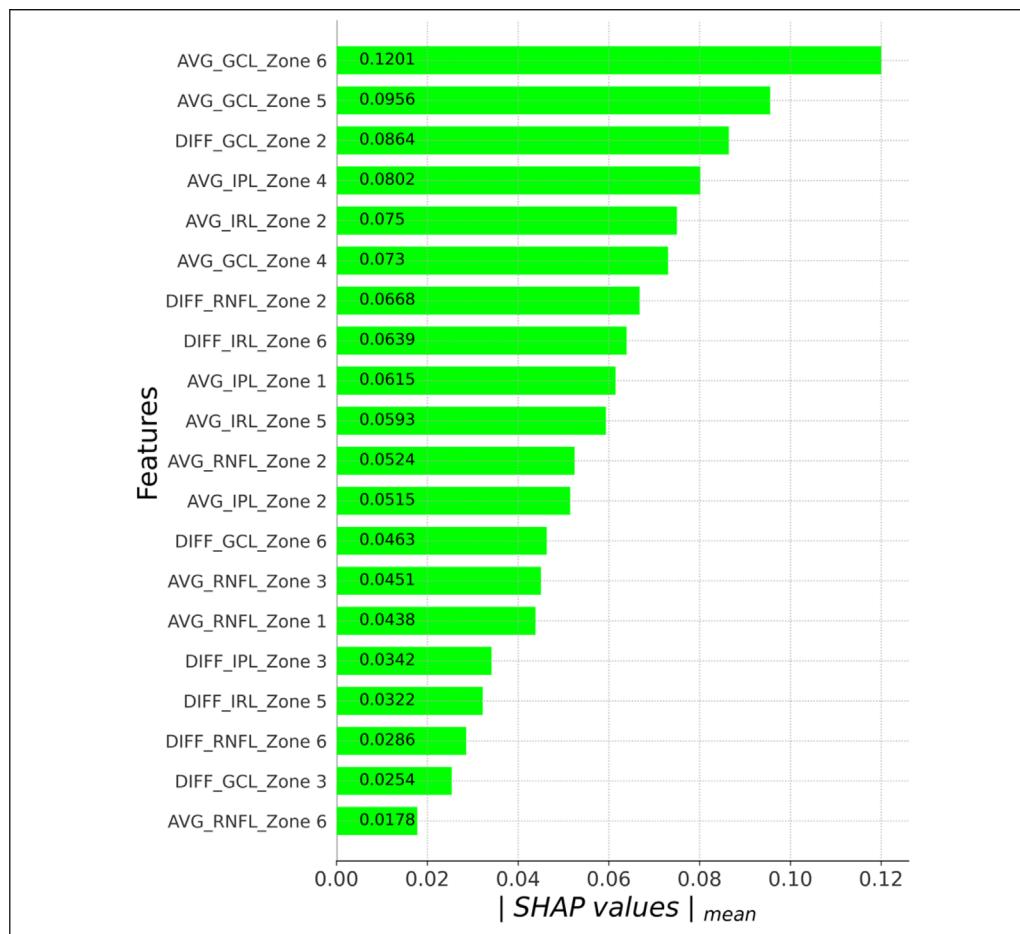


Figure 4. Ranking of relevance of the selected features as per absolute mean SHAP value.

Table 2. Classifier Metrics for the 20 Most Relevant Features

Metric	Value
Accuracy	0.9257
Sensitivity	0.9747
Specificity	0.8696
F1-score	0.9333

greater influence on the classification than inter-eye difference.

In Figure 4, the retinal structures most commonly identified are the GCL ($n_{GCL} = 6$; the top 3 values correspond to this layer and the aggregate value for the entire GCL is 0.4468), the RNFL ($n_{RNFL} = 6$; aggregate value = 0.2545), the IPL ($n_{IPL} = 4$; aggregate value = 0.2274), and the IRL ($n_{IRL} = 4$; aggregate value = 0.2304). Consequently, the structures that most condition the classifier decision are, in order of relevance, the GCL, RNFL, IPL, and IRL.

The results of the SHAP ranking for the zones are as follows: zone 2 ($n_{ZONE2} = 5$; aggregate value = 0.3321), zone 6 ($n_{ZONE6} = 5$; aggregate value = 0.2767), zone 5 ($n_{ZONE5} = 3$; aggregate value = 0.1871), zone 4 ($n_{ZONE4} = 2$ aggregate value = 0.1532), zone 1 ($n_{ZONE1} = 2$; aggregate value = 0.1053), and zone 3 ($n_{ZONE3} = 3$; aggregate value = 0.1047). It can be deduced that the zones that most influence the classifier are zone 2 (peripheral papillomacular bundle) and zone 6 (superotemporal quadrant).

Training the classifier with the 20 features of interest and using the LOO validation method obtains the metrics presented in Table 2.

Discussion

This paper implements a recursive OCT feature selection method based on SHAP ranking that, using only retinal structure data, enhances early RRMS diagnosis and also explains the decisions made by the classifier.

The main outcome of this approach is that it improves on prior methods of automatic RRMS diagnosis^{12,13} and provides insights into how RRMS affects the neuroretina in the initial stages of the disease in patients without ON.

A previous approach that applied a convolutional neural network to the same database obtained an accuracy value of 0.87¹² but did not perform feature selection or analyze the explainability of the results.

The classic RFE-SVM method identifies 7 relevant variables and achieves an accuracy of 0.88.¹³ In

contrast, RFE-SHAP achieves an accuracy of 0.9257 and identifies 20 relevant features. The difference in the results produced by the two RFE methods can be explained by the fact that they are built on different mathematical foundations: in the first case, relevance is assessed using support vector distance, whereas, in the second method, assessment is based on the SHAP value, which analyzes the marginal value each feature contributes to the classifier output.

Most studies using OCT to analyze retinal structure in early-stage RRMS provide evidence of the effect on the average thickness of certain retinal structures, including in patients with no history of ON.^{12,13,21} Inter-eye difference has recently gained importance in RRMS diagnosis in both advanced^{22,23} and early stages of the disease.^{12,13}

Among the 20 features considered relevant, greatest relevance is assigned to values that assess the average of the left and right eye thicknesses in each subject. It is also evident that inter-eye difference contributes to a lesser degree to differentiation between patients with RRMS.

This paper analyses data from the four structures that have the greatest capacity to differentiate patients with RRMS evaluated using the area under the curve (AUC): mRNFL, mGCL, mIPL, and mIRL. Our findings are in line with those obtained in reference²⁴, which analyzed patients with RRMS without ON and with a disease duration of 23 ± 13 months, observing that mRNFL, mGCL, mIPL, and mGCIPL thickness (the authors did not analyze IRL) were significantly reduced. Martucci et al.²⁵ use the posterior pole protocol to analyze patients (97 eyes without ON and a disease duration of 43.2 ± 66.7 months versus 106 control eyes). The findings indicate that there is a significant difference in GCL thickness ($P < 0.001$) and, although it is not statistically significant, the INL ($P = 0.072$) and RNFL ($P = 0.074$) are thinner.

Most previous papers observe that the greatest discriminant capacity is found in the GCL or GCIPL complex (GCL+IPL), generally followed by the RNFL. The findings of this study support this observation because, according to the SHAP values, the GCL plays a greater role in classifier decision making than the RNFL. This may be due to the fact that significant thinning of the GCL appears in the first month from symptom onset, whereas RNFL thinning is not detected until at least 3 months after symptom onset.²⁶ GCL is made up of ganglion cell somas, which decrease in size in neurodegenerative pathologies due to the anterograde atrophy caused by the axonal damage produced in the central nervous system (third neuron of the visual pathway) and which advances to the second neuron of the visual pathway

(which is the ganglion cell; the first neuron is the retinal bipolar cell). Another complementary explanation may be that segmentation of the RNFL often depends on certain ocular conditions (e.g. edema or inflammation), which makes the values in this layer less consistent when compared with the greater reproducibility of the GCL measurements.

According to the feature ranking obtained using the SHAP method, the areas with greatest influence on the classifier are the peripheral papillomacular bundle and the supero-temporal quadrant. Few papers identify defect patterns in retina structures in patients with early-stage RRMS and no history of ON. There are several ways of identifying zonal differences between patients and control subjects: statistical methods, Cohen distance, AUC, the SHAP method, etc.

The study by Albano et al. analyzes 34 eyes of patients with RRMS without ON (disease duration: 0.3 [0.05–1.98] years) and 34 control eyes and uses a statistical method to identify defects in the inner retinal layer complex (macular ganglion cell–inner plexiform layer complex).²⁷ In line with our findings, the study reports that the most frequently occurring defect presents in the proximity of the peripheral papillomacular bundle.

Meanwhile, using the Cohen distance and SD-OCT recordings, Garcia-Martin et al.¹⁰ and López-Dorado et al.¹¹ detect clear thinning of the perimacular area. This thinning principally affects the papillomacular bundle in the GCL+ (GCL+IPL) and GCL++ (RNFL+GCL+IPL) whereas little alteration is detected in the RNFL. Ortiz et al. uses the area under the receiver operating characteristic curve (AUROC) and observes that the defects mainly present in the GCL, IPL, and IRL, with the greatest thinning occurring in the perimacular areas.¹²

Also in line with our findings, using RFE with a random forest classifier identifies a set of eight retinal features in children (<18 years old) with MS (disease duration: 0.6 ± 1.5 years).²⁸ An accuracy value of 0.80 is obtained in diagnosis of MS versus control subjects and it is observed that ganglion cell–inner plexiform layer thickness in the supero-temporal sector and retinal nerve fiber layer thickness in the temporal quadrant are among the features with greatest capacity to distinguish between patients with MS and healthy control subjects.

This study confirms the utility of the inner layers of the neuroretina in reliably diagnosing early-stage RRMS. It also serves to confirm the trend of previous studies in which artificial intelligence (AI) is used to optimize OCT-based diagnosis of this pathology. It likewise improves on previous outcomes and focuses on macular regions, which, according to the latest studies,

are the ones that provide the most relevant information about this pathology. Furthermore, because this method relies less on inter-eye difference it may be more useful in subjects who have not presented ON in either eye, who are in the early stages of the disease, or in whom only one eye is assessable because the other eye has a history of trauma, anisometropia, lazy eye, etc.

A cross-sectional study with larger numbers of subjects in both groups would give the findings greater weight. In addition, because only patients with the relapsing-remitting phenotype were evaluated (as this is the most common phenotype of this disease), similar studies would be necessary for other MS phenotypes. Inclusion of other clinical characteristics in addition to ocular characteristics could further improve the performance of the algorithm. In future studies, incorporation of Angio OCT using new software designed by Heidelberg would provide a wealth of 3D information, including analysis of retinal irrigation and vascularization.

In conclusion, this approach improves both the success rate of automatic diagnosis and the transparency of the decision, thus enhancing the potential clinical utility and acceptance of OCT analysis in RRMS diagnosis. Building on previous papers^{8,9,29} it further contributes to the evidence supporting the inclusion of OCT-assessed optic nerve topography as a fifth analysis region capable of meeting the dissemination in space condition necessary when diagnosing MS according to the McDonald criteria.

Acknowledgments

The authors thank Luis E. Pablo for his help with use of the devices.

Supported by Carlos III Health Institute grants PI23/00935, the Inflammatory Disease Network (RICORS) (RD21/0002/0050 and RD24-0007-0022) (Carlos III Health Institute), and by the Government of Aragon through group B23_23R and PROY_B50_24.

Data Availability: The data collected and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Disclosure: **M. Ortiz**, None; **A. Pueyo**, None; **F.J. Dongil**, None; **L. Boquete**, None; **E.M. Sánchez-Morla**, None; **R. Barea**, None; **J.M. Miguel-Jimenez**, None; **A. López-Dorado**, None; **E. Vilades**, None; **M.J. Rodrigo**, None; **B. Cordon**, None; **E. García-Martin**, None

References

1. Green AJ, McQuaid S, Hauser SL, Allen I V, Lyness R. Ocular pathology in multiple sclerosis: retinal atrophy and inflammation irrespective of disease duration. *Brain*. 2010;133(6):1591–1601.
2. Cordon B, Vilades E, Orduna E, et al. Angiography with optical coherence tomography as a biomarker in multiple sclerosis. *PLoS One*. 2020;15(12):e0243236.
3. Mirmosayeb O, Zivadinov R, Weinstock-Guttman B, Benedict RHB, Jakimovski D. Optical coherence tomography (OCT) measurements and cognitive performance in multiple sclerosis: a systematic review and meta-analysis. *J Neurol*. 2023;270(3):1266–1285.
4. Petzold A, Balcer LJ, Calabresi PA, et al. Retinal layer segmentation in multiple sclerosis: a systematic review and meta-analysis. *Lancet Neurol*. 2017;16(10):797–812.
5. Cujba L, Stan C, Samoila O, Drugan T, Benedec Cutas A, Nicula C. Identifying optical coherence tomography markers for multiple sclerosis diagnosis and management. *Diagnostics (Basel)*. 2023;13(12):2077.
6. Thompson AJ, Banwell BL, Barkhof F, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol*. 2018;17(2):162–173.
7. Solomon AJ, Marrie RA, Viswanathan S, et al. Global barriers to the diagnosis of multiple sclerosis. *Neurology*. 2023;101(6):e624–e635.
8. Bsteh G, Hegen H, Altmann P, et al. Diagnostic performance of adding the optic nerve region assessed by optical coherence tomography to the diagnostic criteria for multiple sclerosis. *Neurology*. 2023;101(8):e784–e793.
9. Vidal-Jordana A, Sastre-Garriga J, Tintoré M, Rovira À, Montalban X. Optic nerve topography in multiple sclerosis diagnostic criteria: existing knowledge and future directions. *Mult Scler*. 2024;30:139–149.
10. Garcia-Martin E, Ortiz M, Boquete L, et al. Early diagnosis of multiple sclerosis by OCT analysis using Cohen's d method and a neural network as classifier. *Comput Biol Med*. 2021;129:104165.
11. López-Dorado A, Ortiz M, Satue M, et al. Early diagnosis of multiple sclerosis using swept-source optical coherence tomography and convolutional neural networks trained with data augmentation. *Sensors (Basel)*. 2021;22(1):167.
12. Ortiz M, Mallen V, Boquete L, et al. Diagnosis of multiple sclerosis using optical coherence tomography supported by artificial intelligence. *Mult Scler Relat Disord*. 2023;74:104725.
13. Dongil-Moreno FJ, Ortiz M, Pueyo A, et al. Diagnosis of multiple sclerosis using optical coherence tomography supported by explainable artificial intelligence. *Eye (Lond)*. 2024;38(8):1502–1508.
14. Guyon I, Weston J, Barnhill S, Vapnik V. Gene selection for cancer classification using support vector machines. *Mach Learn*. 2002;46(1/3):389–422.
15. Lundberg S, Lee S-I. A unified approach to interpreting model predictions. In: *NIPS '17: Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS2017)*. Long Beach, CA, USA. 2017:4768–4777.
16. Bi Y, Xiang D, Ge Z, et al. An interpretable prediction model for identifying N7-Methylguanosine sites based on XGBoost and SHAP. *Mol Ther Nucleic Acids*. 2020;22:362–372.
17. Liu Y, Liu Z, Luo X, Zhao H. Diagnosis of Parkinson's disease based on SHAP value feature selection. *Biocybern Biomed Eng*. 2022;42(3):856–869.
18. Cruz-Herranz A, Balk LJ, Oberwahrenbrock T, et al. The APOSTEL recommendations for reporting quantitative optical coherence tomography studies. *Neurology*. 2016;86(24):2303–2309.
19. Asrani S, Rosdahl JA, Allingham RR. Novel software strategy for glaucoma diagnosis: asymmetry analysis of retinal thickness. *Arch Ophthalmol*. 2011;129(9):1205–1211.
20. Tewarie P, Balk L, Costello F, et al. The OSCAR-IB consensus criteria for retinal OCT quality assessment. *PLoS One*. 2012;7(4):e34823.
21. Paul F, Calabresi PA, Barkhof F, et al. Optical coherence tomography in multiple sclerosis: a 3-year prospective multicenter study. *Ann Clin Transl Neurol*. 2021;8(12):2235–2251.
22. Nij Bijvank J, Uitdehaag BMJ, Petzold A. Retinal inter-eye difference and atrophy progression in multiple sclerosis diagnostics. *J Neurol Neurosurg Psychiatry*. 2022;93(2):216–219.
23. Nolan-Kenney RC, Liu M, Akhand O, et al. Optimal intereye difference thresholds by optical coherence tomography in multiple sclerosis: an international study. *Ann Neurol*. 2019;85(5):618–629.
24. Pietroboni AM, Carandini T, Dell'Arti L, et al. Evidence of retinal anterograde neurodegeneration in the very early stages of multiple sclerosis: a longitudinal OCT study. *Neurol Sci*. 2020;41(11):3175–3183.
25. Martucci A, Landi D, Cesareo M, et al. Complex rearrangement of the entire retinal posterior pole

in patients with relapsing remitting multiple sclerosis. *J Clin Med.* 2021;10(20):4693.

26. Graham EC, You Y, Yiannikas C, et al. Progressive loss of retinal ganglion cells and axons in nonoptic neuritis eyes in multiple sclerosis: a longitudinal optical coherence tomography study. *Invest Ophthalmol Vis Sci.* 2016;57(4):2311.

27. Albano V, Dammacco R, Manni A, et al. Macular ganglion cell-inner plexiform layer defect patterns in multiple sclerosis patients without optic neuritis: a spectral-domain-optical coherence tomography cross-sectional, case-control, pilot study. *Eur J Ophthalmol.* 2023;33(1):546–555.

28. Ciftci Kavaklıoglu B, Erdman L, Goldenberg A, et al. Machine learning classification of multiple sclerosis in children using optical coherence tomography. *Mult Scler.* 2022;28(14):2253–2262.

29. Vidal-Jordana A, Rovira A, Calderon W, et al. Adding the optic nerve in multiple sclerosis diagnostic criteria: a longitudinal, prospective, multi-center study. *Neurology.* 2024;102(1):e200805.