RESEARCH ARTICLE OPEN ACCESS

Deep Learning Modeling to Differentiate Multiple Sclerosis From MOG Antibody–Associated Disease

Rosa Cortese,[1] Francesco Sforazzini,[2] Giordano Gentile,[1,2] Anna de Mauro,[1] Ludovico Luchetti,[1,2] Maria Pia Amato[3,4] et al.

Correspondence
Dr. Cortese
rosa.cortese@unisi.it

Neurology® 2025;105:e214075. doi:10.1212/WNL.0000000000214075

Abstract

Background and Objectives

Multiple sclerosis (MS) is common in adults while myelin oligodendrocyte glycoprotein antibody–associated disease (MOGAD) is rare. Our previous machine-learning algorithm, using clinical variables, ≤6 brain lesions, and no Dawson fingers, achieved 79% accuracy, 78% sensitivity, and 80% specificity in distinguishing MOGAD from MS but lacked validation. The aim of this study was to (1) evaluate the clinical/MRI algorithm for distinguishing MS from MOGAD, (2) develop a deep learning (DL) model, (3) assess the benefit of combining both, and (4) identify key differentiators using probability attention maps (PAMs).

Methods

This multicenter, retrospective, cross-sectional MAGNIMS study included scans from 19 centers. Inclusion criteria were as follows: adults with non-acute MS and MOGAD, with high-quality T2-fluid-attenuated inversion recovery and T1-weighted scans. Brain scans were scored by 2 readers to assess the performance of the clinical/MRI algorithm on the validation data set. A DL-based classifier using a ResNet-10 convolutional neural network was developed and tested on an independent validation data set. PAMs were generated by averaging correctly classified attention maps from both groups, identifying key differentiating regions.

Results

We included 406 MRI scans (218 with relapsing remitting MS [RRMS], mean age: 39 years ±11, 69% F; 188 with MOGAD, age: 41 years ±14, 61% F), split into 2 data sets: a training/testing set (n = 265: 150 with RRMS, age: 39 years ±10, 72% F; 115 with MOGAD, age: 42 years ±13, 61% F) and an independent validation set (n = 141: 68 with RRMS, age: 40 years ±14, 65% F; 73 with MOGAD, age: 40 years ±15, 63% F). The clinical/MRI algorithm predicted RRMS over MOGAD with 75% accuracy (95% CI 67–82), 96% sensitivity (95% CI 88–99), and specificity 56% (95% CI 44–68) in the validation cohort. The DL model achieved 77% accuracy (95% CI 64–89), 73% sensitivity (95% CI 57–89), and 83% specificity (95% CI 65–96) in the training/testing cohort, and 70% accuracy (95% CI 63–77), 67% sensitivity (95% CI 55–79), and 73% specificity (95% CI 61–83) in the validation cohort without retraining. When combined, the classifiers reached 86% accuracy (95% CI 81–92), 84% sensitivity (95% CI 75–92), and 89% specificity (95% CI 81–96). PAMs identified key region volumes: corpus callosum (1872 mm^3), left precentral gyrus (341 mm^3), right thalamus (193 mm^3), and right cingulate cortex (186 mm^3) for identifying RRMS and brainstem (629 mm^3), hippocampus (234 mm^3), and parahippocampal gyrus (147 mm^3) for identifying MOGAD.

Discussion

Both classifiers effectively distinguished RRMS from MOGAD. The clinical/MRI model showed higher sensitivity while the DL model offered higher specificity, suggesting complementary roles. Their combination improved diagnostic accuracy, and PAMs revealed distinct damage patterns. Future prospective studies should validate these models in diverse, real-world settings.

Classification of Evidence

This study provides Class III evidence that both a clinical/MRI algorithm and an MRI-based DL model accurately distinguish RRMS from MOGAD.

MORE ONLINE

Class of Evidence
Criteria for rating therapeutic and diagnostic studies
NPub.org/coe

The Author Byline is continued at the end of the article.

Author affiliations appear at the end of the article.

The Article Processing Charge was funded by the authors.

Glossary

ADEM = acute disseminated encephalomyelitis; **AUC** = area under the curve; **CBA** = cell-based assay; **CNN** = convolutional neural network; **DL** = deep learning; **FLAIR** = fluid-attenuated inversion recovery; **FN** = false negative; **FP** = false positive; **MAGNIMS** = Magnetic Resonance Imaging in Multiple Sclerosis; **MNI** = Montreal Neurological Institute; **MOG** = myelin oligodendrocyte glycoprotein; **MOG-Ab** = MOG antibody; **MOGAD** = MOG antibody–associated disease; **MS** = multiple sclerosis; **NMOSD** = neuromyelitis optica spectrum disorder; **NPV** = negative predictive value; **OCB** = oligoclonal band; **PAM** = probability attention map; **PPV** = positive predictive value; **ROC** = receiver operating characteristic; **RRMS** = relapsing remitting MS; **TN** = true negative; **TP** = true positive.

Introduction

Myelin oligodendrocyte glycoprotein (MOG) antibody–associated disease (MOGAD) is distinct from multiple sclerosis (MS),[1] but their clinical and radiologic differentiation remains challenging. First, early clinical presentations may overlap, despite differences in pathogenesis, prognosis, and treatment.[1,2] Second, while MRI has a crucial role in differentiating the 2 disorders,[3] MOGAD lesions evolve dynamically, forming during attacks and resolving in remission, complicating marker identification.[4,5] Third, although MOG antibodies (MOG-Abs) can be detected through cell-based assays (CBAs), titers may fluctuate and MOG-Abs occasionally can be found in patients diagnosed with MS (2% with fixed CBAs and approximately 0.5% with live CBAs).[6,7] Given the low relative prevalence of MOGAD compared with MS, indiscriminate testing for MOG-Abs in large, unselected populations inevitably leads to false-positive results.[8,9] Finally, there are numerous unresolved questions regarding MOGAD pathogenesis.[10] Addressing these challenges is essential for improving diagnostic accuracy.

Artificial intelligence–based algorithms applied to clinical MRI have shown promise in classifying neurologic diseases.[11] In a recent MAGNIMS study, we used machine learning on a large cohort to identify markers differentiating MOGAD from MS using conventional scans. A combination of fewer brain lesions (≤6) and absence of Dawson fingers distinguished 162 MOGAD cases from 189 relapsing remitting MS (RRMS) cases with 79% accuracy, 78% sensitivity, and 80% specificity. Sensitivity improved when the first clinical presentation included bilateral optic neuritis, optic neuritis with transverse myelitis, or acute disseminated encephalomyelitis (ADEM).[12] Validation on new cohorts is still needed to confirm the model's replicability.

The integration of MRI with deep learning (DL) has gained traction for its ability to tackle complex classification challenges by automatically extracting key features from raw data. In MS, DL, particularly convolutional neural networks (CNNs), has successfully classified patients by disability levels and predicted progression.[13-15] While DL shows promise in distinguishing CNS demyelinating diseases such as MS and neuromyelitis optica spectrum disorder (NMOSD),[16,17] its application to differentiate MS from MOGAD remains underexplored. In addition, explainable DL techniques, such as attention maps, could provide deeper insights into disease characteristics, offering a more nuanced understanding of damage distribution in the CNS.[18,19]

The aim of this study was to evaluate the effectiveness of the clinical/MRI algorithm and a DL-based model in distinguishing MS from MOGAD, both independently and in combination, in 3 steps. First, we performed a validation of the previously proposed clinical/MRI algorithm in a new non-acute cohort, which was not part of the previous study. Next, we developed an MRI-based DL model to distinguish MS from MOGAD and explored the underlying factors behind the differentiation using probability attention maps (PAMs). Finally, we combined both approaches to determine whether this integration would enhance diagnostic accuracy.

The primary research question of this study is whether a clinical/MRI algorithm and a DL model, individually or in combination, can accurately distinguish RRMS from MOGAD and identify key differentiating brain regions.

Methods

Study Design and Population

This multicenter, retrospective study used data from 19 European and non-European centers, collected as part of the Magnetic Resonance Imaging in Multiple Sclerosis (MAGNIMS) collaboration[20] project. The study included scans of patients who met the following inclusion criteria: diagnosis of RRMS or MOGAD according to the respective diagnostic criteria[2,21]; age at MRI 18 years or older; being ≥6 months after an acute event; availability of ≥1 sequence for white matter lesion detection (T2-fluid-attenuated inversion recovery [FLAIR]/proton density/T2-weighted) and a 3D T1-weighted sequence of adequate quality. The scans were divided into 2 independent data sets:

1. Training/testing set: this set comprised patients from the original MAGNIMS cohort where a clinical/MRI classifier was previously developed.[12] This set was further split into 2 independent subsets, 80% for training and 20% for testing, ensuring that centers in the training set were not included in the testing set. The training and testing data sets were split randomly, ensuring that both sets reflected the same ratio of patients with MS and MOGAD by incorporating the outcome data during stratification.

2. Validation test set: this set was separately collected for the purpose of this study, with data provided by both previous and new centers that contributed scans from newly enrolled patients.

Although no formal sample size calculation was performed, the data set was considered sufficiently large based on previous DL studies in similar settings and the rarity of MOGAD[16,17]; its adequacy was further supported by the model's robust performance on an independent validation set.

Standard Protocol Approvals, Registrations, and Patient Consents

Each participant provided written consent for research within each center. The final protocol for analyzing fully anonymized scans, acquired independently at each center, was approved by the MAGNIMS collaboration and by local ethics committees. Scans were shared according to the MAGNIMS framework agreement, and centers not covered by this agreement signed a separate data transfer agreement with Siena.

MRI Analysis

MRI protocols and acquisition parameters of the centers were previously described.[12] White matter lesions were semi-automatically segmented on FLAIR/T2 sequences using a lesion prediction algorithm implemented in the lesion segmentation tool (LST) version for statistical parametric mapping version 12 (SPM12), and lesion masks were created. The quality of lesions was manually checked and corrected by 2 readers (R.C. and A.d.M.) through consensus. The MRI analysis was centralized in Siena.

The image-processing pipeline consisted of 4 steps (Figure 1).

Brain Scan Scoring and Clinical/MRI Classification

The number of lesions for each patient was automatically calculated. FLAIR/T2 sequences of the scans from the validation test set were scored by 2 readers (R.C. and A.d.M.) for the presence of Dawson fingers, which were previously reported to be useful differentiators of RRMS and MOGAD.[12] Scans were classified as indicative of RRMS if they exhibited >6 brain lesions and ≥1 Dawson finger. Scans not meeting these criteria were categorized as MOGAD. Then, clinical presentation at onset was added to the classifier, and in particular, when the initial episode was either unilateral optic neuritis, or isolated transverse myelitis, or other symptoms related to brain involvement, the scans were considered indicative of RRMS.

Development of DL-Based Classifier

The DL classifier was developed using a CNN with a 10-layer Residual Network (ResNet-10) architecture.[22] This ResNet-10 classifier had the following stacked layers: 1 convolutional layer (64 convolutional filters, 3 × 3 with stride 2 and padding of 3, and output with shape 64 × 34 × 34), followed by a max pooling layer (64 max pooling filters, 3 × 3 with stride 2 and padding 1, and output with shape 64 × 17 × 17); 4 residual layers (the output after these 4 residual layers has the shape of 512 × 3 × 3); 1 average pooling layer (512 average pooling filters, 2 × 2 with stride 1 and 0 padding, and output is a vector of 512 neurons); and 1 fully connected layer to generate predictions of the 2 classes. The output of each convolutional layer was processed by batch normalization[23] to prevent vanishing/exploding gradient problem. FLAIR and T1 images were used as input channels. Before feeding to the network, both images underwent the following preprocessing: (1) correction for magnetic field inhomogeneities using N420 and normalization of their intensities to a common range (0–10,000); (2) resampling to have an isotropic resolution of 1 mm, nonlinearly registering to the Montreal Neurological Institute (MNI) space, and brain extraction; (3) final crop to remove nonbrain voxels. The input shape of the network was set to 182 × 145 × 155 voxels. The batch size was set to 10, and the network was trained for 500 epochs using an initial learning rate of 0.0001, with polynomial decay, stochastic gradient descent (SGD) as the optimizer, and a dropout of 0.7 to avoid overfitting. To further mitigate overfitting, we applied data augmentation techniques in addition to using a clinically balanced and sufficiently large data set, which collectively contributed to the model's stable performance on unseen data. Cross-entropy was used as a loss function. The images coming from the patients during the evaluation phase underwent the same preprocessing and were then fed to the trained network, without any recalibration or retraining of the model. To evaluate the performance of the DL-based classifier, we used balanced accuracy in correctly classifying each patient by their diagnosis and sensitivity and specificity of correctly classified patients with RRMS vs MOGAD.

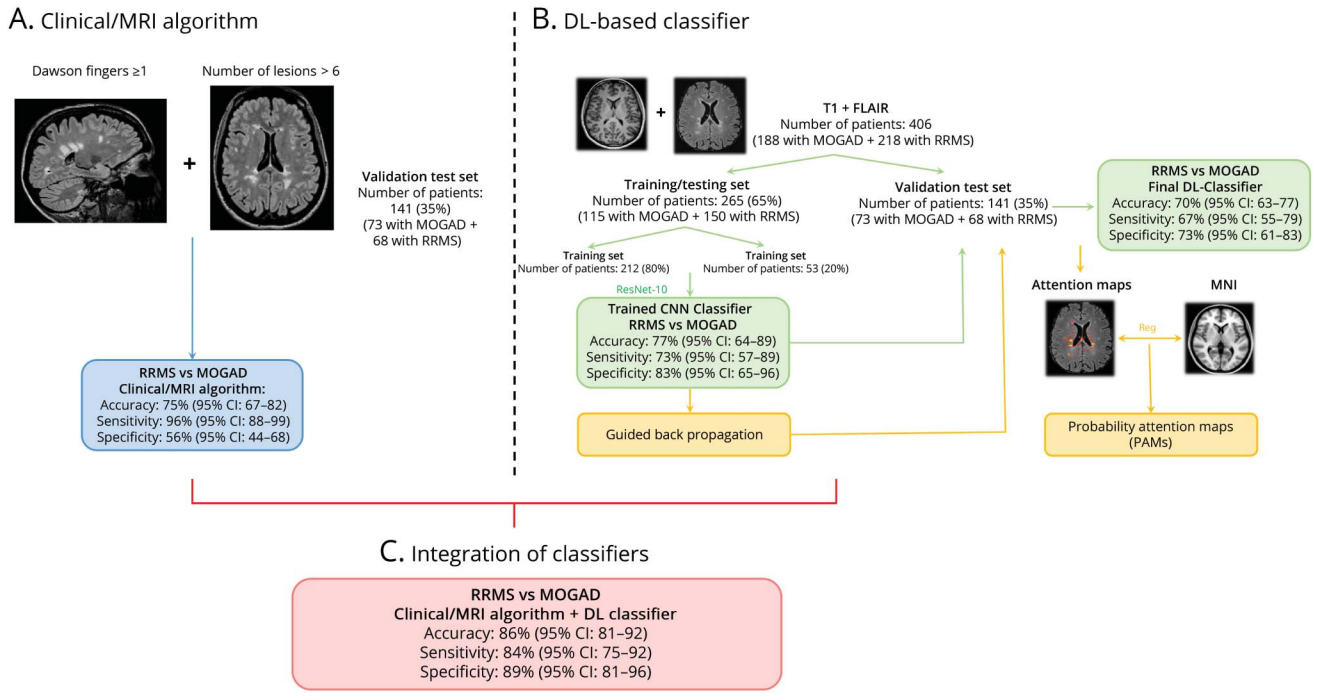
CNN-Derived Attention Map Analysis

To uncover key disease mechanisms leading to differentiation, CNN-derived attention maps were generated to indicate the most relevant anatomical features for CNN-based decisions.[19] The attention within a voxel, indicating its importance for decision making, was calculated using a guided back propagation approach and then normalized using z-scores. Two PAMs were created, one for each disease group by non-linearly registering the maps to MNI space and then averaging all the correctly classified attention maps of RRMS and MOGAD, respectively.

To identify the brain regions referring to the areas significantly contributing to the differentiation, the 2 PAMs were intersected with 2 atlases, the Adult Brain Maximum Probability Map ("Hammersmith") and the Johns Hopkins University diffusion tensor imaging (JHU DTI)-based white matter atlases. The regions within the intersection were tested for differences between RRMS and MOGAD using a *t*-test. This step resulted in 2 maps, 1 for each disease, with the value of the *t* statistic in each area. The volume (mm^3) of each region involved in the classification was calculated.

Finally, to understand the rationale behind the selection of specific regions for identifying different diseases, a group-level

Figure 1 Schematic Representation of the Study Design



The work was structured in 3 steps: (A) FLAIR/T2 sequences of the scans from the validation test set were assessed and scored considering the number of lesions and the presence of Dawson fingers (blue). (B) ResNet-10 architecture was trained using cross-entropy loss function and then accuracy was assessed on the validation test set (green). Then, individual attention maps were created using a guided backpropagation approach and then normalized using z-scores. Finally, 2 PAMs (1 for each disease group) were created by nonlinearly registering to MNI and then averaging all the correctly classified attention maps of RRMS and MOGAD, respectively (yellow). (C) The DL-based and clinical/MRI classifiers were combined to assess their performances in differentiating the 2 disorders (red). In brief, both the number of lesions and the presence of Dawson fingers were concatenated to the features extracted by the DL network and used to discriminate between the 2 classes. CNN = convolutional neural network; DL = deep learning; FLAIR = fluid-attenuated inversion recovery; MNI = Montreal Neurological Institute; MOGAD = myelin oligodendrocyte glycoprotein antibody–associated disease; MRI = magnetic resonance imaging; PAMs = probability attention maps; RRMS = relapsing remitting multiple sclerosis.

comparison of lesion distribution was conducted using the non-parametric permutation testing tool Randomise with cluster-based thresholding.

Integration of DL-Based and Clinical/MRI Classifiers

To explore whether the integration of the clinical/MRI features and DL could provide a more accurate discrimination between the 2 diseases, a fusion model was implemented. Specifically, the binarized total number of brain lesions (1: number >6, 0: otherwise) and the presence of Dawson fingers (1: present, 0: otherwise) for each patient were concatenated to the deep features extracted from the last residual layer of DL classifier. Then, a fully connected layer was added to generate the final class prediction. In addition, to assess whether incorporating other clinical discriminators could further improve model performance, we separately concatenated (1) the binarized presence of CSF-restricted oligoclonal bands (OCBs) and (2) clinical presentation involving the brain with the extracted deep imaging features before final classification.

Statistical Analysis

Means, medians, proportions of demographics, clinical features, and lesion characteristics were calculated for patients. Differences were evaluated using the Kruskal-Wallis or χ^2 test as appropriate. Based on the final diagnosis (RRMS or MOGAD) of the clinical/MRI algorithm, participants were classified as true positive (TP, clinical/MRI classifier fulfilled, diagnosis RRMS), true negative (TN, clinical/MRI classifier negative, diagnosis MOGAD), false positive (FP, clinical/MRI classifier positive, diagnosis MOGAD), or false negative (FN, clinical/MRI classifier negative, diagnosis RRMS). We obtained the accuracy of having RRMS vs MOGAD by the ratio TP + TN/TP + TN + FP + FN, sensitivity by the ratio TP/(TP + FN), specificity by TN/(TN + FP), positive predictive value (PPV) as TP/(TP + FP), negative predictive value (NPV) as TN/(TN + FN), and both the receiver operating characteristic (ROC) and precision recall area under the curve (AUC).

The model outputs consisted of 2 activation units (1 per class), with the predicted class determined by the highest activation. For ROC analysis, outputs were converted to class probabilities using the softmax function.

Data Availability

The anonymized grouped data that support the findings of this study may be available to qualified noncommercial investigators from the corresponding author, on reasonable request and in accordance with consent and ethics requirements.

Results

Study Population

In total, we analyzed 406 T2-FLAIR and T1-weighted brain MRI scans from patients with RRMS (n = 218, 150 F, mean [SD] age: 39 years [±11]) and MOGAD (n = 188, 114 F, mean [SD] age: 41 years [±14]). The training/testing cohort (number of patients: 265/406 [65%]) included 150 patients with RRMS and 115 patients with MOGAD. This set was further split into 2 independent subsets: the training set (number of patients: 212/265 [80%]), including 120 patients with RRMS and 92 patients with MOGAD, and the testing set (number of patients: 53/265 [20%]), including 30 patients with RRMS and 23 patients with MOGAD.

The validation group (number of patients: 141/406 [35%]) included 68 patients with RRMS and 73 patients with MOGAD. Clinical details of patients in the different sets are summarized in Table 1.

Validation of the Clinical/MRI Algorithm on the Validation Data Set

On the validation test set, the combination of >6 brain lesions and the presence of Dawson fingers predicted RRMS instead of MOGAD with 75% accuracy (95% CI 67–82), 96% sensitivity (95% CI 88–99), 56% specificity (95% CI 44–68), 67% PPV (95% CI 57–76), 93% NPV (95% CI 81–99), and a ROC-AUC of 0.76 (95% CI 0.70–0.82). The addition of the clinical presentation at onset (i.e., unilateral optic neuritis/transverse myelitis/other symptoms related to brain involvement) reached an accuracy of 52% (95% CI 44–61), sensitivity of 100% (95% CI 94–100), specificity of 23% (95% CI 11–30), PPV of 56% (95% CI 46–65), NPV of 99% (95% CI 81–99), and ROC-AUC of 0.61 (95% CI 0.56–0.67).

Performance of the DL-Based Classifier

The trained CNN classifier identified RRMS rather than MOGAD on the testing set with an accuracy of 77% (95% CI 64–89), sensitivity of 73% (95% CI 57–89), specificity of 83% (95% CI 65–96), and ROC-AUC of 0.82 (95% CI 0.69–0.93). The final DL model was further tested using the validation, unseen test set, reaching an accuracy of 70% (95% CI 63–77), sensitivity of 67% (95% CI 55–79), specificity of 73% (95% CI 61–83), PPV of 70% (95% CI 59–82), NPV of 69% (95% CI 58–81), and ROC-AUC of 0.74 (95% CI 0.65–0.82).

Probability Attention Maps

Using the validation test set, in the correctly classified RRMS cohort, PAMs revealed the involvement of corpus callosum (total volume: 1872 mm^3), left precentral gyrus (total volume: 341 mm^3), right thalamus (total volume: 193 mm^3), and right cingulate cortex (total volume: 186 mm^3) for the identification of the disease.

In the correctly classified MOGAD cohort, PAMs revealed the involvement of brainstem, particularly midbrain (total volume: 629 mm^3); bilateral hippocampus (total volume: 234 mm^3); and parahippocampal gyrus (total volume: 147 mm^3) for the identification of the disease.

Figure 2 shows the distribution of relevant voxels across the different anatomical regions.

The voxel-wise comparison of the lesion probability maps revealed that RRMS lesions were significantly more likely than MOGAD lesions to be periventricular (Figure 3).

Integration of Classifiers

When tested on the validation set, the combination of the clinical/MRI and DL-based classifiers reached the highest performance in differentiating the 2 diseases. Specifically, when the DL-based model was combined with having >6 brain lesions and presence of ≥ 1 Dawson finger, it was able to identify RRMS vs MOGAD with 86% accuracy (95% CI 81–92), 84% sensitivity (95% CI 75–92), 89% specificity (95% CI 81–96), 89% PPV (95% CI 80–96), 84% NPV (95% CI 75–92), and an ROC-AUC of 0.90 (95% CI 0.84–0.95). Figure 4 shows examples of RRMS and MOGAD cases classified by the DL model, demonstrating how combining the 2 classifiers enhanced differentiation and enabled more accurate classification.

Among the 11 misclassified patients with RRMS, 10 (91%) were CSF-restricted OCBs positive and only 1 had a phenotype at onset involving the brain compartment. Among the 7 misclassified patients with MOGAD, 2 (29%) had unmatched OCBs and 1 had ADEM. All other misclassified patients had optic neuritis and/or myelitis at onset but presented with brain lesions at the time of MRI analysis. The model incorporating the presence of CSF-restricted OCBs achieved an accuracy of 76% (95% CI 67–84), a sensitivity of 70% (95% CI 56–84), a specificity of 82% (95% CI 70–92), a PPV of 78 (95% CI 64–91), and an NPV of 74 (95% CI 61–86). The model incorporating clinical presentation involving the brain achieved an accuracy of 79% (95% CI 71–85), a sensitivity of 77% (95% CI 66–86), a specificity of 81% (95% CI 71–90), a PPV of 82% (95% CI 72–90), and an NPV of 75% (95% CI 64–85).

Table 2 summarizes the results of the validation analysis, highlighting the performance of the clinical/MRI features, the DL model, and their combination in distinguishing RRMS from MOGAD.

Classification of Evidence

This study provides Class III evidence that both a clinical/MRI algorithm and an MRI-based DL model accurately distinguish RRMS from MOGAD.

Discussion

In this multicenter study, we showed that both the clinical/MRI algorithm and the DL model differentiated RRMS from MOGAD with high accuracy, sensitivity, and specificity across

Table 1 Clinical Features of Patients With RRMS and MOGAD in the Whole Cohort, Training/Testing Set, and Validation Test Set

	Whole cohort (n = 406)			Training/testing set (n = 265, 65%)			Validation test set (n = 141, 35%)		
	RRMS (n = 218)	MOGAD (n = 188)	*p* Value[a]	RRMS (n = 150)	MOGAD (n = 115)	*p* Value[a]	RRMS (n = 68)	MOGAD (n = 73)	*p* Value[a]
Sex (F/M)	150/68	114/74	0.10	108/42	70/45	0.07	44/24	46/27	0.97
Age at MRI, y, mean ± SD	39 ± 11	41 ± 14	0.06	39 ± 10	42 ± 13	0.02	40 ± 14	40 ± 15	0.85
Disease duration, y, median (range)	5.7 (0.3–52.5)	3.0 (0.3–47.3)	0.01	5.7 (0.3–29.3)	4.5 (0.3–47.1)	0.13	5.5 (0.3–52.5)	2.2 (0.3–47.3)	0.01
EDSS score at MRI, median (range)[b]	2 (0–8.0)	2 (0–8.5)	0.42	2 (0–8)	2 (0–7.5)	0.43	2 (0–6)	2 (0–8.5)	0.74
Patients with CSF oligoclonal bands, n (%)									
Absence	14 (6)	134 (71)	<0.001	9 (6)	84 (73)	<0.001	5 (7)	50 (68)	<0.001
Presence	145 (67)	26 (14)		90 (60)	16 (14)		54 (80)	10 (14)	
NA	59 (27)	28 (15)		51 (34)	15 (13)		9 (13)	13 (18)	
Patients tested with CBA, n (%)[c]									
Fixed	11 (5)	93 (50)	<0.001	8 (5)	68 (59)	<0.001	3 (4)	25 (34)	<0.001
Live	0 (0)	88 (47)		0 (0)	40 (35)		0 (0)	48 (66)	
NA	207 (95)	7 (3)		142 (95)	7 (6)		65 (96)	0 (0)	
Type of presenting attack, n (%) patients									
Unilateral ON	52 (24)	59 (31)	<0.001	35 (23)	34 (30)	<0.001	17 (25)	25 (34)	0.04
Bilateral ON	4 (2)	34 (18)		2 (1)	22 (19)		2 (3)	12 (16)	
Short myelitis	70 (32)	14 (8)		51 (34)	9 (8)		19 (28)	5 (7)	
LETM	1 (0.5)	27 (14)		1 (1)	19 (17)		0	8 (11)	
Brainstem/cerebellar attack	30 (14)	8 (4)		17 (11)	6 (5)		13 (19)	2 (3)	
Cortical encephalitis	0 (0)	0 (0)		0	0		0	0	
ADEM	2 (1)	10 (5)		2 (2)	5 (4)		0	5 (7)	
Cerebral mono/polyfocal	22 (10)	3 (2)		14 (9)	1 (1)		8 (12)	2 (3)	
Combination of symptoms	12 (5.5)	29 (16)		4 (3)	16 (14)		8 (12)	13 (18)	
NA	25 (11)	4 (2)		24 (16)	3 (2)		1 (1)	1 (1)	
No. of relapses at study inclusion, median (range)	2 (1–13)	1 (1–15)	0.25	2 (1–13)	1 (1–15)	0.75	2 (1–4)	1 (1–5)	0.27
Patients with brain lesions, n (%)	218 (100)	154 (82)	<0.001	150 (100)	87 (76)	<0.001	68 (100)	67 (92)	0.06
No. of brain lesions, mean ± SD	32 ± 27	14 ± 26	<0.001	28 ± 22	9 ± 13	<0.001	41 ± 33	24 ± 37	0.004
No. of patients with ≥6 brain T2 lesions	190	65	<0.001	128	33	<0.001	62	32	<0.001
No. of patients with ≥1 Dawson finger	160	22	<0.001	104	16	<0.001	56	6	<0.001

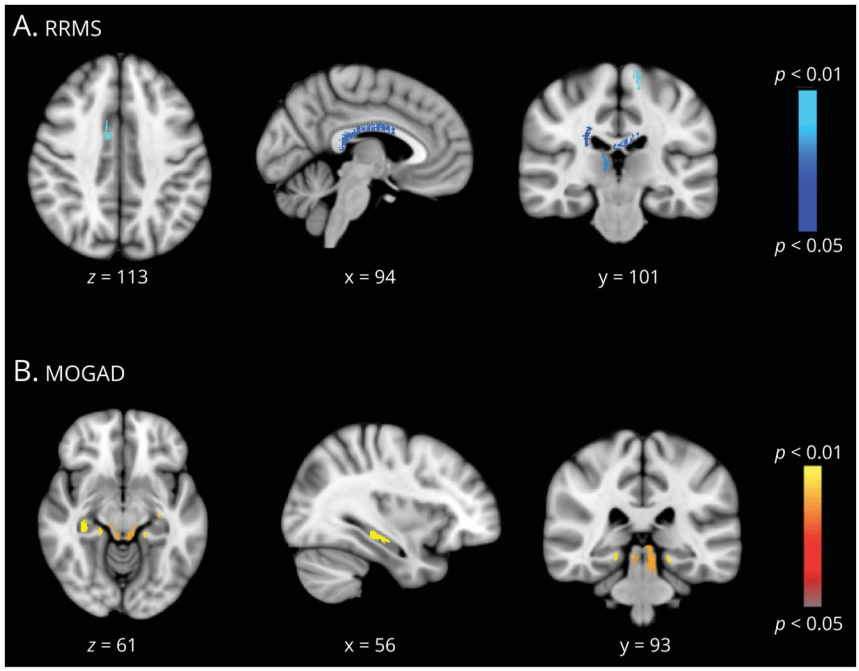
Abbreviations: ADEM = acute disseminated encephalomyelitis; CBA = cell-based assay; EDSS = Expanded Disability Status Scale; LETM = longitudinally extensive transverse myelitis; MOGAD = myelin oligodendrocyte glycoprotein antibody–associated disease; NA = not available; ON = optic neuritis; RRMS = relapsing remitting multiple sclerosis.

[a] Using the Kruskall–Wallis or χ^2 test as appropriate, depending on the nature of the variable. The distribution of continuous variables was evaluated applying the Shapiro-Wilk test.

[b] Time at MRI corresponds to study inclusion.

[c] In line with recent evidence and clinical recommendations,[24] only a subset of patients with MS was tested for MOG-IgG to avoid false positives and diagnostic misclassification.

Figure 2 Topographical Distributions of the Areas Significantly Contributing to the Differentiation Between Patients With RRMS and With MOGAD

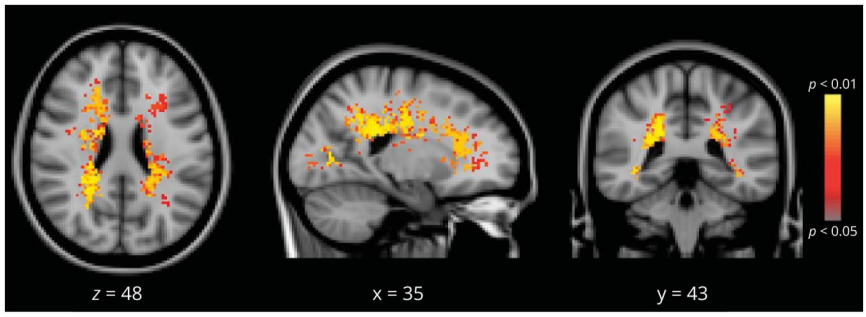


PAMs revealed the involvement of corpus callosum, right cingulate cortex, right thalamus, and left precentral gyrus for the identification of RRMS (A) while the relevant role of the brainstem, bilateral hippocampi, and parahippocampal gyri in the identification of MOGAD (B). Significant voxels are shown in a color scale from light blue to dark blue for RRMS and from yellow to red for MOGAD, from the most to the less significant, respectively (all $p < 0.001$). MOGAD = myelin oligodendrocyte glycoprotein antibody–associated disease; PAMs = probability attention maps; RRMS = relapsing remitting multiple sclerosis.

multiple cohorts. Although the clinical/MRI algorithm exhibited superior sensitivity, the DL model achieved greater specificity. The combination of both methods further improved diagnostic accuracy, indicating their potential for diagnosing challenging cases or identifying patients for research studies. In addition, PAMs offered valuable insights into the distribution and severity of brain damage, providing a deeper understanding of the distinct pathophysiologic mechanisms underlying the 2 conditions.

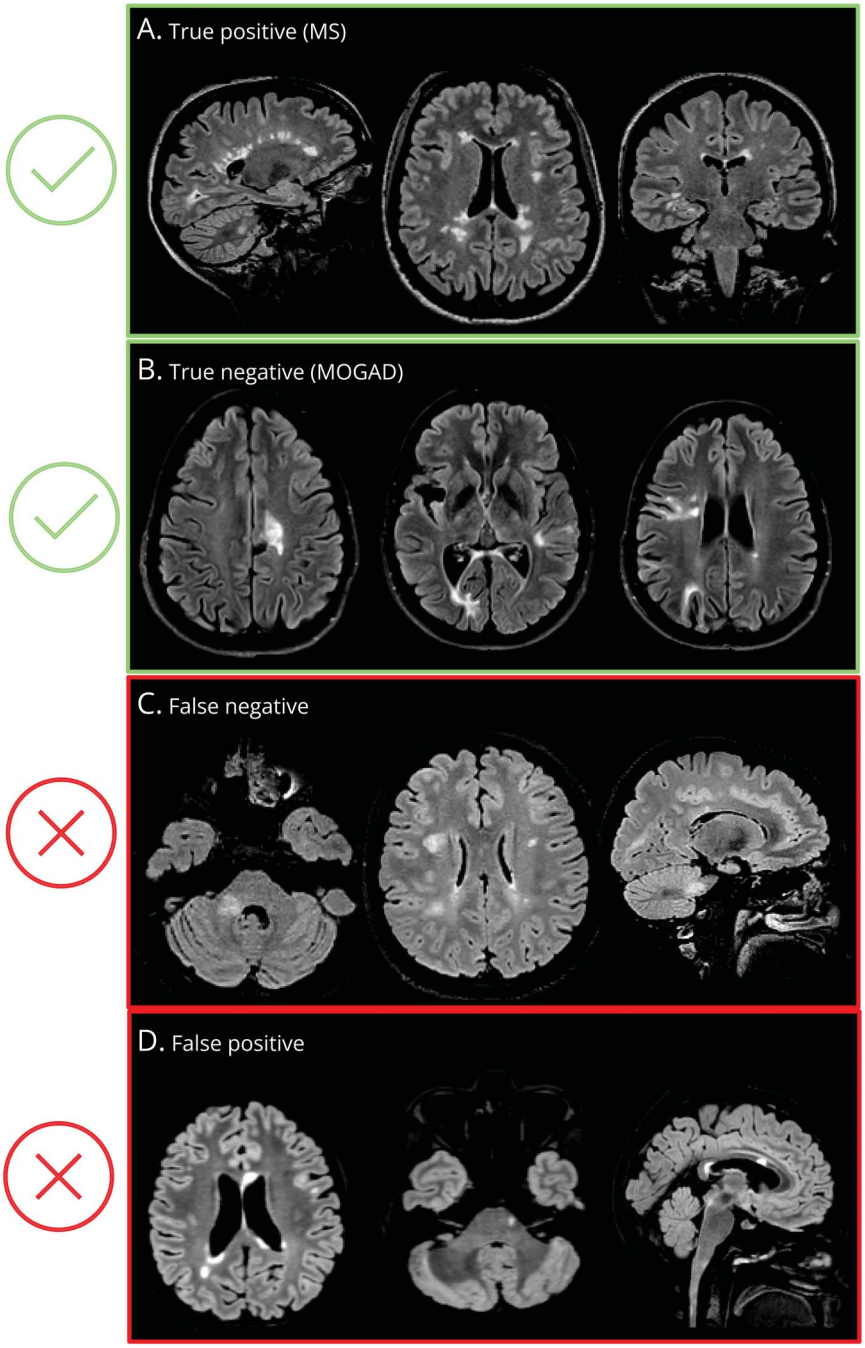
In this study, we retrospectively validated our clinical/MRI algorithm[12] within a new multicenter cohort of patients with RRMS and MOGAD. The results demonstrate that the algorithm can be effectively integrated into the diagnostic process for non-acute adult patients with suspected CNS inflammatory diseases, aiding in the identification of those who should undergo testing for MOG-Abs. Routine testing for MOG-Abs in all patients suspected of having MS is not appropriate in clinical practice because it could lead to an increased incidence of false-positive results.[8] Therefore, it is essential to pinpoint specific clinical and imaging features that increase the likelihood of MOGAD rather than RRMS. Our validation of the clinical/MRI algorithm demonstrated high sensitivity in accurately identifying patients with non-acute RRMS, thereby minimizing FNs, a significant consideration given the differing epidemiology of these 2 conditions. While MS is relatively common in adults (prevalence is increasing, exceeding 200/100,000 in some American and European countries),[25-27] MOGAD remains rare (prevalence 1.3–3.7/100,000).[9,28] As awareness and diagnostic testing for MOGAD continue to improve, its reported incidence and

Figure 3 Voxel-Wise Comparison of RRMS and MOGAD Lesion Maps



The voxel-wise comparison between RRMS and MOGAD showed a highest probability of RRMS lesions in the periventricular area. Significant voxels are shown in a color scale from yellow to red for RRMS, from the most to the least significant, respectively (all $p < 0.001$). MOGAD = myelin oligodendrocyte glycoprotein antibody–associated disease; RRMS = relapsing remitting multiple sclerosis.

Figure 4 Performance of DL-Based Classifiers in Differentiating RRMS and MOGAD



The figure shows examples of MRI, as classified by the DL classifier. (A) For MS, the classifier correctly identified oval-shaped periventricular and deep white matter lesions on axial FLAIR, with Dawson fingers and involvement of the temporal lobe. (B) For MOGAD, the classifier accurately detected large, ill-defined lesions in the juxtacortical/cortical and periventricular regions. (C) An example of MS misclassified as MOGAD revealed a poorly defined hyperintensity in the right cerebellar peduncle, along with more than 6 supratentorial lesions, some of which were round-shaped, and 1 Dawson finger. (D) By contrast, MOGAD misclassified as MS exhibited a well-defined peripheral lesion in the pons, alongside 4 supratentorial confluent lesions, including 2 round-shaped lesions involving the corpus callosum. The integration of clinical and imaging classifiers enabled accurate diagnosis in (C) and (D). DL = deep learning; FLAIR = fluid-attenuated inversion recovery; MNI = Montreal Neurological Institute; MOGAD = myelin oligodendrocyte glycoprotein antibody–associated disease; MS = multiple sclerosis; RRMS = relapsing remitting MS.

prevalence are expected to rise, potentially surpassing NMOSD over time, as suggested by recent studies.[9]

Our algorithm selectively identifies patients who may not need MOG-Ab testing, focusing on those more likely to have MOGAD. In non-acute adult patients with suspected CNS involvement, testing is questionable if there are >6 brain T2 lesions and ≥1 Dawson finger. Notably, adding information on onset presentation, CSF-restricted oligoclonal bands, or brain involvement did not improve the model's performance. This suggests that deep imaging features, especially those captured in the remission phase, may provide more robust discriminative power between MS and MOGAD, particularly in those MOGAD cases with overlapping symptoms with MS.

It is important to note that we developed a DL model that accurately differentiates between the 2 diseases, demonstrating high specificity. This has the potential to refine clinical workflows by identifying patients who would benefit most from MOG-Ab testing while minimizing false positives. Recent studies validating

Table 2 Results From the Validation Analysis to Predict RRMS Over MOGAD Using the Clinical/MRI Features and the Deep Learning Model, Alone and in Combination

Features	ROC-AUC (95% CI)	PR-AUC (95% CI)	Accuracy (%) (95% CI)	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	PPV (%) (95% CI)	NPV (%) (95% CI)
Dawson fingers ≥1 + no. of brain T2 lesions >6	0.76 (0.70–0.82)	0.89 (0.79–0.95)	75 (67–82)	96 (88–99)	56 (44–68)	67 (57–76)	93 (81–99)
Number of brain T2 lesions >6 + clinical presentation at onset[a]	0.61 (0.56–0.67)	0.73 (0.60–0.86)	52 (44–61)	100 (94–100)	23 (11–30)	56 (46–65)	99 (81–99)
DL model only	0.74 (0.65–0.82)	0.70 (0.58–0.82)	70 (63–77)	67 (55–79)	73 (61–83)	70 (59–82)	69 (58–81)
DL model + Dawson fingers ≥1 + no. of brain T2 lesions >6	0.90 (0.84–0.95)	0.91 (0.83–0.96)	86 (81–92)	84 (75–92)	89 (81–96)	89 (80–96)	84 (75–92)
DL model + Dawson fingers ≥1 + no. of brain T2 lesions >6 + CSF oligoclonal bands	0.81 (0.71–0.90)	0.77 (0.60–0.90)	76 (67–84)	70 (56–84)	82 (70–92)	78 (64–91)	74 (61–86)
DL model + Dawson fingers ≥1 + no. of brain T2 lesions >6 + clinical presentation involving brain	0.85 (0.78–0.92)	0.82 (0.72–0.93)	79 (71–85)	77 (66–86)	81 (71–90)	82 (72–90)	75 (64–85)

Abbreviations: AUC = area under curve; CSF = cerebrospinal fluid; DL = deep learning; MOGAD = myelin oligodendrocyte glycoprotein antibody–associated disease; NPV = negative predictive value; PPV = positive predictive value; PR = precision recall; ROC = receiver operating characteristic; RRMS = relapsing remitting multiple sclerosis.
[a] Either unilateral optic neuritis, or isolated transverse myelitis, or other symptoms related to brain involvement.

the 2023 MOGAD criteria[2] have shown better classification during attacks than remissions,[29] underscoring the challenge of diagnosing MOGAD in non-acute phases when typical clinical features may be absent. Therefore, our DL model offers a promising tool for improving the accuracy of MOGAD diagnosis, particularly in challenging non-acute phases.

A key advantage of DL is its ability to identify imaging features beyond visible lesion distribution. Because 18% of patients with MOGAD were lesion-free, the CNN model likely classified disorders using non-lesion features, a notable benefit. This is critical because optic nerve and spinal cord lesions, common in adult MOGAD, are less frequently imaged than the brain.[24] A DL tool using only brain MRI, which is more routinely acquired, has significant clinical relevance. Using multicenter data improved the model's robustness and generalizability, performing well on validation tests without fine-tuning. Integrating DL with our clinical/MRI algorithm enhanced performance, demonstrating its potential to aid clinicians while emphasizing the importance of combining clinical and imaging insights for final decisions. However, translating this model to clinical settings requires rigorous validation on data acquired prospectively to confirm its performance across diverse populations and clinical scenarios.

Notably, the use of PAMs facilitated the visualization of distinct patterns in damage severity within the brain, providing deeper insights into the underlying pathophysiologic differences between the 2 conditions. To elucidate the rationale behind selecting specific regions for distinguishing diseases, we conducted a voxel-wise comparison of lesion probability maps to differentiate RRMS from MOGAD. This comparison revealed that lesions in RRMS were significantly more likely periventricular than those in MOGAD. This finding likely explains why the corpus callosum was chosen as a key region for identifying MS. Moreover, our recent comparison of gray matter volumes indicated widespread atrophy in MS, particularly pronounced in the deep gray matter compared with MOGAD, accounting for the involvement of the thalamus in identifying MS.[30]

In MOGAD, the characteristic fluffy infratentorial lesions during acute phases tend to resolve at a higher rate than in other regions.[4,5] Therefore, we can speculate that brainstem involvement may be linked to microstructural damage after lesion resolution. However, further longitudinal studies are needed to confirm this observation. The role of hippocampal volume in differentiating MOGAD from MS has been demonstrated in various studies, though with conflicting results. Generally, lower hippocampal and parahippocampal volumes are associated with greater disability in MOGAD,[31] and these volumes tend to decrease over time at a higher atrophy rate compared with MS.[32] However, PAMs identify regions most responsible for DL-based differentiation without indicating the direction of damage, making our interpretations speculative. To strengthen these findings, additional studies with larger sample sizes and advanced imaging techniques are required.

This study has several limitations. MOGAD is a dynamic disease, with lesions evolving over time,[4] and it remains unclear whether microstructural damage may also change. Consequently, the cross-sectional design of our study may be limiting. Future research could consider integrating combined attack and remission scans with DL models to enhance sensitivity and specificity, particularly given the strong predictive value of T2 lesion resolution.[5] Moreover, longitudinal studies are needed to assess the potential evolution of microstructural damage in MOGAD and to determine whether the DL model can adapt to these changes

over time. In addition, we did not account for other clinical and imaging features, including MOG-Abs titers, and presence of optic nerve and cord lesions, because our focus was on using mostly the previously identified factors that distinguish the 2 diseases. Only patients aged 18 years or older were included, and future studies could investigate whether early disease presentations affect model predictions and evaluate its utility in younger populations. Regarding the methodological aspects, the study's generalizability may be limited by the reader's experience in assessing certain variables, such as Dawson fingers, although most variables remain objectively measurable. Finally, DL model training requires a large data set to provide robust results and avoid overfitting. Despite the use in this study of regularization techniques and data augmentation to mitigate the aforementioned limitations, more extensive tests on unseen data are needed to assess the generalizability of the proposed solution. Additional training/fine-tuning might be also required. Our future goal is to develop an online platform that, based on uploaded clinical data and MRI scans, can apply the most appropriate model to support diagnosis, particularly in complex cases. This approach could help guide antibody testing in uncertain cases and assist treatment decisions while awaiting test results in cases of possible overlap between MS and MOGAD.

In conclusion, this multicenter study demonstrates that both the clinical/MRI algorithm and the DL model effectively distinguish RRMS from MOGAD, with each showing complementary strengths in sensitivity and specificity. The integration of these approaches enhances diagnostic accuracy, underscoring their potential to refine clinical decision making. In addition, the application of PAMs has provided valuable insights into the distinct pathophysiologic features of these conditions, contributing to a deeper understanding of their underlying mechanisms. Future research should focus on prospectively validating the clinical/MRI algorithm and DL model in clinical scenarios using larger, diverse populations, including pediatric patients and those with atypical presentations of RRMS and MOGAD.

Author Byline *(Continued)*

Samira Luisa Apóstolos-Pereira,[5] Georgina Arrambide,[6] Barbara Bellenberg,[7] Alessia Bianchi,[1,8] Alvino Bisecco,[9] Benedetta Bodini,[10,11] Massimiliano Calabrese,[12] Valentina Camera,[12] Elisabeth G. Celius,[13] Carolina de Medeiros Rimkus,[14] Yunyun Duan,[15] Françoise Durand-Dubief,[16] Massimo Filippi,[17,18,19,20,21] Antonio Gallo,[9] Claudio Gasperini,[22] Cristina Granziera,[23,24,25] Sergiu Groppa,[26] Matthias Grothe,[27] Mor Gueye,[17,18,21] Matilde Inglese,[28,29] Anu Jacob,[30,31] Caterina Lapucci,[29] Andrea Lazzarotto,[10,11] Yaou Liu,[15] Sara Llufriu,[32] Carsten Lukas,[7,33] Romain Marignier,[16] Silvia Messina,[34] Jannis Müller,[23,24] Jacqueline Palace,[34] Luisa Pastó,[35] Friedemann Paul,[36] Ferran Prados,[8,37,38] Anne-Katrin Pröbstel,[24] Àlex Rovira,[39] Maria Assunta Rocca,[17,18,21] Serena Ruggieri,[22] Jaume Sastre-Garriga,[6] Douglas Kazutoshi Sato,[40] Ruth Schneider,[7,33] Maria Sepulveda,[32] Piotr Sowa,[41] Bruno Stankoff,[10,11] Carla Tortorella,[22] Frederik Barkhof,[42,43] Olga Ciccarelli,[8,44] Marco Battaglini,[1,2] Nicola De Stefano,[1] for the MAGNIMS Study Group

Affiliation

[1]Department of Medicine, Surgery and Neuroscience, University of Siena, Siena, Italy; [2]Siena Imaging SRL, 53100 Siena, Italy; [3]Department NEUROFARBA, University of Florence, Italy; [4]IRCCS Fondazione Don Carlo Gnocchi, Florence, Italy; [5]Departamento de Neurologia, Faculdade de Medicina, Universidade de São Paulo, São Paulo SP, Brazil; [6]Neurology-Neuroimmunology Department Multiple Sclerosis Centre of Catalonia (Cemcat), Vall d'Hebron Barcelona Hospital Campus. Universitat Autònoma de Barcelona, Barcelona, Spain; [7]Institute of Neuroradiology, St. Josef Hospital, Ruhr University Bochum, Germany; [8]Queen Square MS Centre, Department of Neuroinflammation, UCL Queen Square Institute of Neurology, Faculty of Brain Sciences, University College London, London, United Kingdom; [9]Department of Advanced Medical and Surgical Sciences, University of Campania "Luigi Vanvitelli", Naples, Italy; [10]Sorbonne Université, Paris Brain Institute, CNRS, Inserm, France; [11]AP-HP, Hôpital Universitaire Pitié-Salpêtrière, Paris, France; [12]Neurology Section of Department of Neuroscience, Biomedicine and Movement, University of Verona, Italy; [13]Department of Neurology, Oslo University Hospital and Faculty of Medicine, University of Oslo, Norway; [14]Departamento de Radiologia e Oncologia, Faculdade de Medicina, Universidade de São Paulo, São Paulo SP, Brazil; [15]Department of Radiology, Beijing Tiantan Hospital, Capital Medical University, Beijing, China; [16]Service de Neurologie, Sclérose en Plaques, Pathologies de la Myéline et Neuro-inflammation, Hôpital Neurologique Pierre Wertheimer, Hospices Civils de Lyon, France; [17]Neuroimaging Research Unit, IRCCS San Raffaele Scientific Institute, Milan, Italy; [18]Neurology Unit, IRCCS San Raffaele Scientific Institute, Milan, Italy; [19]Neurorehabilitation Unit, IRCCS San Raffaele Scientific Institute, Milan, Italy; [20]Neurophysiology Service, IRCCS San Raffaele Scientific Institute, Milan, Italy; [21]Vita-Salute San Raffaele University, Milan, Italy; [22]Department of Neurosciences, S. Camillo-Forlanini Hospital, Rome, Italy; [23]Translational Imaging in Neurology (ThINk) Basel, Department of Biomedical Engineering, Faculty of Medicine, University Hospital Basel and University of Basel, Switzerland; [24]Multiple Sclerosis Centre, Departments of Neurology, Clinical Research and Biomedicine, University Hospital and University Basel, Switzerland; [25]Research Center for Clinical Neuroimmunology and Neuroscience Basel (RC2NB), University Hospital Basel and University of Basel, Switzerland; [26]Department of Neurology, University Medical Center of the Johannes Gutenberg University Mainz, Germany; [27]Department of Neurology, University Medicine of Greifswald, Germany; [28]Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health (DiNOGMI), University of Genoa, Italy; [29]IRCCS Ospedale Policlinico San Martino, Genoa, Italy; [30]NMO Clinical Service at the Walton Centre, Liverpool, United Kingdom; [31]Department of Neurology, Cleveland Clinic, AbuDhabi, UAE; [32]Neuroimmunology and Multiple Sclerosis Unit and Laboratory of Advanced Imaging in Neuroimmunological Diseases (ImaginEM), Hospital Clinic and Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), University of Barcelona, Spain; [33]Department of Neurology, St. Josef Hospital, Ruhr University Bochum, Germany; [34]Department of Clinical Neurology, John Radcliffe Hospital, Oxford, United Kingdom; [35]Careggi University Hospital of Florence, Italy; [36]Experimental and Clinical Research Center, Max Delbrueck Center for Molecular Medicine and Charité - Universitaetsmedizin Berlin, Germany; [37]UCL Hawkes Institute, London, United Kingdom; [38]E-Health Center Universitat Oberta de Catalunya, Barcelona, Spain; [39]Section of Neuroradiology, Department of Radiology, Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona, Spain; [40]Pontifícia Universidade Católica do Rio Grande do Sul, School of Medicine, Porto Alegre RS, Brazil; [41]Division of Radiology and Nuclear Medicine, Oslo University Hospital, Norway; [42]Department of Radiology & Nuclear Medicine, Amsterdam UMC, Vrije Universiteit, the Netherlands; [43]Queen Square Institute of Neurology and Centre for Medical Image Computing, University College, London, United Kingdom; and [44]National Institute for Health Research (NIHR) University College London Hospitals (UCLH) Biomedical Research Centre, London, United Kingdom.

Author Contributions

R. Cortese: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design; analysis or interpretation of data. F. Sforazzini: drafting/revision of the manuscript for content, including medical writing for content; study concept or design; analysis or interpretation of data. G. Gentile: analysis or interpretation of data. A. de Mauro: major role in the acquisition of data. L. Luchetti: analysis or interpretation of data. M.P. Amato: major role in the acquisition of data. S.L. Apóstolos-Pereira: major role in the acquisition of data. G. Arrambide: major role in the acquisition of data. B. Bellenberg: major role in the acquisition of data. A. Bianchi: major role in the acquisition of data. A. Bisecco: major role in the acquisition of data. B. Bodini: major role in the acquisition of data. M. Calabrese: major role in the acquisition of data. V. Camera: major role in the acquisition of data. E.G. Celius: major role in the acquisition of data. C. de Medeiros Rimkus: major role in the acquisition of data. Y. Duan: major role in the acquisition of data. F. Durand-Dubief: major role in the

acquisition of data. M. Filippi: major role in the acquisition of data. A. Gallo: major role in the acquisition of data. C. Gasperini: major role in the acquisition of data. C. Granziera: major role in the acquisition of data. S. Groppa: major role in the acquisition of data. M. Grothe: major role in the acquisition of data. M. Gueye: major role in the acquisition of data. M. Inglese: major role in the acquisition of data. A. Jacob: major role in the acquisition of data. C. Lapucci: major role in the acquisition of data. A. Lazzarotto: major role in the acquisition of data. Y. Liu: major role in the acquisition of data. S. Llufriu: major role in the acquisition of data. C. Lukas: major role in the acquisition of data. R. Marignier: major role in the acquisition of data. S. Messina: major role in the acquisition of data. J. Müller: major role in the acquisition of data. J. Palace: drafting/revision of the manuscript for content, including medical writing for content. L. Pastó: major role in the acquisition of data. F. Paul: drafting/revision of the manuscript for content, including medical writing for content. F. Prados: analysis or interpretation of data. A.-K. Pröbstel: major role in the acquisition of data. À. Rovira: drafting/revision of the manuscript for content, including medical writing for content. M.A. Rocca: drafting/revision of the manuscript for content, including medical writing for content. S. Ruggieri: major role in the acquisition of data. J. Sastre-Garriga: major role in the acquisition of data. D.K. Sato: drafting/revision of the manuscript for content, including medical writing for content. R. Schneider: major role in the acquisition of data. M. Sepulveda: major role in the acquisition of data. P. Sowa: major role in the acquisition of data. B. Stankoff: major role in the acquisition of data. C. Tortorella: major role in the acquisition of data. F. Barkhof: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. O. Ciccarelli: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. M. Battaglini: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design; analysis or interpretation of data. N. De Stefano: drafting/revision of the manuscript for content, including medical writing for content; study concept or design; analysis or interpretation of data.

Study Funding

The authors report no targeted funding.

Disclosure

R. Cortese was awarded a MAGNIMS-ECTRIMS fellowship in 2019; received speaker honoraria/travel support from Roche, Merck Serono, UCB, Sanofi Genzyme, Alexion, Novartis, and Janssen; and received a research grant from the Italian Ministry of University and Research, Research Project of Relevant National Interest (PRIN), project code: 2022PR3PEY. M.P. Amato has served on Scientific Advisory Boards for Biogen, Novartis, Roche, Merck, Sanofi Genzyme, and Teva; received speaker honoraria from Biogen, Merck, Sanofi Genzyme, Roche, Novartis, and Teva; and received research grants for her Institution from Biogen, Merck, Sanofi Genzyme, Novartis, and Roche. G. Arrambide has received speaking honoraria; has received compensation for consulting services or participation in advisory boards from Merck, Roche, and Horizon Therapeutics; has received travel support for scientific meetings from Novartis, Roche, and ECTRIMS; is editor of *Europe of the Multiple Sclerosis Journal – Experimental, Translational and Clinical*; is a member of the International Women in Multiple Sclerosis (iWiMS) network executive committee; and is a member of the European Biomarkers in MS (BioMS-eu) consortium steering committee. A. Bisecco received speaker's honoraria and/or compensation for consulting service and/or speaking activities and/or travel grants from Biogen, Roche, Novartis, Merck, Alexion, Amgen, UCB, and Genzyme. B. Bodini has received traveling and speaker's honoraria from Novartis, Genzyme, Roche, and Merck Serono; and received research support from Biogen. M. Calabrese received speaker honoraria from Biogen, Bristol Myers Squibb, Celgene, Genzyme, Merck Serono, Novartis, and Roche; and receives research support from the Progressive MS Alliance and Italian Minister of Health. E.G. Celius has received honoraria for lecturing and advice from Biogen, Merck, and Roche, and grants and honoraria from Novartis and Sanofi Genzyme. C. de Medeiros Rimkus was awarded an MSIF-ECTRIMS fellowship in 2021; is an associate editor of *Arquivos de Neuropsiquiatria*, and has received speaker honoraria from Roche Pharma, Novartis, and Guerbet. M. Filippi is Editor-in-Chief of the *Journal of Neurology*, an associate editor of *Human Brain Mapping*, *Radiology*, and *Neurological Sciences*; has received compensation for consulting services from Alexion, Almirall, Biogen, Merck, Novartis, Roche, and Sanofi Genzyme; for speaking activities from Bayer, Biogen, Celgene, Chiesi Italia SpA, Eli Lilly, Genzyme, Janssen, Merck-Serono, Neopharmed Gentili, Novartis, Novo Nordisk, Roche, Sanofi Genzyme, Takeda, and TEVA; for participation in advisory boards for Alexion, Biogen, Bristol Myers Squibb, Merck, Novartis, Roche, Sanofi Aventis, Sanofi Genzyme, and Takeda; and for scientific direction of educational events for Biogen, Merck, Roche, Celgene, Bristol Myers Squibb, Lilly, Novartis, and Sanofi Genzyme; and receives research support from Biogen Idec, Merck-Serono, Novartis, Roche, Italian Ministry of Health, Fondazione Italiana Sclerosi Multipla, and ARiSLA (Fondazione Italiana di Ricerca per la SLA). C. Gasperini received fees as an invited speaker or travel expenses for attending meeting from Biogen, Merck-Serono, Teva, Mylan, Sanofi Genzyme, and Novartis. The University Hospital Basel (USB), as the employer of C. Granziera, has received the following fees that were used exclusively for research support: (1) advisory board and consultancy fees from Actelion, Sanofi Genzyme, Novartis, GeNeuro, and Roche; (2) speaker fees from Sanofi Genzyme, Novartis, GeNeuro, and Roche; (3) research support from Siemens, GeNeuro, and Roche. C. Granziera is supported by the Swiss National Science Foundation (SNSF) grant PP00P3_176984, the Stiftung zur Förderung der gastroenterologischen und allgemeinen klinischen Forschung, and the EUROSTAR E!113682

HORIZON2020. M. Grothe has received compensation for serving on scientific advisory boards for Novartis, Roche, and Sanofi Genzyme; received speaker honoraria and travel support from Merck, Novartis, Roche, Sanofi Genzyme, and Teva; and received research support from Novartis. M. Inglese received grants from NIH, NMSS, FISM, and Ministero dell'Università e ricerca (MUR); and received fees for consultation from BMS, Janssen, Roche, Genzyme, Merck, Biogen, and Novartis. C. Lapucci has received honoraria for speaking, travel grants, and participating in advisory boards from Merck, Sanofi, Novartis, Roche, and Alexion. S. Llufriu received compensation for consulting services and speaker honoraria from Biogen Idec, Novartis, Bristol Myers Squibb, Sanofi, Johnson & Johnson, and Merck. C. Lukas received a research grant from the German Federal Ministry for Education and Research, BMBF, and German Competence Network Multiple Sclerosis (KKNMS, grant 01GI1601I); and has received consulting and speaker's honoraria from Biogen Idec, Bayer Schering, Daiichi Sanykyo, Merck Serono, Novartis, Sanofi, Genzyme, and TEVA. J. Müller has received funding from the Swiss National Science Foundation (grants P500PM_214230 and P5R5PM_225288). J. Palace has received support for scientific meetings and honoraria for advisory work from Merck Serono, Novartis, Chugai, Alexion, Roche, Medimmune, Argenx, UCB, Mitsubishi, Amplo, Janssen, and Sanofi; has received grants from Alexion, Roche, Medimmune, UCB, and Amplo Biotechnology; has patent ref. P37347WO and licence agreement Numares multimarker MS diagnostics Shares in AstraZeneca; and acknowledges partial funding by highly specialized services NHS England. F. Paul serves on scientific advisory boards for Novartis, Viela Bio, and Alexion; and has received speaker honoraria from Bayer, Teva, Merck, Viela, Alexion, Roche, and Novartis. F. Prados received a Guarantors of Brain fellowship 2017–2020 and is supported by National Institute for Health Research (NIHR), Biomedical Research Centre initiative at University College London Hospitals (UCLH). A.K. Pröbstel has participated as speaker in meetings sponsored by and received consulting fees and/or grant support from Biogen and Roche. A. Rovira serves/ed on scientific advisory boards for Novartis, Sanofi Genzyme, Synthetic MR, TensorMedical, Roche, and Biogen; has received speaker honoraria from Bayer, Sanofi Genzyme, Merck-Serono, Teva Pharmaceutical Industries Ltd., Novartis, Roche, Bristol Myers, and Biogen; is CMO and cofounder of TensorMedical; and receives research support from Fondo de Investigación en Salud (PI19/00950) from Instituto de Salud Carlos III, Spain. M.A. Rocca received consulting fees from Biogen, Bristol Myers Squibb, Eli Lilly, Janssen, and Roche; received speaker honoraria from AstraZeneca, Biogen, Bristol Myers Squibb, Bromatech, Celgene, Genzyme, Horizon Therapeutics Italy, Merck Serono SpA, Novartis, Roche, Sanofi, and Teva; receives research support from the MS Society of Canada, the Italian Ministry of Health, the Italian Ministry of University and Research, and Fondazione Italiana Sclerosi Multipla; and is an associate editor of *Multiple Sclerosis and Related Disorders*. J. Sastre-Garriga serves as a coeditor of Europe for the *Multiple Sclerosis Journal* and as Editor-in-Chief of *Revista de Neurología*; receives research support from Fondo de Investigaciones Sanitarias (19/950 and 22/750); and in the last 12 months, has served as a consultant/speaker for BMS, Roche, Sanofi, Janssen, and Merck. D.K. Sato received research support from CNPq/Brazil (425331/2016-4, 308636/2019-8, and 306016/2022-2), FAPERGS/Brazil (17/2551-0001391-3 and 21/2551-0000077-5), Teva, Merck, and Biogen for investigator-initiated studies and speaker honoraria from Horizon, Amgen, Alexion, AstraZeneca, Novartis, and Roche; and participated in advisory boards for Horizon, Roche, Alexion, and AstraZeneca. R. Schneider has received consulting and speaker honoraria from Merck, Biogen-Idec, GmbH, Bayer HealthCare, Alexion Pharma, Novartis Pharma, and Roche Pharma AG; and has received research scientific grant support from Novartis Pharma. M. Sepulveda received speaking honoraria from Roche and Biogen; and received travel reimbursement from Roche, Biogen, and Sanofi for national and international meetings. B. Stankoff has received grants and personal fees for lectures from Roche, Sanofi Genzyme, and Merck-Serono; and personal fees for lectures from Novartis and Janssen. C. Tortorella received honoraria for speaking and travel grants from Biogen, Sanofi-Aventis, Merck Serono, Bayer-Schering, Teva, Genzyme, Roche, Alexion, Horizon, Celgene, Almirall, and Novartis. F. Barkhof is a steering committee or data safety monitoring board member for Biogen, Merck, Eisai, and Prothena; is an advisory board member for Combinostics, Scottish Brain Sciences, and Alzheimer Europe; is a consultant for Roche, Celltrion, Rewind Therapeutics, Merck, and Bracco; has research agreements with ADDI, Merck, Biogen, GE Healthcare, and Roche; and is cofounder and shareholder of Queen Square Analytics LTD. O. Ciccarelli is an NIHR Research Professor (RP-2017-08-ST2-004); acts as a consultant for Biogen, Merck, Novartis, Roche, and Lundbeck; and has received research grant support from the MS Society of Great Britain and Northern Ireland, the NIHR UCLH Biomedical Research Centre, the Rosetree Trust, the National MS Society, and the NIHR-HTA. N. De Stefano has received honoraria from Biogen-Idec, Bristol Myers Squibb, Celgene, Genzyme, Immunic, Merck Serono, Novartis, Roche, and Teva for consulting services, speaking, and travel support; serves on advisory boards for Merck, Novartis, Biogen-Idec, Roche, Genzyme, and Immunic; and has received research grant support from the Italian MS Society. All other authors report no relevant disclosures. Go to Neurology.org/N for full disclosures.

Publication History

Received by *Neurology*® January 16, 2025. Accepted in final form June 25, 2025. Submitted and externally peer reviewed. The handling editor was Editor-in-Chief José Merino, MD, MPhil, FAAN.

References

1. Marignier R, Hacohen Y, Cobo-Calvo A, et al. Myelin-oligodendrocyte glycoprotein antibody-associated disease. *Lancet Neurol*. 2021;20(9):762-772. doi:10.1016/S1474-4422(21)00218-0
2. Banwell B, Bennett JL, Marignier R, et al. Diagnosis of myelin oligodendrocyte glycoprotein antibody-associated disease: International MOGAD Panel proposed criteria. *Lancet Neurol*. 2023;22(3):268-282. doi:10.1016/s1474-4422(22)00431-8

3. Cortese R, Prados Carrasco F, Tur C, et al. Differentiating multiple sclerosis from AQP4-neuromyelitis optica spectrum disorder and MOG-antibody disease with imaging. *Neurology*. 2023;100(3):E308-E323. doi:10.1212/WNL.0000000000201465
4. Sechi E, Krecke KN, Messina SA, et al. Comparison of MRI lesion evolution in different central nervous system demyelinating disorders. *Neurology*. 2021;97(11):e1097-e1109. doi:10.1212/WNL.0000000000012467
5. Cacciaguerra L, Abdel-Mannan O, Champsas D, et al. Radiologic lag and brain MRI lesion dynamics during attacks in MOG antibody-associated disease. *Neurology*. 2024;102(10):e209303. doi:10.1212/WNL.0000000000209303
6. Waters PJ, Komorowski L, Woodhall M, et al. A multicenter comparison of MOG-IgG cell-based assays. *Neurology*. 2019;92(11):e1250-e1255. doi:10.1212/WNL.0000000000007096
7. Villacieros-Álvarez J, Espejo C, Arrambide G, et al. Myelin oligodendrocyte glycoprotein antibodies in adults with a first demyelinating event suggestive of multiple sclerosis. *Ann Neurol*. 2024;95(1):116-128. doi:10.1002/ANA.26793
8. Sechi E, Buciuc M, Pittock SJ, et al. Positive predictive value of myelin oligodendrocyte glycoprotein autoantibody testing. *JAMA Neurol*. 2021;78(6):741-746. doi:10.1001/jamaneurol.2021.0912
9. Cacciaguerra L, Sechi E, Komla-Soukha I, et al. MOG antibody-associated disease epidemiology in Olmsted County, USA, and Martinique. *J Neurol*. 2025;272(2):118. doi:10.1007/s00415-024-12861-9
10. Höftberger R, Guo Y, Flanagan EP, et al. The pathology of central nervous system inflammatory demyelinating disease accompanying myelin oligodendrocyte glycoprotein autoantibody. *Acta Neuropathol*. 2020;139(5):875-892. doi:10.1007/s00401-020-02132-y
11. O'Connell S, Cannon DM, Broin P. Predictive modelling of brain disorders with magnetic resonance imaging: a systematic review of modelling practices, transparency, and interpretability in the use of convolutional neural networks. *Hum Brain Mapp*. 2023;44(18):6561-6574. doi:10.1002/HBM.26521
12. Cortese R, Battaglini M, Prados F, et al. Clinical and MRI measures to identify non-acute MOG-antibody disease in adults. *Brain*. 2023;146(6):2489-2501. doi:10.1093/BRAIN/AWAC480
13. Roca P, Attye A, Colas L, et al. Artificial intelligence to predict clinical disability in patients with multiple sclerosis using FLAIR MRI. *Diagn Interv Imaging*. 2020;101(12):795-802. doi:10.1016/J.DIII.2020.05.009
14. Storelli L, Rocca MA, Pagani E, et al. Measurement of whole-brain and gray matter atrophy in multiple sclerosis: assessment with MR imaging. *Radiology*. 2018;288(2):554-564. doi:10.1148/radiol.2018172468
15. Collorone Sara. Artificial intelligence applied to MRI data to tackle key challenges in multiple sclerosis. *Mult Scler*. 2024. doi:10.1177/13524585241249422
16. Cacciaguerra L, Storelli L, Radaelli M, et al. Application of deep-learning to the seronegative side of the NMO spectrum. *J Neurol*. 2022;269(3):1546-1556. doi:10.1007/s00415-021-10727-y
17. Rocca MA, Anzalone N, Storelli L, et al. Deep learning on conventional magnetic resonance imaging improves the diagnosis of multiple sclerosis mimics. *Invest Radiol*. 2021;56(4):252-260. doi:10.1097/RLI.0000000000000735
18. Bach S, Binder A, Montavon G, Klauschen F, Müller KR, Samek W. On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *PLoS One*. 2015;10(7):e0130140. doi:10.1371/JOURNAL.PONE.0130140
19. Coll L, Pareto D, Carbonell-Mirabent P, et al. Deciphering multiple sclerosis disability with deep learning attention maps on clinical MRI. *Neuroimage Clin*. 2023;38:103376. doi:10.1016/j.nicl.2023.103376
20. magnims.eu
21. 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. doi:10.1016/S1474-4422(17)30470-2
22. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*. 2016;2016:770-778. doi:10.1109/CVPR.2016.90
23. Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. *32nd Int Conf Mach Learn ICML 2015*. 2015;1:448-456. Accessed September 27, 2024. arxiv.org/abs/1502.03167v3
24. Geraldes R, Arrambide G, Banwell B, et al. The influence of MOGAD on diagnosis of multiple sclerosis using MRI. *Nat Rev Neurol*. 2024;20(10):620-635. doi:10.1038/S41582-024-01005-2
25. Walton C, King R, Rechtman L, et al. Rising prevalence of multiple sclerosis worldwide: insights from the Atlas of MS, third edition. *Mult Scler*. 2020;26(14):1816-1821. doi:10.1177/1352458520970841
26. Carlson AM, Sollero CV, Nair KV, et al. Prevalence of multiple sclerosis and treatment utilization in a large, highly diverse population. *Mult Scler Relat Disord*. 2022;61:103784. doi:10.1016/J.MSARD.2022.103784
27. Portaccio E, Magyari M, Havrdova EK, et al. Multiple sclerosis: emerging epidemiological trends and redefining the clinical course. *Lancet Reg Health Eur*. 2024;44:100977. doi:10.1016/J.LANEPE.2024.100977
28. O'Connell K, Hamilton-Shield A, Woodhall M, et al. Prevalence and incidence of neuromyelitis optica spectrum disorder, aquaporin-4 antibody-positive NMOSD and MOG antibody-positive disease in Oxfordshire, UK. *J Neurol Neurosurg Psychiatry*. 2020;91(10):1126-1128. doi:10.1136/JNNP-2020-323158
29. Fonseca E, Olivé-Cirera G, Martinez-Hernandez E, et al. Investigating the 2023 MOGAD criteria in children and adults with MOG-antibody positivity within and outside attacks. *Neurology*. 2024;103(6):e209682. doi:10.1212/WNL.0000000000209682
30. Cortese R, Battaglini M, Prados F, et al. Grey matter atrophy and its relationship with white matter lesions in patients with myelin oligodendrocyte glycoprotein antibody-associated disease, aquaporin-4 antibody-positive neuromyelitis optica spectrum disorder, and multiple sclerosis. *Ann Neurol*. 2024;96(2):276-288. doi:10.1002/ANA.26951
31. Zhuo Z, Duan Y, Tian D, et al. Brain structural and functional alterations in MOG antibody disease. *Mult Scler J*. 2021;27(9):1350-1363. doi:10.1177/1352458520964415
32. Lotan I, Billiet T, Ribbens A, et al. Volumetric brain changes in MOGAD: a cross-sectional and longitudinal comparative analysis. *Mult Scler Relat Disord*. 2023;69:104436. doi:10.1016/j.msard.2022.104436