2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

An interpretable machine learning model to explain the interplay between brain lesions and cortical atrophy in multiple sclerosis

A. Conti, C.A. Treaba, A. Mehndiratta, V.T. Barletta, C. Mainero[#], and N. Toschi[#], *Senior Member, IEEE*

***Abstract*— Multiple Sclerosis (MS) is the most common cause, (after trauma) of neurological disability in young adults in Western countries. While several Magnetic Resonance Imaging (MRI) studies have demonstrated a strong association between the presence of cortical grey matter atrophy and the progression of neurological impairment in MS patients, the neurobiological substrates of cortical atrophy in MS, and in particular its relationship with white matter (WM) and cortical lesions, remain unknown. The aim of this study was to investigate the interplay between cortical atrophy and different types of lesions at Ultra-High Field (UHF) 7 T MRI, including cortical lesions and lesions with a susceptibility rim (a feature which histopathological studies have associated with impaired remyelination and progressive tissue destruction). We combined lesion characterization with a recent machine learning (ML) framework which includes explainability, and we were able to predict cortical atrophy in MS from a handful of lesion-related features extracted from 7 T MR imaging. This highlights not only the importance of UHF MRI for accurately evaluating intracortical and rim lesion load, but also the differential contributions that these types of lesions may bring to determine disease evolution and severity. Also, we found that a small subset of features [WM lesion volume (not considering rim lesions), patient age and WM lesion count (not considering rim lesions), intracortical lesion volume] carried most of the prediction power. Interestingly, an almost opposite pattern emerged when contrasting cortical with WM lesion load: WM lesion load is most important when it is small, whereas cortical lesion load behaves in the opposite way.**

***Clinical Relevance*— Our results suggest that disconnection and axonal degeneration due to WM lesions and local cortical demyelination are the main factors determining cortical thinning. These findings further elucidate the complexity of MS pathology across the whole brain and the need for both statistical and mechanistic approaches to understanding the etiopathogenesis of lesions.**

I. Introduction

Multiple sclerosis (MS) is one of the most common causes of neurological disability in young adults in the Western world [1]. Different radiological features, such as brain Magnetic Resonance Imaging (MRI) "demyelinating" lesions and grey matter (GM) atrophy, are commonly used to diagnose and evaluate disease progression in MS patients [2].

Several MRI studies have shown that GM atrophy arises early in the course of the disease and accelerates with disease progression [3], and additional studies have also demonstrated a strong association between GM atrophy and neurological impairment in MS patients [4] evaluable through functional MRI [5]. As cortical atrophy seems to be the main driver of GM atrophy [6], the understanding of the neurobiological substrates and main determinants of cortical atrophy in MS could be instrumental in predicting disease progression and stratifying the disease subtypes.

Cortical (both intracortical and leukocortical) demyelinated lesions constitute a substantial part of the total lesion load in MS brain [7]. In addition, MS patients may exhibit chronically active white matter lesions, which are identifiable on susceptibility-weighted MR images by their characteristic paramagnetic rim (commonly called "rim lesions") [8], [9]. While both cortical and rim lesions load as well as cortical atrophy are relevant for the diagnosis and the evaluation of MS progression, very little is known about their interplay and, in particular, about how the differential occurrence of one or more types of lesion may be related to cortical atrophy. In this context, it is not clear whether cortical atrophy is mainly the result of local pathological processes or, instead, disconnection from other brain regions which may result by the disruption caused by white matter (WM) lesions. A strong limitation in the investigation of this question is the ability to actually detect and differentiate cortical and rim lesions (as well as of evaluating their extension) when employing MR scanners equipped with static fields with intensities typically found in clinical centers (3T or even 1.5T) [10]. This often allows clinicians to detect and evaluate only a small portion of lesions. In this context, recent studies have demonstrated that ultra-high field (UHF) human MRI (7T) significantly improves *in vivo* imaging of both cortical and rim lesions in MS patients [11], [12]. UHF radiological findings are therefore of strong clinical relevance and may represent the best candidates for investigating the differential role of all lesion types in the progression of cortical atrophy in patients affected by MS.

The aim of this study was to understand the interplay between cortical thickness and different types of lesions, by leveraging radiological markers (cortical and rim lesion load

[*]This work was supported by grants from the National Multiple Sclerosis Society (NMSS 4281-RG-A1 and NMSS RG 4729A2/1), National Institutes of Health R01NS078322-01-A1, and United States Army W81XWH-13-1-0122.

# These authors contributed equally to this work

A.C. and N.T. are with the Department of Biomedicine and Prevention, University of Rome "Tor Vergata", C.A.T., A. M., V.T.B. and C.M. are with Massachusetts General Hospital, Boston, United States. C.M., A. M., V.T.B and N.T. are also with A. A. Martinos Center for Biomedical Imaging, Boston, United States.
Corresponding author: Allegra Conti. E-mail: allegra.conti@uniroma2.it.

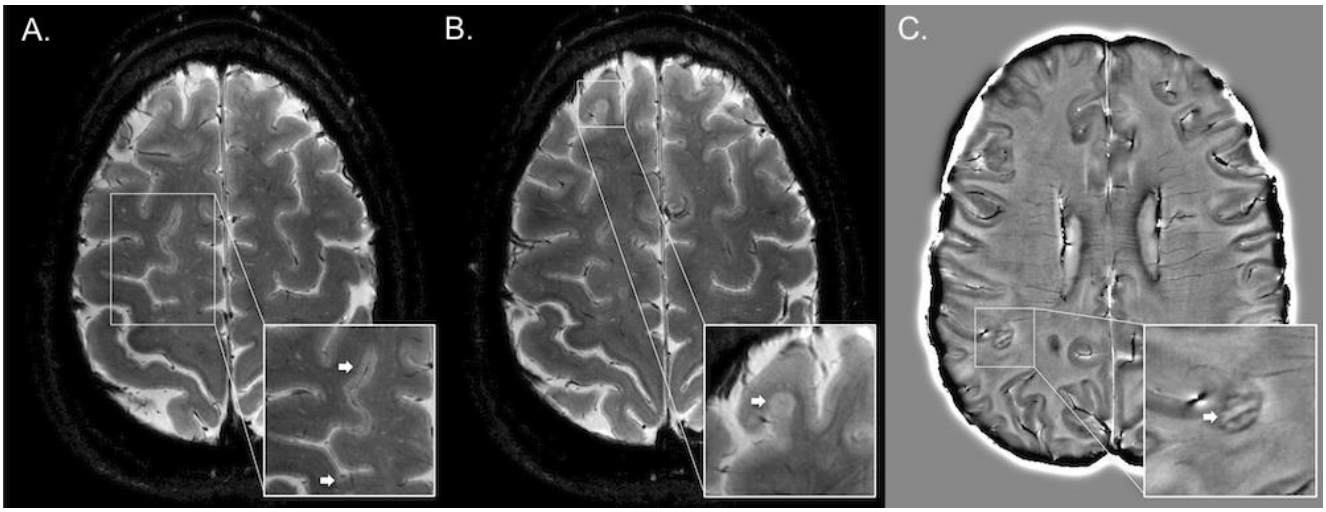

Figure 1. Examples of intracortical (A) and leukocortical (B) lesions on the 7T magnitude T2* images in a 40 years old female with secondary progressive multiple sclerosis. One of the white matter lesions from the same patient is surrounded by a peripheral paramagnetic rim that is easily identifiable on the 7T phase image (C)

MRI metrics) evaluated at UHF MRI as well as conventional global white matter lesion load. To this end, we combine lesions characterization with a recent machine learning (ML) framework which includes explainability, i.e. the ability to rank neuroradiological lesion signatures according to their unique contribution in predicting cortical atrophy.

II. METHODS

A. Patient population

100 MS patients (74 with relapsing remitting MS and 26 with secondary progressive MS; 24 F, 76 M; Age=43±10, Age at onset=33±9) were recruited at Massachusetts General Hospital, in Boston. Figure 2 shows inclusion criteria. The institutional review board approved all protocols, and a written informed consent was obtained from all participants.

B. Magnetic Resonance Imaging

2D- T2*-weighted (T2*-w) MR images (fast low-angle shot [FLASH], TR/TE = 1700/21.8 msec, 0.33x0.3x1 mm^3 resolution) were acquired at 7T to evaluate WM and cortical lesions load as well as the presence of paramagnetic rims at lesions' periphery. Anatomical 3D T1-weighted MR images were acquired on a 3T MRI scanner (TR/TE=2530/1200 msec, 0.9x0.9x0.9 mm^3 resolution) for Freesurfer reconstruction, co-registration with 7 T MR images [13] and regional cortical thickness evaluation in 150 brain regions determined by a predefined parcellation (Destrieux atlas [14]). Mean thickness in single hemispheres and in the whole brain has been calculated by averaging local thickness values.

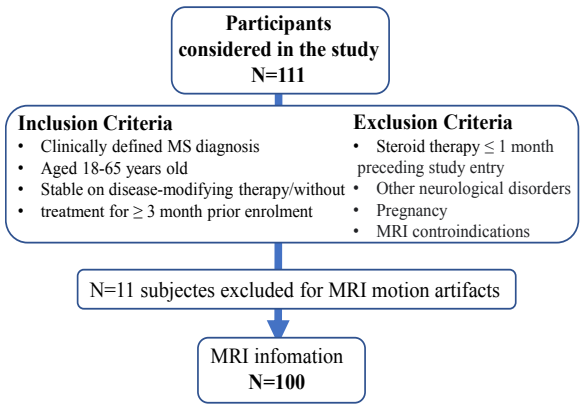

Figure 2 : Study inclusion criteria

C. Lesion Identification

Lesions were segmented with Slicer (version 4.4.0; http://www.slicer.org) by an expert radiologist (CAT) in collaboration with an expert neurologist (CM) with experience in cortical lesion detection. Focal cortical hyperintensities extending for at least 3 voxels across two consecutive slices on magnitude 7T images were classified as intracortical lesions if subpial/confined to the cortex (Fig.1(A)), or leukocortical if they also involved the white matter (Fig.1(B)). Rim lesions were segmented using Slicer on phase images (Fig.1(C)). A MS lesion was defined as "rim lesion" when a "susceptibility rim" (i.e. a hypointense peripheral margin) was detected and was encircling an isointense to extralesional center [15]. FreeSurfer and FSL (version 5.0; http://fsl.fmrib.ox.ac.uk) tools were used to quantify the lesion counts and volumes.

D. Predictive model Development

We designed prediction models for cortical thickness based on 13 demographic and lesional features (gender, patients age, age at onset, rim lesions presence [binary variable, rim lesions/no rim lesions] rim lesions load [binary variable, <4 rim lesions/≥4 rim lesions [16]], rim lesions count and volume, leukortical lesions count and volume, intracortical lesions count and volume, rimless WM lesions count and volume) based on recent gradient boosting technique (Extreme Gradient Boosting-XGBoost). XGBoost provides a parallel tree boosting which has been shown to perform very well in a number of data science problems in a fast and accurate way [17]. We employed an XGBoost classifiers as follows:

i) The original dataset was split randomly into training (70%) and test (30%) sets.
ii) For each training split, a grid search was executed in a 5-fold a cross-validation fashion for hyperparameters

optimization (Hyperparameters optimized: Step size shrinkage used in update to prevents overfitting; maximum depth of a tree; minimum sum of instance weight needed in a child)

iii) After training, performance was assessed in each corresponding (held out) test set by calculating the Pearson correlation (r) as well as related p-values between the real and predicted values.

iv) The above procedure was repeated 50 times, each time sampling the 70/30 split in a random manner. This allowed us to assess the confidence level of the prediction metrics evaluated on the test sets.

Also, for each repetition of the 70/30 split, the unique contribution of each feature to the final prediction performance of the model was evaluated by computing the Shapley Additive explanations (SHAP) values, derived from coalitional game theory [18]. Shap values were derived both for single features as a function of the feature value itself (i.e. for each patient), and as a global average across patients. Both of these metrics were averaged across the 50 repetitions. All predictive regression analyses were implemented in Python 3.6 using the scikit-learn python module [19].

III. RESULTS

Table I shows model performances in the prediction of the mean cortical thickness in the right and left hemispheres as well as in the whole brain. The prediction performance was satisfactory (average $p<0.02$ and average Pearson $r>0.4$ evaluated in 30%-sized test across 50 repetitions) in all experiments, confirming a strong relationship between lesional features and cortical atrophy.

TABLE I. Models performances

Mean Thickness	*r-value*	*p-value*
Right Hemisphere	0.47 (0.15)	0.009 (0.0013)
Left Hemisphere	0.44 (0.18)	0.016 (0.020)
Whole Brain	0.48 (0.17)	0.008 (0.011)

Mean (across 50 repetitions) Pearson correlation (r) as well as related p-values between the real and predicted values evaluated in a 70.30 test/train split. Standard deviations across 50 repetitions are shown in brackets.

Figure 3, depicts an exemplar relationship between real and predicted values cortical thickness averaged across the whole brain (extracted from one single test set).

Figure 4(A-C) shows the resulting SHAP feature importance raking derived from XGBoost model when predicting average thickness in the right-left hemispheres (Fig.4(A-B)), and in the whole brain (Fig.4(C)). These plots list the most significant features in thickness prediction, in descending order. The top variables contribute more to the model than the bottom ones and thus have highest predictive power. The four most important features for the predictions in each hemisphere and in the whole brain were, in order, WM lesion volume (not considering rim lesions), patient age and WM lesion count (not considering rim lesions), intracortical lesion volume.

Finally, Fig. 5 shows partial SHAP dependence plots (median and confidence intervals across repetitions) of the features which displayed the highest rank in terms of contribution to the prediction of globally averaged cortical thickness (Fig. 3(C)). Several insights can be drawn from these plots. For example, the smaller the WM lesions volumes and counts, the higher their importance in predicting mean thickness values, pointing to an almost binary, absence/presence effect. On the other hand, an increase of intracortical lesion volume corresponds to an increase in its importance in predicting atrophy. AS expected, a similar effect is observed with Age.

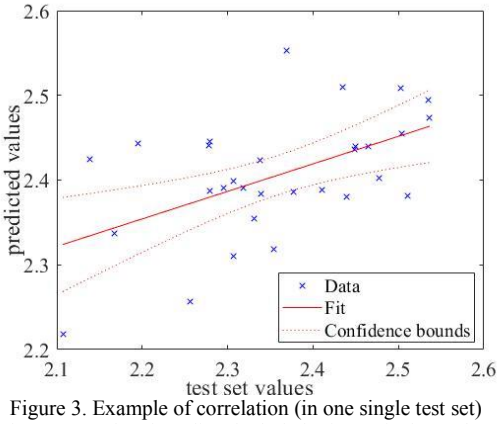

Figure 3. Example of correlation (in one single test set) between real and predicted whole brain cortical atrophy values (r=0.51, p=0,007).

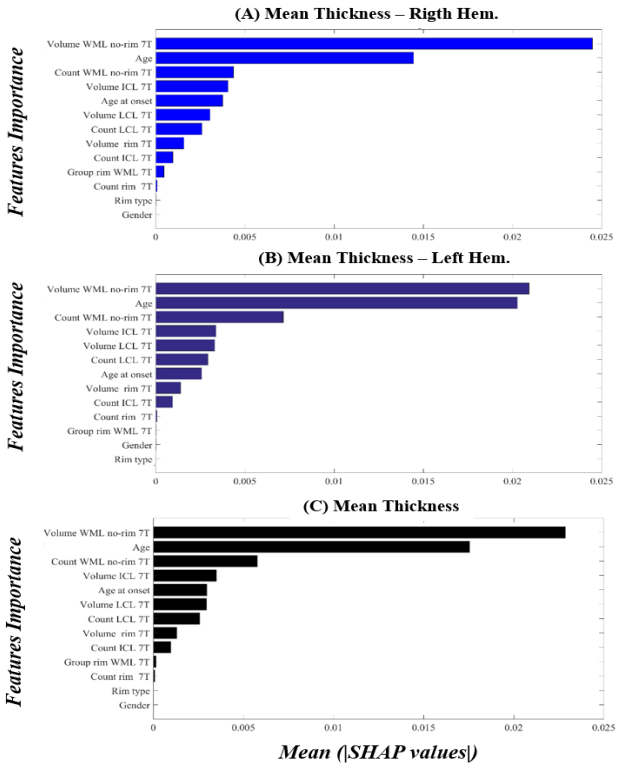

Figure 4. Mean feature importances across 50 repetitions for thickness prediction in the two hemispheres (A,B) and in the whole brain (C)

IV. DISCUSSION AND CONCLUSION

Through an interpretable machine learning approach, we were able to predict cortical thickness in MS from a handful of lesion-related features extracted from ultra-high field imaging (7 T), highlighting not only the importance of UHF MRI for

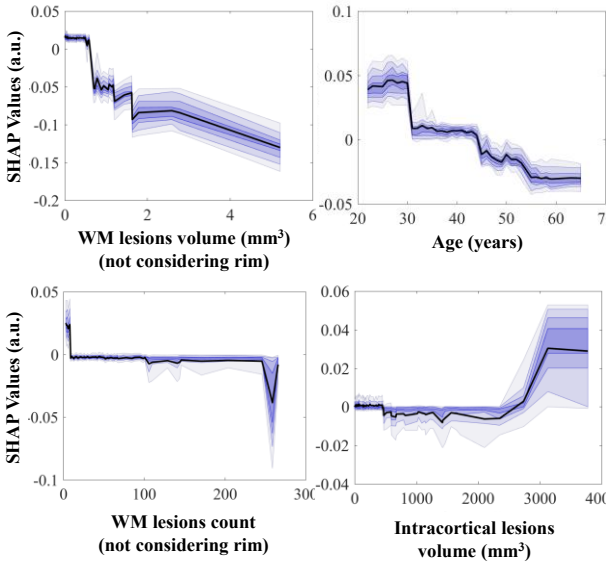

Figure 5. SHAP dependence plots of the four most important features in the prediction of globally averaged mean thickness (in order of importance: WM lesions volume, not considering rim lesions; patient age; WM lesions counts, not considering rim lesions and intracortical lesions volume.

accurately evaluating intracortical and rim lesion load [12], [20], but also the differential contributions that these type of lesions may bring to determining disease evolution and severity. We found that a small subset of features [WM lesion volume (not considering rim lesions), patient age and WM lesion count (not considering rim lesions), intracortical lesion volume] carried most of the prediction power.

The fact that both WM lesions and intracortical lesions were included in those top ranking predictors points to a complex interplay of local pathology and distant disconnection which should be elucidated in targeted studies evaluating e.g. the lesional distance and geometry with respect to the reconstruction of the cortical mantle. However, on the bases of our results we can speculate that disconnection and retrograde Wallerian (e.g. axonal) degeneration due to WM lesions and local cortical demyelination seem to be the main factors determining cortical thinning. Interestingly, rim lesions *per se* are not associated with cortical thinning, meaning that any type of WM lesions could determine disconnection. Locally, neurodegeneration is probably not related to local cortical-cortical disconnection but possibly to a progressive accumulation of pathology.

We also assessed the individual unique contributions of each lesional feature to our predictions. Interestingly, an almost opposite pattern emerged when contrasting cortical with WM lesion load: WM lesion load is most important when it is small, whereas cortical lesion load behaves in the opposite way. This finding further elucidates the complexity of MS pathology across the whole brain and the need for both statistical and mechanistic approaches to understanding the etiopathogenesis of lesions. In the future deep learning techniques tailored to neurodegenerative disorders [21] might prove useful to complement our ML approach.

REFERENCES

[1] M. T. Wallin et al., "Global, regional, and national burden of multiple sclerosis 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016", The Lancet Neurology, vol. 18, n. 3, pagg. 269–285, mar. 2019.

[2] J.M. Honce, "Gray Matter Pathology in MS: Neuroimaging and Clinical Correlations", Multiple Sclerosis International, vol. 2013, Article ID 627870, 16 pages, 2013.

[3] A. Andravizou et al., "Brain atrophy in multiple sclerosis: mechanisms, clinical relevance and treatment options", Auto Immun Highlights, vol. 10, n. 1, pag. 7, dic. 2019.

[4] M. A. Rocca et al., "Brain MRI atrophy quantification in MS: From methods to clinical application", Neurology, vol. 88, n. 4, pagg. 403–413, gen. 2017.

[5] A. Conti et al., "Variability and Reproducibility of Directed and Undirected Functional MRI Connectomes in the Human Brain", Entropy, vol. 21, n.7 pagg. 661

[6] E. Fisher et al., "Gray matter atrophy in multiple sclerosis: a longitudinal study". Ann Neurol, vol. 64, n.3, pagg: 255-265, 2018.

[7] J. J. G. Geurts, L. Bö, P. J. W. Pouwels, J. A. Castelijns, C. H. Polman, e F. Barkhof, "Cortical lesions in multiple sclerosis: combined postmortem MR imaging and histopathology", AJNR Am J Neuroradiol, vol. 26, n. 3, pagg. 572–577, mar. 2005.

[8] P. Maggi et al., "Paramagnetic Rim Lesions are Specific to Multiple Sclerosis: An International Multicenter 3T MRI Study", Ann Neurol, vol. 88, n. 5, pagg. 1034–1042, nov. 2020.

[9] CA Treaba, MD, "Cortical and phase rim lesions on 7 tesla MRI as markers of multiple sclerosis disease progression," Brain Communications, 2021;fcab134, DOI:10.1093/braincomms/fcab134

[10] M. Filippi et al., "Cortical Lesions on 7-T MRI in Multiple Sclerosis: A Window into Pathogenetic Mechanisms?". Radiology. Vol 291, n. 13., pagg: 750-751., june 2019

[11] C. A. Treaba et al., "Longitudinal Characterization of Cortical Lesion Development and Evolution in Multiple Sclerosis with 7.0-T MRI", Radiology, vol. 291, n. 3, pagg. 740–749, giu. 2019.

[12] C. Mainero et al., "In vivo imaging of cortical pathology in multiple sclerosis using ultra-high field MRI", Neurology, vol. 73, n. 12, pagg. 941–948, set. 2009.

[13] R. Guidotti et al., "Optimized 3D co-registration of ultra-low-field and high-field magnetic resonance images", PLoS ONE, vol. 13, n. 3, e0193890, mar. 2018

[14] C. Destrieux et al., "Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. ", NeuroImage vol. 53, n.1,pagg.1-15, (2010).

[15] B. Yao et al., "Chronic Multiple Sclerosis Lesions: Characterization with High-Field-Strength MR Imaging", Radiology, vol. 262, n. 1, pagg. 206–215, gen. 2012.

[16] M. Absinta, et al, "Association of Chronic Active Multiple Sclerosis Lesions With Disability In Vivo. " JAMA neurology vol. 76, n.12, pagg.1474-1483. 2019.

[17] T. Chen e C. Guestrin, "XGBoost: A Scalable Tree Boosting System", in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, ago. 2016, pagg. 785–794.

[18] S. Lundberg e S.-I. Lee, "A Unified Approach to Interpreting Model Predictions", arXiv:1705.07874 [cs, stat], nov. 2017, Consultato: mag. 01, 2021. [Online]. Disponibile su: http://arxiv.org/abs/1705.07874.

[19] Pedregosa F., et al., "Scikit-learn: Machine Learning in Python", The Journal of Machine Learning Research. Vol 12, pagg. 2825-2830, Nov. 2011

[20] C. Mainero et al., "A gradient in cortical pathology in multiple sclerosis by in vivo quantitative 7 T imaging. Brain." Vol. 138, n. 14, pagg.932–945. 2015

[21] S. Spasov et al., "A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease" Neuroimage. vol. 1, n. 189, pagg: 276-287. Apr 2019