



## OPEN ACCESS

## EDITED BY

Fahmi Khalifa,  
Morgan State University, United States

## REVIEWED BY

Sneha Pandya,  
Cornell University, United States  
João Valente Duarte,  
University of Coimbra, Portugal

## \*CORRESPONDENCE

Antonio Ricciardi  
✉ antonio.ricciardi@ucl.ac.uk

RECEIVED 03 October 2022

ACCEPTED 28 February 2023

PUBLISHED 23 March 2023

## CITATION

Ricciardi A, Grussu F, Kanber B, Prados F, Yiannakas MC, Solanky BS, Riemer F, Golay X, Brownlee W, Ciccarella O, Alexander DC and Gandini Wheeler-Kingshott CAM (2023) Patterns of inflammation, microstructural alterations, and sodium accumulation define multiple sclerosis subtypes after 15 years from onset. *Front. Neuroinform.* 17:1060511. doi: 10.3389/fninf.2023.1060511

## COPYRIGHT

© 2023 Ricciardi, Grussu, Kanber, Prados, Yiannakas, Solanky, Riemer, Golay, Brownlee, Ciccarella, Alexander and Gandini Wheeler-Kingshott. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Patterns of inflammation, microstructural alterations, and sodium accumulation define multiple sclerosis subtypes after 15 years from onset

Antonio Ricciardi<sup>1\*</sup>, Francesco Grussu<sup>1,2</sup>, Baris Kanber<sup>1,3</sup>, Ferran Prados<sup>1,3,4</sup>, Marios C. Yiannakas<sup>1</sup>, Bhavana S. Solanky<sup>1</sup>, Frank Riemer<sup>5</sup>, Xavier Golay<sup>6</sup>, Wallace Brownlee<sup>1</sup>, Olga Ciccarella<sup>1,7</sup>, Daniel C. Alexander<sup>3</sup> and Claudia A. M. Gandini Wheeler-Kingshott<sup>1,8,9</sup>

<sup>1</sup>NMR Research Unit, Queen Square MS Centre, Department of Neuroinflammation, UCL Queen Square Institute of Neurology, Faculty of Brain Sciences, University College London, London, United Kingdom, <sup>2</sup>Radiomics Group, Vall d'Hebron Institute of Oncology, Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain, <sup>3</sup>Centre for Medical Image Computing, Department of Medical Physics and Biomedical Engineering, University College London, London, United Kingdom, <sup>4</sup>eHealth Center, Universitat Oberta de Catalunya, Barcelona, Spain, <sup>5</sup>Mohn Medical Imaging and Visualization Centre, Department of Radiology, Haukeland University Hospital, Bergen, Norway, <sup>6</sup>Department of Brain Repair and Rehabilitation, UCL Queen Square Institute of Neurology, Faculty of Brain Sciences, University College London, London, United Kingdom, <sup>7</sup>NIHR UCLH Biomedical Research Centre, London, United Kingdom, <sup>8</sup>Department of Brain and Behavioural Sciences, University of Pavia, Pavia, Italy, <sup>9</sup>Brain Connectivity Research Center, IRCCS Mondino Foundation, Pavia, Italy

**Introduction:** Conventional MRI is routinely used for the characterization of pathological changes in multiple sclerosis (MS), but due to its lack of specificity is unable to provide accurate prognoses, explain disease heterogeneity and reconcile the gap between observed clinical symptoms and radiological evidence. Quantitative MRI provides measures of physiological abnormalities, otherwise invisible to conventional MRI, that correlate with MS severity. Analyzing quantitative MRI measures through machine learning techniques has been shown to improve the understanding of the underlying disease by better delineating its alteration patterns.

**Methods:** In this retrospective study, a cohort of healthy controls (HC) and MS patients with different subtypes, followed up 15 years from clinically isolated syndrome (CIS), was analyzed to produce a multi-modal set of quantitative MRI features encompassing relaxometry, microstructure, sodium ion concentration, and tissue volumetry. Random forest classifiers were used to train a model able to discriminate between HC, CIS, relapsing remitting (RR) and secondary progressive (SP) MS patients based on these features and, for each classification task, to identify the relative contribution of each MRI-derived tissue property to the classification task itself.

**Results and discussion:** Average classification accuracy scores of 99 and 95% were obtained when discriminating HC and CIS vs. SP, respectively; 82 and 83% for HC and CIS vs. RR; 76% for RR vs. SP, and 79% for HC vs. CIS. Different patterns of alterations were observed for each classification task, offering key insights in the understanding of MS phenotypes

pathophysiology: atrophy and relaxometry emerged particularly in the classification of HC and CIS vs. MS, relaxometry within lesions in RR vs. SP, sodium ion concentration in HC vs. CIS, and microstructural alterations were involved across all tasks.

#### KEYWORDS

**MRI, multiple sclerosis, quantitative, multi-modal, diffusion, sodium, machine learning, random forest**

## 1. Introduction

Multiple sclerosis (MS) is an immune-mediated, inflammatory, neurodegenerative disease of the central nervous system characterized by inflammatory demyelination and heterogeneous accrual of physical disability (Lucchinetti et al., 2000). The onset is determined by the first inflammatory episode suggestive of MS, referred to as clinically isolated syndrome (CIS), with CIS being recognized as the first clinical instance in the MS spectrum (Lublin et al., 2014). Further neurological symptoms may lead to a clinically defined diagnosis, as determined by the updated McDonald criteria (Thompson et al., 2018). Based on the clinical course, patients can be categorized into three types of relapse-onset MS: CIS, relapsing remitting (RR) and secondary progressive (SP). RR is characterized by clinically defined focal activity followed by periods of total or partial remission of neurological deficit, and the lack of disease progression between attacks; SP may follow from an initial RR course, with progressive worsening of neurological symptoms, with or without acute relapses. Primary progressive MS is associated with a progressive deterioration of clinical symptoms from onset (Lublin et al., 2014). Understanding why patients may develop different MS phenotypes over the years, or why only a small fraction of the diversity of clinical disability in MS can be explained by radiological evidence (*clinico-radiological paradox*) (Barkhof, 1999, 2002), are cause for further research.

Magnetic resonance imaging (MRI) is instrumental in the diagnosis and prognosis of MS, routinely used in clinical practice for the acquisition of qualitative images, e.g., proton density-(PD), T2- and T1-weighted, for lesion assessment. In the research environment, a much wider spectrum of dedicated, quantitative MRI techniques are employed for the study and characterization of MS pathophysiology, investigating the complex relationship between radiological evidence and clinical disability (Chard and Trip, 2017; Filippi et al., 2019). *In vivo* imaging biomarkers can be sensitive to inflammation, microstructural alterations, and even sodium ions accumulation, providing a window into the disease pathophysiology over time. Brain atrophy is a known indicator of disease progression since the early stages of MS, with recent studies providing further insight into the hierarchical recruitment of different brain regions over time (Eshaghi et al., 2018), although the integration of specifically cortical and sub-cortical regional volumetric measurements in clinical practice has yet to reach a consensus (Sastre-Garriga et al., 2020). Relaxometry and quantitative PD imaging have been shown to provide good biomarkers for inflammation and demyelination in normal appearing tissue (Neema et al., 2007; Mezer et al., 2013), invisible to the standard qualitative imaging. Through sensitivity

to the diffusion of water molecules within the structured axonal environment of the brain, diffusion weighted imaging (DWI) has shown microstructural alterations in both lesions and normal appearing tissues, correlating with physical disability in progressive MS (Filippi et al., 2001; Collorone et al., 2020); recent studies have reported abnormalities at the early stages (Tur et al., 2020) and potential links to cognitive disability as well (Savini et al., 2019). Sodium ( $^{23}\text{Na}$ ) imaging has been used to access the signal induced by sodium ions, showing promise in probing axonal function directly (Gandini Wheeler-Kingshott et al., 2018), with evidence of increased total sodium concentration (TSC) being reported in MS, correlating with disability and disease progression (Inglese et al., 2010; Paling et al., 2013; Maarouf et al., 2014). Whilst the potential of advanced MRI modalities is evident, they lack a unified consensus about their implementation, optimization and interpretation, and require, when compared to standard routine scans, additional acquisition times, costs and expertise, which make their application in clinics limited.

In this work, we explored a multi-modal dataset acquired in a cohort of patients with the same disease duration, where clinical and MRI assessments were performed 15 years from CIS, comprehending both routine and advanced MRI metrics sensitive to inflammation, microstructural alteration and sodium ions accumulation. Using a machine learning approach, we aimed to gain further understanding of which modalities are more likely to carry biophysically meaningful information for different classification tasks. Machine learning indeed has shown to be a key tool in the data-driven exploration of MRI datasets for the identification of patterns and biomarkers of disease, including the ability to identify discriminating factors of disease phenotypes against each other and healthy controls (HC) (Wottschel et al., 2015; Eshaghi et al., 2016). We therefore trained and tested a *random forest* algorithm to classify different subtypes of MS vs. HC and between each other, using a rich array of quantitative imaging features extracted from both clinical and advanced MRI data. Feature importance was calculated for each task and used to assess which metrics mostly contributed to the decision-making process. This provided us with novel insights into the pathophysiology of different MS subtypes, while also informing future studies toward more task-efficient MRI acquisitions.

## 2. Methods

A retrospective (Brownlee et al., 2019) multi-modal MRI dataset of HC, CIS, RR, and SP patients with same disease duration

was analyzed to provide evidence on what MRI features are best representative for different classification tasks.

## 2.1. Cohort

The cohort consisted of a total of 123 subjects: 29 HC (10 men, age:  $35 \pm 10$  years old), 18 CIS (6 men, age:  $47 \pm 10$  years old, EDSS:  $0.4 \pm 0.5$ ), 63 RR (15 men, age:  $47 \pm 8$  years old, EDSS:  $2.2 \pm 1.1$ ), and 13 SP (4 men, age:  $48 \pm 8$  years old, EDSS:  $5.5 \pm 1.2$ ). All MS patients (CIS, RR, SP) attended the MS center for clinical and radiological follow-up after a mean of 15 years from onset (Brownlee et al., 2019).

## 2.2. MRI protocol

Data were acquired on a 3T Philips Achieva MR system. The acquisition protocol included:

- PD/T2-w.** Dual-echo 2D PD/T2-weighted turbo spin-echo (resolution:  $1 \times 1 \times 3\text{mm}^3$ , echo time TE: 19/85ms, repetition time TR: 3,500ms, turbo factor: 10, echo spacing: 9.4ms, scan time: 4'2").
- T1-w.** 2D T1-weighted spin-echo (resolution:  $1 \times 1 \times 3\text{mm}^3$ , TE: 10ms, TR: 625ms, scan time: 5'43").
- DWI.** Cardiac-gated, multi-shell, diffusion-weighted echo-planar imaging, with {8, 15, 30} isotropically distributed directions at  $b$ -values: {300, 711, 2000}  $\text{s/mm}^2$  (resolution:  $2.3 \times 2.3 \times 2.3\text{mm}^3$ , TE: 82ms, nominal TR (12 heart-beats): 13846ms, scan time:  $\sim 16'$ ).
- Sodium.**  $^{23}\text{Na}$  imaging with 3D-cone sampling trajectory (resolution:  $3 \times 3 \times 3\text{mm}^3$ , TE: 0.22ms, TR: 120ms, scan time:  $\sim 18'$ ). Two 4% agar phantoms with sodium concentration of 40 and 80mM were placed near the subject's head during the image acquisition for calibration purposes (Riemer et al., 2014).
- 3DT1.** 3D sagittal T1-weighted magnetization-prepared rapid gradient echo (resolution:  $1 \times 1 \times 1\text{mm}^3$ , TE: 3.1ms, TR: 6.9ms, inversion delay time: 823ms, flip angle:  $8'$ , scan time: 6'32").

All proton scans were acquired using a 32 channel head coil, whilst sodium imaging was performed using a single channel transmit-receive volume head coil (Rapid Biomedical, Rimpur, Germany). Patients were repositioned prior to the sodium imaging scans to allow for the coil change.

## 2.3. Image analysis

Lesion masks from Brownlee et al. (2019) studies were used. Brain tissue segmentation was performed on lesion-filled (Prados et al., 2016a) 3DT1 using the *Geodesic Information Flows* (GIF) tool (Cardoso et al., 2015), obtaining masks of white matter (WM), deep gray matter (dGM), and cortical gray matter (cGM).

PD/T2-weighted scans were initially acquired for lesion segmentation only; however, given the availability of T1-weighted scans with similar readout, they were also used to extract quantitative estimates of PD, T2, and T1 maps by fitting the relevant

Bloch equations, using the *MyRelax* toolbox (Grussu et al., 2020). Further details are reported in the [Supplementary material](#).

DWI data were corrected for motion and eddy current distortion using *FSL* (Andersson and Sotiropoulos, 2016). The *spherical mean technique* (SMT) multi-compartment model (Kaden et al., 2016) was used to analyse the DWI data, producing maps of intra-neurite volume fraction, intrinsic diffusivity and orientation dispersion entropy.

TSC maps were calculated by calibrating the  $^{23}\text{Na}$  images by the signal intensity within the phantoms (Inglese et al., 2010; Riemer et al., 2014), which were segmented automatically (Prados et al., 2016b).

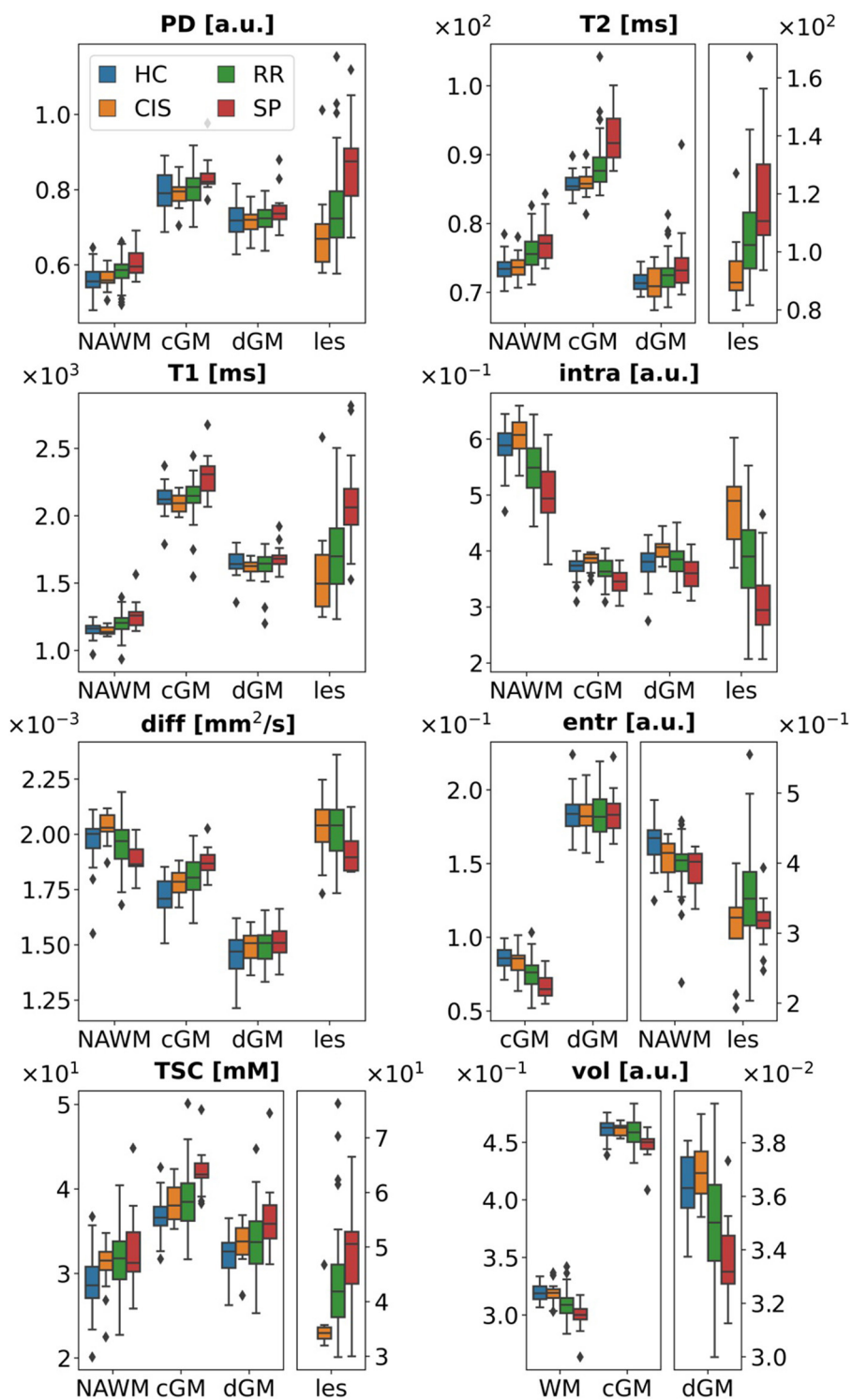
Mean values for quantitative PD, T2, and T1, intra-neurite volume fraction, intrinsic diffusivity, entropy, and TSC were calculated in normal appearing white matter (NAWM), dGM, cGM, and lesions, when present. Details about the calculation of the summary statistics are reported in the [Supplementary material](#). The volume for the three tissue classes (WM, dGM, cGM) was also calculated from the brain segmentation, and divided by the total intra-cranial volume to take into account variability in head-sizes. In total, for each of the 123 subjects, 31 regional variables, or *biophysically meaningful features*, were therefore calculated.

Due to the HC group being significantly younger than the rest (average age 12 years lower,  $p = 1 \times 10^{-6}$  from Kruskal–Wallis test), all features other than lesions-based ones were corrected for age using the HC as reference. A linear model, with age as independent variable and  $\{\beta_0, \beta_1\}$  as intercept and slope, respectively, was fitted feature-wise on the HC data: the features that resulted significantly ( $p < 0.05$ ) correlated with age were corrected by subtracting  $\beta_1 \times \text{age}$  from the original data.

## 2.4. Classification analysis

After correcting for age, the data was *standardized* feature-wise such that the value distribution for each feature had mean of zero and standard deviation of one. The dataset was used to train and test a random forest algorithm over different binary classification tasks: HC vs. MS (that is RR and SP), CIS vs. MS, and all binary permutations of HC, CIS, RR, and SP. All HC and most CIS had no lesions, therefore lesion features were included only for the RR vs. SP classification task.

Classification was implemented using Python 3.7.4 (VanRossum and Drake, 2010) and the *scikit-learn* package (Pedregosa et al., 2011). Default parameters for the `ensemble.RandomForestClassifier` function were selected, with the number of trees set to 1,000 based on the available literature and previous experience on datasets with similar dimensionality. For each classification task, a 10-fold stratified cross-validation with 10 repetitions was implemented, for a total of 100 iterations. The classification performance was assessed by the average *receiver operating characteristic* (ROC) *area under the curve* (AUC) score on the test set across the 100 train/test iterations. Variable importance is defined by the improvement in the split-criterion attributed to each variable (feature) during training of the random forest. Variable importances were averaged across iterations, returning the mean feature ranking for the task;



**FIGURE 1**  
 Age-corrected data. Data points for some tissue types have been plotted against different y-axes to better visualize boxplots across different ranges. PD, proton density; intra, intra-neurite volume fraction; diff, intrinsic diffusivity; entr, orientation dispersion entropy; TSC, total sodium concentration; vol, tissue volume; WM, white matter; NAWM, normal appearing white matter; cGM, cortical gray matter; dGM, deep gray matter; les, lesions; a.u., arbitrary units.

this allowed to identify the features that most contributed to each classification task, and thus are more likely to be biophysically meaningful with respect to the groups characterization.

In order to assess the significance of the classification results, the training and testing process was repeated identically 1,000 times with randomly permuted labels of the subjects at each repetition.

TABLE 1 ROC AUC classification results.

Tasks	Mean	Median [Q <sub>1</sub> , Q <sub>3</sub> ]	p-value
HC - RR	0.82	0.83[0.72, 0.94]	< 0.001
HC - SP	0.99	1.00[1.00, 1.00]	< 0.001
HC - MS	0.84	0.88[0.76, 0.92]	< 0.001
CIS - RR	0.83	0.89[0.75, 1.00]	< 0.001
CIS - SP	0.95	1.00[1.00, 1.00]	< 0.001
CIS - MS	0.84	0.87[0.71, 1.00]	< 0.001
RR - SP	0.76	0.83[0.67, 1.00]	< 0.01
HC - CIS	0.79	0.83[0.67, 1.00]	< 0.01

p-values calculated through permutation test. HC, healthy controls; CIS, clinically isolated syndrome; RR, relapsing-remitting MS; SP, secondary progressive MS; MS, RR and SP; Q<sub>1,3</sub>, 1st, 3rd quartile, respectively.

The distribution of the 1,000 mean ROC AUC scores defined the random classifier performance profile, which was used as reference to calculate the p-value associated to the classification performances on the original data.

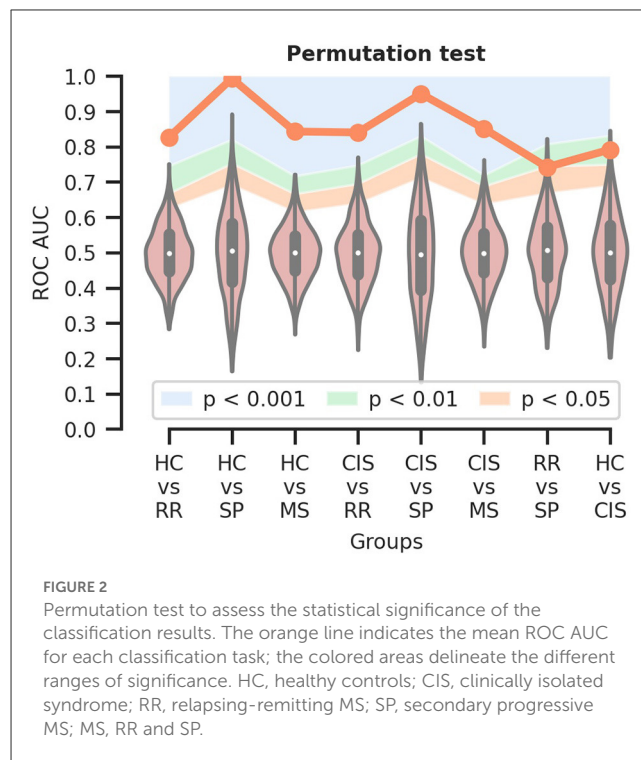
## 3. Results

### 3.1. Age correction

Of the 24 non-lesion features, 5 resulted significantly correlated with age: quantitative T2 in dGM ( $\beta_1 = -0.09$ ,  $p = 0.002$ ), and cGM ( $\beta_1 = -0.07$ ,  $p = 0.02$ ), intrinsic diffusivity in dGM ( $\beta_1 = 5 \times 10^{-6}$ ,  $p = 0.02$ ), volume of dGM ( $\beta_1 = 6.6 \times 10^{-5}$ ,  $p = 0.006$ ) and cGM ( $\beta_1 = 4.8 \times 10^{-4}$ ,  $p = 0.009$ ). Fitting results for all features are reported in the [Supplementary material](#). Age-corrected feature distributions are shown in [Figure 1](#).

### 3.2. Classification results

ROC AUC scores for each task are reported in [Table 1](#). In addition to the mean, the median and interquartile range [Q<sub>1</sub>, Q<sub>3</sub>], with Q<sub>1,3</sub> indicating the 25-th and 75-th percentiles respectively, are also reported to assess dispersion instead of standard deviation, as the ROC AUC distribution over the 100 iterations was not symmetric, but skewed toward better-than-chance performance values. The best classification performances were obtained for the HC vs. SP and CIS vs. SP tasks, with mean ROC AUC scores of 0.99 and 0.95, respectively. Mean ROC AUC scores for HC vs. RR and CIS vs. RR were 0.82 and 0.83, and when discriminating HC and CIS against both the clinically defined MS groups, the performance scores fell in between. The lowest scores were observed for the RR vs. SP and HC vs. CIS tasks, with mean ROC AUC scores of 0.76 and 0.79, respectively. Mean ROC AUC, sensitivity and specificity scores have been also calculated with a random under-sampling method to correct for group imbalance, and reported in the [Supplementary material](#).



### 3.3. Permutation test

The random classifier performance profiles for the different tasks are shown in [Figure 2](#). Statistical significance of  $0.001 < p < 0.01$  was observed for the RR vs. SP and HC vs. CIS classification tasks, whilst  $p < 0.001$  was recorded for all others.

### 3.4. Feature importance

Average feature ranking for all classification tasks is shown in [Figures 3, 4](#). Features have been color-coded to group MRI modalities from the same biophysical source (e.g., relaxometry parameters are in orange, diffusion microstructure in blue, sodium concentrations in green and tissue volume in gray); the top-ranking features contributing to 50% of the decision process for each task have been highlighted by a striped block and considered for interpretation. Overall, tissue volumes were the most meaningful when discriminating HC and CIS against clinically-defined MS; relaxometry parameters mainly in lesions had a role when discriminating the clinically-defined MS subtypes against each others; diffusion metrics were meaningful across all tasks, although particularly prominent in differentiating CIS vs. HC; TSC was mostly relevant when discriminating CIS against HC.

- **HC vs. RR.** RR patients showed a reduced WM and dGM volume, as well as reduced orientation dispersion entropy and increased T2 in NAWM and cGM with respect to HC. Increased intrinsic diffusivity in cGM also contributed to the classification task, with reduced intra-neurite volume fraction in NAWM at the 50% cumulative importance threshold.

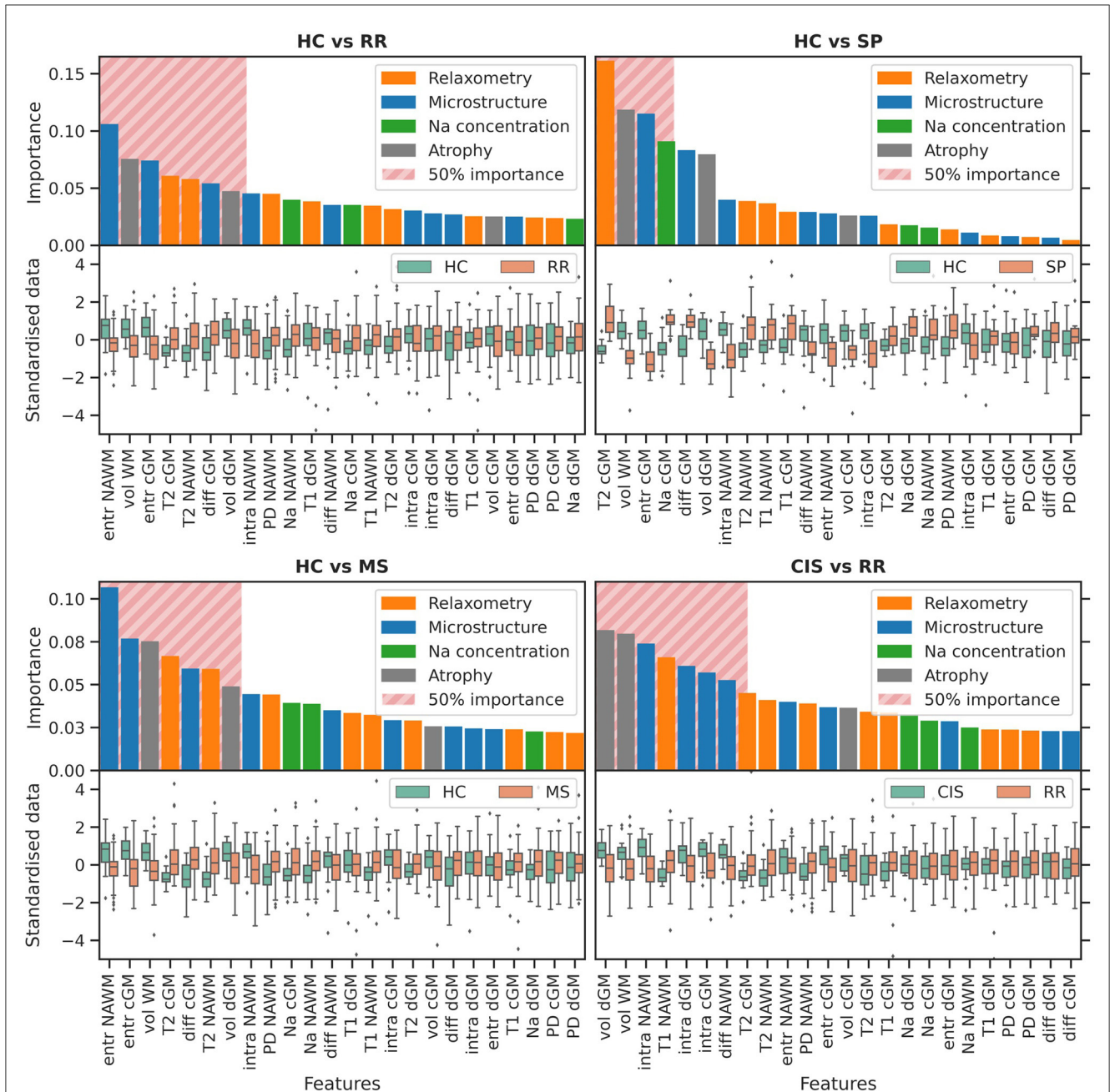
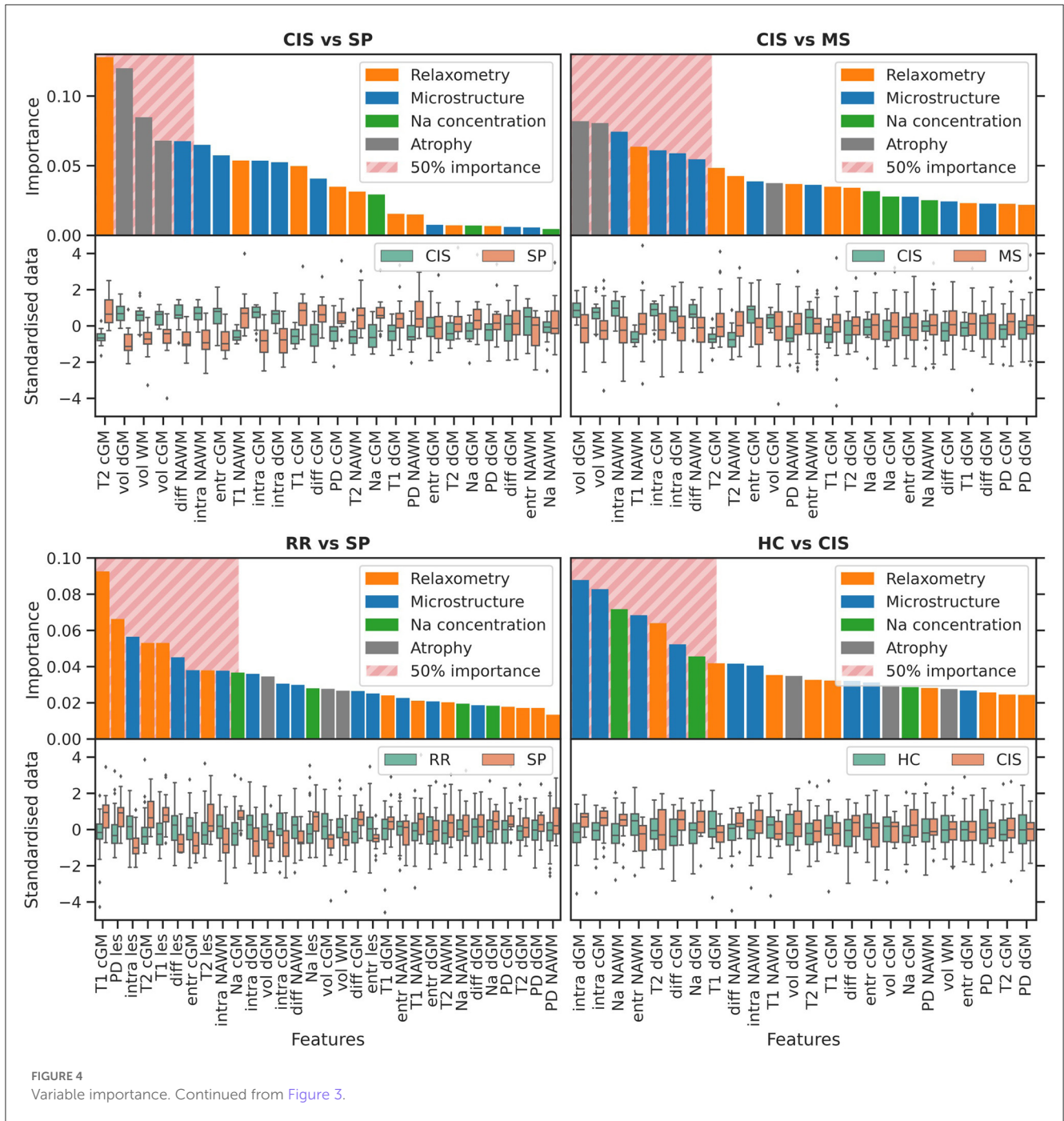


FIGURE 3

Variable importance. The bar-plot is color-coded to group MRI modalities from the same biophysical source. The striped block highlights the features contributing to 50% of the classification process. The y-axis is shared between columns. PD, proton density; intra, intra-neurite volume fraction; diff, intrinsic diffusivity; entr, orientation dispersion entropy; TSC, total sodium concentration; vol, tissue volume; WM, white matter; NAWM, normal appearing white matter; cGM, cortical gray matter; dGM, deep gray matter; les, lesions; HC, healthy controls; CIS, clinically isolated syndrome; RR, relapsing-remitting MS; SP, secondary progressive MS; MS, RR and SP. Continuing to Figure 4.

- **HC vs. SP.** The decision task was mostly driven by increased T2 in cGM, and reduced WM volume and entropy in cGM of SP compared to HC. Increased TSC and diffusivity in cGM were also observed in SP at the 50% threshold.
- **HC vs. MS.** Top-ranking features were distributed similarly to the HC vs. RR task.
- **CIS vs. RR.** Reduced dGM and WM volume mostly characterized RR compared to CIS, together with reduced intra-neurite volume fraction across all tissues, and increased T1 in NAWM. Reduced intrinsic diffusivity also emerged in NAWM in RR.
- **CIS vs. SP.** The task was driven mostly by reduced volume of all tissues in SP, and increased T2 in cGM. Reduced diffusivity and intra-neurite volume fraction in NAWM was also observed at the 50% threshold in SP compared to CIS.
- **CIS vs. MS.** Similar top-ranking features to the CIS vs. RR task were observed.



- **RR vs. SP.** The task was driven mostly by relaxometry—increased PD, T2 and T1—and diffusion metrics—reduced intra-neurite volume fraction, diffusivity—alterations in lesions in SP compared to RR. Increased T1, T2 and entropy in cGM were also observed, with increased TSC in NAWM at the 50% threshold.
- **HC vs. CIS.** Increased TSC in NAWM and dGM in CIS compared to HC appeared as top-ranking features, together with increased intra-neurite volume fraction in dGM and cGM, reduced intra- in NAWM, and increased diffusivity in cGM. Reduced T2 and T1 in dGM were also observed in CIS.

## 4. Discussion

In this work, we used random forest classifiers to study the interaction between features extracted from both routine scans and advanced diffusion and sodium weighted imaging for the purpose of characterizing CIS and clinically-defined MS subtypes. The results show that a combination of advanced quantitative MRI and clinical features achieve classification scores between 76% and 99% depending on the task. Moreover, it is apparent that after 15 years from the initial CIS episode, features reflecting inflammation, microstructure changes and sodium accumulation play a very

different role in each MS subtype. Whether these alterations are present from the start or are the result of a 15 years evolution, it is not possible to say from this dataset, and requires targeted longitudinal studies.

- **Atrophy.** Brain atrophy was observed in the MS groups and emerged as a meaningful feature in all the classification tasks discriminating MS patients against HC and CIS. No strong involvement of brain volume features was observed instead when classifying MS phenotypes against each other, and in the HC vs. CIS task. Tissue volume loss in MS showed heterogeneous behavior across the different tissues, particularly with dGM consistently scoring higher than cGM, which is in line with previous findings of progressive recruitment of gray matter structures as part of MS neurodegeneration (Eshaghi et al., 2018; Soares et al., 2020). dGM significant involvement is well-known in the scientific community, however a consensus for the incorporation of *global* gray matter volumetrics into clinical practice has only recently been reached, and the specific inclusion of dGM structures (e.g., thalami, basal ganglia) in particular is still debated (Sastre-Garriga et al., 2020). Further research is recommended, with this work providing evidence in this direction.
- **Relaxometry.** Relaxometry features contributed partially to most tasks involving the clinically defined MS population, with prolonged T2 in cGM emerging particularly when discriminating SP against HC and CIS, a possible biomarker for the advanced cortical demyelination observed in the progressive stages of MS (Magliozzi et al., 2018). The strongest contribution was however observed in the RR vs. SP classification task across all relaxometry features, both in cGM and lesions: the involvement of quantitative parameters measured in lesions in the discrimination of MS phenotypes is indicative of the heterogeneous nature of MS pathogenesis and neurodegeneration (Lucchinetti et al., 2000). This result supports the need for adopting a more quantitative approach to lesion characterization in clinics than mere lesion load assessment. Reduced T2 in dGM was also observed in CIS with respect to HC: this reduction goes against a possible demyelination effect and could be due to residual (after age correction) iron deposition (Aquino et al., 2009). Further studies of iron deposition in MS, using for example quantitative susceptibility mapping and magnetic susceptibility source separation approaches, are recommended (Shin et al., 2021).
- **Diffusion imaging.** Diffusion imaging metrics were involved in all classification tasks, which is expected given that microstructural alterations are at the core of MS demyelination and neurodegeneration. Despite being the most ubiquitous set of feature across all tasks, diffusion metrics from multi-compartment models are also strongly model-dependent, hence prone to modeling artifacts and limitations, e.g., the lack of a myelin compartment, and results should be interpreted with care. Against HC, MS patients exhibited overall reduced orientation dispersion entropy, and reduced intra-neurite volume fraction against CIS.

These results are in line with findings of reduced *fractional anisotropy* associated to higher fiber dispersion and neurite loss (Roosendaal et al., 2009). Reduced intrinsic diffusivity in SP was also observed with respect to CIS and RR, both in NAWM and lesions, but not in gray matter, which may be spurious, or an indicator of new lesion formation compatible with axonal undulation (Grussu et al., 2016). Increased intra-neurite volume fraction in CIS emerged in the classification against HC: whilst counter-intuitive in the context of neurite loss, this may be suggestive of axonal swelling, as further discussed below.

- **Sodium imaging.** TSC was particularly meaningful when discriminating CIS vs. HC, with increased TSC being observed in CIS in NAWM and dGM. Increased TSC was also observed in MS patients, albeit with a lower contribution with respect to other features, and has been reported in literature from the early stages of the disease (Maarouf et al., 2017). It has been associated with the over-expression and redistribution of sodium-potassium channels from the Ranvier nodes to newly demyelinated membrane: this is an adaptive response to the disruption of saltatory conduction caused by demyelination, apt to preserve action potential transmission, limit the onset of neurological deficits, and facilitate recovery. This however also increases the axonal metabolism, as the proliferation of the sodium-potassium *active* pumps comes with higher energy expenditure which, if not satisfied, causes the accumulation of intra-cellular sodium. In MS, the impaired trophic support from oligodendrocytes and mitochondrial dysfunction contribute to energy underproduction which, coupled with the increased metabolic need, can lead to axonal degeneration due to metabolic failure secondary to chronic energy deprivation (Petracca et al., 2016). In the case of CIS, the increased TSC might be explained as a long lasting effect established in the brain following the initial inflammatory event. It may be speculated that this might be due to an over-expression of sodium channels at the onset of CIS to support neuronal function, which may also explain the increased intra-neurite volume fraction detected with diffusion imaging: indeed, the intra-cellular accumulation of sodium might induce axonal swelling through osmosis (Armstrong, 2003). To what extent this can happen before functional derangement accrues, leading to a more severe MS phenotype, is to be investigated.

Overall, the results of this unique dataset with MS patients of same disease duration, and a rich multi-modal quantitative MRI protocol, have shown that atrophy and relaxometry features contribute significantly to the discrimination of MS patients from HC and CIS; relaxometry in lesions emerges as particularly involved in the classification of MS phenotypes, which highlights the heterogeneity of MS pathophysiology. With both brain volumetry and relaxometry features being extracted from routine scans readily available in clinical practice, we have offered evidence of the hidden potential qualitative MRI data holds beyond lesion and tissue segmentation. Whilst advanced MRI acquisitions ought to be preferred when available, they are far from being routinely introduced in clinics; on the other hand, the use of routine scans



can pave the way to quantitative studies on large historical datasets otherwise lacking dedicated quantitative modalities. Advanced diffusion and sodium imaging have proven particularly sensitive to the characterization of MS phenotypes against each other, and CIS against HC, where differences in atrophy or relaxometry scores in normal appearing tissues were not as important, or not present at all. In these cases, dedicated quantitative MRI modalities showed their role in the quantification of subtle tissue microstructural and physiological alterations, otherwise invisible to conventional MRI, offering further insights on MS heterogeneous neurodegeneration. Specifically, CIS presenting subtle alterations compatible with MS histopathology (sodium ions accumulation and possible axonal swelling) may mark long lasting subtle damage accrued as a result of the first episode of neurologic symptoms. Alternatively, the observed alterations suggest that neuroprotective mechanisms may be at play in the stable CIS population, but, unlike with the clinically-defined MS patients, they do not lead to meaningful atrophy, inflammation, demyelination, and axonal loss. In other words, the ability to adapt to the increased metabolic demand without succumbing to energy failure, or avoiding axonal degeneration by excessive osmotic swelling, might be compensatory or even protective mechanisms, and as such key factors in what determines conversion, or lack thereof, to clinically defined MS.

Interestingly, what differentiates RR from a progressive form of the disease characterizing SP are changes in relaxometry parameters in the lesions. The classification task ranked as highest not microstructure changes or sodium accumulation in normal appearing tissue, but alterations of relaxometry parameters in the lesions of SP patients compared to RR. This could give an insight into the possible source of disease progression, driven not by the number or volume of the lesions, nor by diffuse damage of tissue, but by the severity and biophysical nature of lesion alterations. This therefore calls on monitoring relaxometry, as well as potentially others quantitative biophysically meaningful features, in the lesions as potential predictor of risks of progression.

The interpretation of these findings is of course conditional on this study's limitations. The statistical significance is hindered by the small sample size, especially tasks involving the SP group—only 13 subjects: although spurious results due to the many features may be expected, we strove to minimize their impact on the final outcome through careful examination of the data and corroboration with the published literature. In terms of classification tasks, the class imbalance between RR and SP likely caused the RR group to drive the classification results when discriminating HC or CIS against the whole MS cohort (RR and SP). Feature selection was performed implicitly by the random forest based on the relative contribution of each feature to the classification. No prior feature selection was performed as it would have reduced the exploratory power of this study. Each MRI modality came with its own limitations, which also must be taken into consideration, e.g., the multi-compartment diffusion model lacking a myelin volume fraction, or the use of surrogate quantitative PD, T2, and T1 extracted from routine scans not optimized to the scope. Particularly, the *MyRelax* algorithm for the calculation of quantitative T1 diverges in regions of cerebrospinal fluid, which were not of interest to this work; however, lesions

also exhibit a similar behavior at their core, therefore the summary statistics for quantitative T1 in lesions is to be intended as representative more of the peripheral part of lesions, where the partial volume effect with cerebrospinal fluid is less pronounced, than the central part. Additional studies with larger sample size and histological evidence are required to substantiate these findings.

We showed that different MRI features appear to be biophysically meaningful when discriminating CIS and clinically defined MS phenotypes, with qualitative and quantitative MRI modalities offering specific insights for different classification tasks. Key to our results is highlighting the need for further studies focused on the role of quantitative MRI in the lesions of early CIS and MS subjects to score risks of progression. These findings can help in further understanding MS pathophysiology, as well as inform future studies toward more efficient acquisition protocols, better tailored to the scope.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## Ethics statement

The studies involving human participants were reviewed and approved by Research Ethics Committee 13/LO/1413. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

AR, OC, DA, and CGWK designed the study. WB enrolled the subjects and collected all the clinical data. MY acquired the MRI data, with support from BS, FR, XG, and CGWK. AR designed and performed the analyses with support from FG, BK, FP, DA, and CGWK. OC and CGWK provided support and guidance with the interpretation of the results. AR wrote the manuscript, with support from OC and CGWK, and comments from all authors. All authors contributed to the article and approved the submitted version.

## Funding

This project has received funding under the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 634541. FG received the support of a fellowship from "la Caixa" Foundation (ID 100010434). The fellowship code is "LCF/BQ/PR22/11920010". FG has also received support from the Beatriu de Pinós (2020 BP 00117) programme, funded by the Secretary of Universities and Research (Government of Catalonia). BK, FP, and OC are supported by the National Institute of Health Research Biomedical Research Centre at UCL and UCLH. EPSRC grants EP/M020533/1 and EP/J020990/01, MRC MR/T046422/1 and MR/T046473/1, Wellcome Trust award 221915/Z/20/Z, and the NIHR UCLH BRC support DCA's work in this area. CGWK

also receives funding from Horizon 2020 [Research and Innovation Action Grants Human Brain Project 945539 (SGA3)], BRC (#BRC704/CAP/CGW), MRC (#MR/S026088/1), and Ataxia UK.

## Acknowledgments

We thank Patricia Mota for contributing in the development of the sodium imaging protocol, and Enrico Kaden for supporting the SMT image analysis.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

- Andersson, J. L., and Sotiropoulos, S. N. (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage* 125, 1063–1078. doi: 10.1016/j.neuroimage.2015.10.019
- Aquino, D., Bizzi, A., Grisoli, M., Garavaglia, B., Bruzzone, M. G., Nardocci, N., et al. (2009). Age-related iron deposition in the basal ganglia: quantitative analysis in healthy subjects. *Radiology* 252, 165–172. doi: 10.1148/radiol.2522081399
- Armstrong, C. M. (2003). The Na/K pump, Cl ion, and osmotic stabilization of cells. *Proc. Natl. Acad. Sci. U.S.A.* 100, 6257–6262. doi: 10.1073/pnas.0931278100
- Barkhof, F. (1999). MRI in multiple sclerosis: correlation with expanded disability status scale (EDSS). *Multiple Scler. J.* 5, 283–286. doi: 10.1177/135245859900500415
- Barkhof, F. (2002). The clinico-radiological paradox in multiple sclerosis revisited. *Curr. Opin. Neurol.* 15, 239–245. doi: 10.1097/00019052-200206000-00003
- Brownlee, W. J., Solanky, B., Prados, F., Yiannakas, M., Da Mota, P., Riemer, F., et al. (2019). Cortical grey matter sodium accumulation is associated with disability and secondary progressive disease course in relapse-onset multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry* 90, 755–760. doi: 10.1136/jnnp-2018-319634
- Cardoso, M. J., Modat, M., Wolz, R., Melbourne, A., Cash, D., Rueckert, D., et al. (2015). Geodesic information flows: spatially-variant graphs and their application to segmentation and fusion. *IEEE Trans. Med. Imaging* 34, 1976–1988. doi: 10.1109/TMI.2015.2418298
- Chard, D., and Trip, S. A. (2017). Resolving the clinico-radiological paradox in multiple sclerosis. *F1000Res.* 6:1828. doi: 10.12688/f1000research.11932.1
- Collorone, S., Cawley, N., Grussu, F., Prados, F., Tona, F., Calvi, A., et al. (2020). Reduced neurite density in the brain and cervical spinal cord in relapsing-remitting multiple sclerosis: a Noddi study. *Multiple Scler. J.* 26, 1647–1657. doi: 10.1177/1352458519885107
- Eshaghi, A., Marinescu, R. V., Young, A. L., Firth, N. C., Prados, F., Jorge Cardoso, M., et al. (2018). Progression of regional grey matter atrophy in multiple sclerosis. *Brain* 141, 1665–1677. doi: 10.1093/brain/awy088
- Eshaghi, A., Wotschel, V., Cortese, R., Calabrese, M., Sahraian, M. A., Thompson, A. J., et al. (2016). Gray matter MRI differentiates neuromyelitis optica from multiple sclerosis using random forest. *Neurology* 87, 2463–2470. doi: 10.1212/WNL.0000000000003395
- Filippi, M., Brück, W., Chard, D., Fazekas, F., Geurts, J. J., Enzinger, C., et al. (2019). Association between pathological and MRI findings in multiple sclerosis. *Lancet Neurol.* 18, 198–210. doi: 10.1016/S1474-4422(18)30451-4
- Filippi, M., Cercignani, M., Inglesse, M., Horsfield, M., and Comi, G. (2001). Diffusion tensor magnetic resonance imaging in multiple sclerosis. *Neurology* 56, 304–311. doi: 10.1212/WNL.56.3.304
- Gandini Wheeler-Kingshott, C. A., Riemer, F., Palesi, F., Ricciardi, A., Castellazzi, G., Golay, X., et al. (2018). Challenges and perspectives of quantitative functional sodium imaging (fNaI). *Front. Neurosci.* 12, 810. doi: 10.3389/fnins.2018.00810
- Grussu, F., Battiston, M., Veraart, J., Schneider, T., Cohen-Adad, J., Shepherd, T. M., et al. (2020). Multi-parametric quantitative *in vivo* spinal cord MRI with unified signal readout and image denoising. *Neuroimage* 217, 116884. doi: 10.1016/j.neuroimage.2020.116884
- Grussu, F., Schneider, T., Yates, R. L., Zhang, H., Wheeler-Kingshott, C. A. G., DeLuca, G. C., et al. (2016). A framework for optimal whole-sample histological quantification of neurite orientation dispersion in the human spinal cord. *J. Neurosci. Methods* 273, 20–32. doi: 10.1016/j.jneumeth.2016.08.002
- Inglesse, M., Madelin, G., Oesingmann, N., Babb, J., Wu, W., Stoeckel, B., et al. (2010). Brain tissue sodium concentration in multiple sclerosis: a sodium imaging study at 3 Tesla. *Brain* 133, 847–857. doi: 10.1093/brain/awp334
- Kaden, E., Kelm, N. D., Carson, R. P., Does, M. D., and Alexander, D. C. (2016). Multi-compartment microscopic diffusion imaging. *Neuroimage* 139, 346–359. doi: 10.1016/j.neuroimage.2016.06.002
- Lublin, F. D., Reingold, S. C., Cohen, J. A., Cutter, G. R., Sorensen, P. S., Thompson, A. J., et al. (2014). Defining the clinical course of multiple sclerosis: the 2013 revisions. *Neurology* 83, 278–286. doi: 10.1212/WNL.0000000000000560
- Lucchinetti, C., Brück, W., Parisi, J., Scheithauer, B., Rodriguez, M., and Lassmann, H. (2000). Heterogeneity of multiple sclerosis lesions: implications for the pathogenesis of demyelination. *Ann. Neurol.* 47, 707–717. doi: 10.1002/1531-8249(200006)47:6<707::AID-ANA3>3.0.CO;2-Q
- Maarouf, A., Audoin, B., Konstantin, S., Rico, A., Soulier, E., Reuter, F., et al. (2014). Topography of brain sodium accumulation in progressive multiple sclerosis. *Magn. Reson. Mater. Phys. Biol. Med.* 27, 53–62. doi: 10.1007/s10334-013-0396-1
- Maarouf, A., Audoin, B., Pariollaud, F., Gherib, S., Rico, A., Soulier, E., et al. (2017). Increased total sodium concentration in gray matter better explains cognition than atrophy in MS. *Neurology* 88, 289–295. doi: 10.1212/WNL.0000000000003511
- Magliozzi, R., Reynolds, R., and Calabrese, M. (2018). MRI of cortical lesions and its use in studying their role in MS pathogenesis and disease course. *Brain Pathol.* 28, 735–742. doi: 10.1111/bpa.12642
- Mezer, A., Yeatman, J. D., Stikov, N., Kay, K. N., Cho, N.-J., Dougherty, R. F., et al. (2013). Quantifying the local tissue volume and composition in individual brains with magnetic resonance imaging. *Nat. Med.* 19, 1667–1672. doi: 10.1038/nm.3390
- Neema, M., Stankiewicz, J., Arora, A., Dandamudi, V. S., Batt, C. E., Guss, Z. D., et al. (2007). T1- and T2-based MRI measures of diffuse gray matter and white matter damage in patients with multiple sclerosis. *J. Neuroimaging* 17, 16S–21S. doi: 10.1111/j.1552-6569.2007.00131.x
- Paling, D., Solanky, B. S., Riemer, F., Tozer, D. J., Wheeler-Kingshott, C. A., Kapoor, R., et al. (2013). Sodium accumulation is associated with disability and a progressive course in multiple sclerosis. *Brain* 136, 2305–2317. doi: 10.1093/brain/awt149
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830.
- Petracca, M., Vancea, R. O., Fleysher, L., Jonkman, L. E., Oesingmann, N., and Inglesse, M. (2016). Brain intra- and extracellular sodium concentration in multiple sclerosis: a 7 t MRI study. *Brain* 139, 795–806. doi: 10.1093/brain/awv386
- Prados, F., Cardoso, M. J., Kanber, B., Ciccarelli, O., Kapoor, R., Wheeler-Kingshott, C. A. G., et al. (2016a). A multi-time-point modality-agnostic patch-based method for lesion filling in multiple sclerosis. *Neuroimage* 139, 376–384. doi: 10.1016/j.neuroimage.2016.06.053
- Prados, F., Solanky, B. S., Alves Da Mota, P., Cardoso, M. J., Brownlee, W. J., Riemer, F., et al. (2016b). “Automatic sodium maps reconstruction using patchmatch algorithm for phantom detection,” in *ISMRM* (Singapore).
- Riemer, F., Solanky, B. S., Stehning, C., Clemence, M., Wheeler-Kingshott, C. A., and Golay, X. (2014). Sodium (<sup>23</sup>Na) ultra-short echo time imaging in the human

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fninf.2023.1060511/full#supplementary-material>

brain using a 3D-cones trajectory. *Magn. Reson. Mater. Phys. Biol. Med.* 27, 35–46. doi: 10.1007/s10334-013-0395-2

Roosendaal, S., Geurts, J. J., Vrenken, H., Hulst, H., Cover, K. S., Castelijns, J., et al. (2009). Regional DTI differences in multiple sclerosis patients. *Neuroimage* 44, 1397–1403. doi: 10.1016/j.neuroimage.2008.10.026

Sastre-Garriga, J., Pareto, D., Battaglini, M., Rocca, M. A., Ciccarelli, O., Enzinger, C., et al. (2020). Magnims consensus recommendations on the use of brain and spinal cord atrophy measures in clinical practice. *Nat. Rev. Neurol.* 16, 171–182. doi: 10.1038/s41582-020-0314-x

Savini, G., Pardini, M., Castellazzi, G., Lascialfari, A., Chard, D., D'Angelo, E., et al. (2019). Default mode network structural integrity and cerebellar connectivity predict information processing speed deficit in multiple sclerosis. *Front. Cell. Neurosci.* 13, 21. doi: 10.3389/fncel.2019.00021

Shin, H.-G., Lee, J., Yun, Y. H., Yoo, S. H., Jang, J., Oh, S.-H., et al. (2021).  $\chi$ -separation: magnetic susceptibility source separation toward iron and myelin mapping in the brain. *Neuroimage* 240, 118371. doi: 10.1016/j.neuroimage.2021.118371

Soares, J., Sousa, T., d'Almeida, O. C., Batista, S., Sousa, L., Castelo-Branco, M., et al. (2020). "Investigating whole-brain MRI markers in multiple sclerosis—emerging dimensions in morphometric space," in *XV Mediterranean Conference on Medical and Biological Engineering and Computing—MEDICON 2019: Proceedings of MEDICON 2019* (Coimbra), 1644–1652.

Thompson, A. J., Banwell, B. L., Barkhof, F., Carroll, W. M., Coetzee, T., Comi, G., et al. (2018). Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol.* 17, 162–173. doi: 10.1016/S1474-4422(17)30470-2

Tur, C., Grussu, F., Prados, F., Charalambous, T., Collorone, S., Kanber, B., et al. (2020). A multi-shell multi-tissue diffusion study of brain connectivity in early multiple sclerosis. *Multiple Scler. J.* 26, 774–785. doi: 10.1177/1352458519845105

VanRossum, G., and Drake, F. L. (2010). *The Python Language Reference*. Python Software Foundation, Amsterdam.

Wottschel, V., Alexander, D., Kwok, P., Chard, D., Stromillo, M., De Stefano, N., et al. (2015). Predicting outcome in clinically isolated syndrome using machine learning. *Neuroimage Clin.* 7, 281–287. doi: 10.1016/j.nicl.2014.11.021