OPEN ACCESS

Original research

Predicting disability progression and cognitive worsening in multiple sclerosis using patterns of grey matter volumes

Elisa Colato ,[1] Jonathan Stutters ,[1] Carmen Tur ,[1] Sridar Narayanan ,[2] Douglas L Arnold ,[2] Claudia A M Gandini Wheeler-Kingshott ,[1,3,4] Frederik Barkhof ,[1,5,6] Olga Ciccarelli ,[1,7] Declan T Chard ,[1,7] Arman Eshaghi [1,5]

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/jnnp-2020-325610).

For numbered affiliations see end of article.

Correspondence to
Elisa Colato, Russell Square House, 10-12 Russell Square, Queen Square MS Centre, Department of Neuroinflammation, UCL Queen Square Institute of Neurology, Faculty of Brain Sciences, University College London, London, UK WC1B 5EH; elisa.colato.18@ucl.ac.uk

Received 12 November 2020
Revised 18 March 2021
Accepted 20 March 2021

To cite: Colato E, Stutters J, Tur C, *et al*. *J Neurol Neurosurg Psychiatry* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/jnnp-2020-325610

ABSTRACT

Objective In multiple sclerosis (MS), MRI measures at the whole brain or regional level are only modestly associated with disability, while network-based measures are emerging as promising prognostic markers. We sought to demonstrate whether data-driven patterns of covarying regional grey matter (GM) volumes predict future disability in secondary progressive MS (SPMS).

Methods We used cross-sectional structural MRI, and baseline and longitudinal data of Expanded Disability Status Scale, Nine-Hole Peg Test (9HPT) and Symbol Digit Modalities Test (SDMT), from a clinical trial in 988 people with SPMS. We processed T1-weighted scans to obtain GM probability maps and applied spatial independent component analysis (ICA). We repeated ICA on 400 healthy controls. We used survival models to determine whether baseline patterns of covarying GM volume measures predict cognitive and motor worsening.

Results We identified 15 patterns of regionally covarying GM features. Compared with whole brain GM, deep GM and lesion volumes, some ICA components correlated more closely with clinical outcomes. A mainly basal ganglia component had the highest correlations at baseline with the SDMT and was associated with cognitive worsening (HR=1.29, 95% CI 1.09 to 1.52, p<0.005). Two ICA components were associated with 9HPT worsening (HR=1.30, 95% CI 1.06 to 1.60, p<0.01 and HR=1.21, 95% CI 1.01 to 1.45, p<0.05). ICA measures could better predict SDMT and 9HPT worsening (C-index=0.69–0.71) compared with models including only whole and regional MRI measures (C-index=0.65–0.69, p value for all comparison <0.05).

Conclusions The disability progression was better predicted by some of the covarying GM regions patterns, than by single regional or whole-brain measures. ICA, which may represent structural brain networks, can be applied to clinical trials and may play a role in stratifying participants who have the most potential to show a treatment effect.

INTRODUCTION

Multiple sclerosis (MS) is an inflammatory and neurodegenerative disease of the central nervous system. The most recognised pathological feature of MS is an inflammatory demyelinating white matter (WM) lesion, whose formation is associated with relapses.[1–3] However, the principal driver of irreversible disability, and progressive MS, is thought to be neurodegeneration.[4 5] We now have many treatments that reduce the risk of MS relapses, but only have two licensed treatments for progressive MS, and their efficacy appears to be mainly in people who still show evidence of ongoing inflammatory lesion activity.[6 7]

Neurodegeneration manifests as brain atrophy and this can be measured with MRI.[8] Grey matter (GM) volume loss contributes to brain atrophy.[5] In patients with secondary progressive MS (SPMS), GM volume loss is faster in the deep grey matter (DGM) than the cortex, and within the cortex, it preferentially affects temporal and parietal regions.[5 9 10] However, regional and global brain atrophy, and other conventional MRI measures, only partly correlate with and predict disability progression in people with progressive MS.[11] In part, this is explained by pathology being assessed at a whole or regional brain level, while the disability occurs as a result of impaired connections between clinically eloquent regions and brain networks.

Pathology in MS affects some parts of the brain more than others, and ideally, we should seek to measure pathology where it is most likely to affect clinical outcomes. Network-based measures have the potential to add value to conventional MRI measures, and have already proven promising in explaining motor disability.[12] Data-driven GM network measures are also a good candidate to be used as prognostic markers in clinical trials. Data-driven measures of brain networks have the potential to select those more likely to progress to enrich clinical trials and so demonstrate treatment effects.

Independent component analysis (ICA) is a robust data-driven technique that has been used to identify brain networks on structural MRI.[13 14] Spatial ICA can identify separate brain regions whose volume covaries, which can be linked by a common biological or pathological property.[15 16] In a mainly relapsing-remitting (RR) MS cohort, in a cross-sectional study Steenwijk *et al* identified covarying patterns in cortical thickness associated with clinical outcomes. A previous study in early RRMS showed that covarying patterns of GM intensities at baseline did not predict confirmed disability

progression (CDP) within 10 years, or at 10 years differentiate between patients with CDP and without CDP.[17] These studies were weighted towards RRMS, and while atrophy occurs early in MS, it is more prominent, and thought to be more clinically relevant,[18 19] during the progressive phase. No study so far has looked at the predictive value of baseline network-based measures of the cortex and DGM.

The overarching goal of our study was to apply network-based MRI measures of GM volume changes, seeking to better predict disability progression in SPMS when compared with conventional regional or whole brain volumes. We applied spatial ICA to identify covarying patterns of GM from structural MRI in 988 people with SPMS. Our specific aims were to (1) identify clinically relevant measures of covarying GM volume at study entry and (2) identify patterns of covarying GM volumes that predict future disability progression. We also aimed to assess stability and reliability of these patterns.

MATERIALS AND METHODS

Participants

We re-analysed data from the ASCEND trial, an international (163 sites across 17 countries), phase III, randomised, double-blind, placebo-controlled trial. Baseline and longitudinal clinical and baseline MRI data from 1003 subjects aged 18– 58 years, who had SPMS, and baseline Expanded Disability Status Scale (EDSS) between 3.0 and 6.5, were acquired.[20] Clinical data were acquired at baseline, every 12 weeks up to week 108 and at week 156. Because ASCEND was a negative trial, we included data from participants on natalizumab and from those in the placebo arm. We included visits that acquired the following MRI sequences: (1) T1-weighted without contrast administration MRI scans and (2) T2-fluid-attenuated inversion recovery (FLAIR) and/or (3) T2-weighted. We excluded data from (n=15) participants with artefacts on the available scans and re-baselined data from subjects (n=57) with artefacts just on their baseline scan (eg, ghosting, magnetic susceptibility and motion artefacts).

To determine whether similar patterns were present in healthy controls, we randomly selected 400 participants from the Human Connectome Project (males=171, females=229, mean age 28.89±3.74) and downloaded the available three-dimensional (3D) T1 and T2 scans.

MRI acquisition and processing

Image acquisition

Brain scans were acquired at either 1.5 or 3.0 T with two-dimensional (2D) T1-weighted sequences with voxel size=0.98×0.98×3 mm^3; (2) fast FLAIR with voxel size=0.98×0.98×3 mm^3 and (3) T2-weighted sequences with voxel size=0.98×0.98×3 mm^3. Details on MRI acquisition from a representative centre are provided in the online supplemental materials.

Image processing

The aim of image processing was to extract GM probability maps which are the input to ICA from T1-weighted MRI. We followed the steps as shown in figure 1.

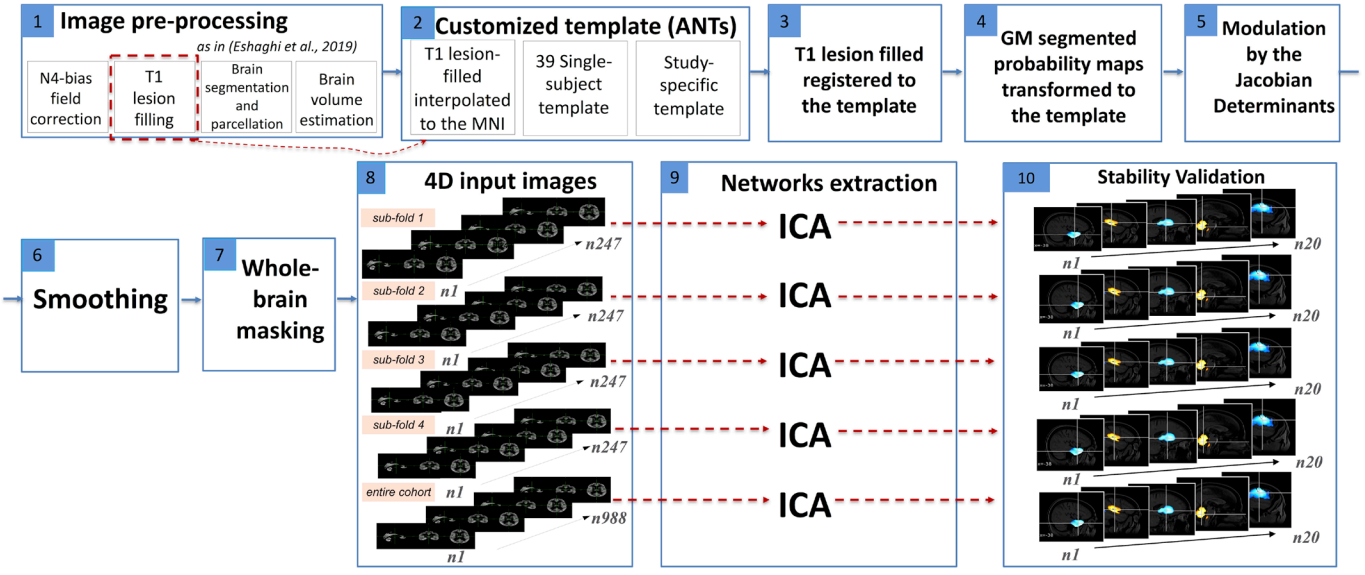


Figure 1 Visual representation of our image-analysis pipeline. Aiming to identify data-driven network-based measures of covarying GM volumes, we initially preprocessed our data as in *Eshaghi et al* (N4 bias field correction, lesion filling, brain segmentation and parcellation). We created a customised template from all the available scans from 39 randomly selected subjects. After having resampled those scans to an isotropic space, we created 39 single subject templates, and from those an average study-specific template. We registered the T1 lesion filled scans to the template and diffeomorphically transformed the GM segmentation maps to the template using the warping matrix generated from the previous step. We modulated the GM segmentation maps by the Jacobian determinants in order to account for possible deformations to the original volumes occurred after the non-linear transformation. We applied an 8 mm smoothing kernel to account for intersubject variability and applied a whole brain mask to constrain the following analysis at the level of the brain. Aiming to prove the stability of our results, we randomly divided our cohort into four folds. For each fold and for the entire cohort, we generated a 4D image by concatenating the available GM maps and ran fast ICA on each of those inputs allowing for 20 components to be identified. For each fold and for the entire cohort, we generated a 4D image by concatenating the 20 generated ICA components and ran cross-sectional correlations between those inputs to identify which components were stable and could be implemented for statistical analysis. 4D, four-dimensional; ANTs, advanced normalisation tools; GM, grey matter; ICA, independent component analysis.

We used an established pipeline as described elsewhere.[21] Briefly, this pipeline included N4 bias field correction,[22] lesion filling[23] (to reduce the effects of hypointense lesions in T1 scans during segmentation), and used the Geodesic Information Flows (GIF) V.3.0[24] to segment the lesion filled T1-weighted images into GM, WM and cerebrospinal fluid (CSF) probability maps, as well as to divide the brain into 120 parcellations according to the Neuromorphometrics atlas.[25] We used GIF because it allows the inclusion of 2D-MRI data and does not require additional manual editing, which for a cohort of this size would have been unfeasible. EC visually inspected these outputs to check for erroneous segmentation or segmentation outputs.

Voxel-based analysis of GM probability maps

We randomly selected 39 participants to create a study-specific template (as described in online supplemental materials), to improve the accuracy of the registration and minimise bias.[26]

For each participant, we registered the baseline T1 lesion filled scans to the study-specific template using rigid, affine and diffeomorphic non-linear registrations.[27] We calculated the cortical and deep GM volumes from GM probability maps in the native space. We transformed GM probability maps to the template by applying the warping matrices obtained from the previous step. We modulated the GM probabilistic maps by the Jacobian determinants estimated in Advanced Normalisation tools V.2.3.1 to adjust for deformations that occurred to the original volumes after the non-linear registration.[17 26 28] We used an 8 mm full width at half maximum smoothing kernel to account for inter-subject variability. We created a whole brain parcellation mask (as described in online supplemental materials) to constrain the ICA to the brain, and to identify and label brain regions in each ICA component.

Analysis of covarying GM volumes with ICA

We used the FastICA algorithm[29] implemented in scikit-learn 0.23.1 to identify the independent components representing spatial maps of GM covariation (GM patterns). We concatenated the GM probability maps from the ASCEND cohort into a 4D volume and fitted the ICA model allowing for 20 components to be identified.[14 30] To assess the stability and repeatability of the identified components, we randomly divided our cohort into four folds (247 subjects each) and repeated the analysis for each fold. We generated a 4D image by concatenating the 20 identified components and assessed pairwise spatial cross-correlations with 'fslcc' in FSL[31] to select components that were spatially stable for each fold (see figure 1). We defined components with statistically significant correlations ($p<0.05$) across folds and entire cohort as stable.

To account for heterogeneity of MRI protocols, we split the cohort based on the manufacturer used to acquire the MRI images, and repeated ICA for each subsample. We assessed pairwise spatial cross-correlations between the entire cohort and each cross-validation fold using '*fslcc*' to determine whether components were stable across manufacturers. To account for between-site variability, we split the cohort based on their geographical region, and repeated ICA for each cross-validation fold. We applied spatial cross-correlations to determine whether the ICA components identified from the entire cohort were stable regardless of the geographical region in which they were acquired and investigate the reliability of data-driven GM networks.

We overlaid the stable components obtained from the entire cohort with our whole-brain mask (obtained as described in online supplemental materials) to label brain areas involved in each component. To infer potential functional relevance, we visually compared the identified GM patterns with functional networks previously reported in the literature.[32 33]

We repeated ICA for healthy controls as described in online supplemental materials.

We used the loading factors of the stable components for further statistical analysis. Loading factors quantify the contribution of a given subject to a particular component.

Statistical analysis

We computed z-scores from the loading factors for each ICA component, whole brain GM, DGM and other brain regions volumes with R V.3.6.1. To identify components that represent overall brain preservation and brain volume loss, we correlated the loading factors of the ICA components with baseline whole brain GM volume. We correlated ICA factors with whole brain GM volumes, rather than with the volumes of brain regions involved in each component, because we aimed to determine the direction of ICA-brain volume associations, not their true magnitude (correlations are likely to be smaller than they would otherwise have been considering just brain volumes comprised in each network). To further identify which brain areas in each ICA component presented volume loss and which were preserved, we obtained parcellation maps for each participant from the preprocessing stage, estimated the baseline volume of each region and by correlating the baseline volume of each region involved in each ICA component, we determined whether the corresponding regions in the ICA patterns were preserved (positive correlation) or represented volume loss (negative correlation) (see online supplemental materials for additional details). We calculated Pearson's correlation coefficients across z-scores of ICA components and baseline average (dominant and non-dominant hands) inverse 9HPT (1/9HPT) and SDMT, and Spearman's correlations between these z-scores and the baseline EDSS, to determine the association between the GM patterns measures and the current clinical status. We calculated correlation coefficients for EDSS, inverse 9HPT and SDMT with the z-scores of lesion load, whole brain GM volume, DGM and thalamus volume.[34 35] To account for the number of comparisons performed, correlations were corrected for multiple comparisons using Bonferroni correction ($a=0.05$).

To calculate time to worsening of physical and cognitive disability, we estimated the EDSS progression as an increase of 1 point from a baseline EDSS score of 5.5 or below, or of 0.5 points from baseline EDSS score >5.5, and these scores were confirmed at least at 3 months.[36] We excluded from this estimation all the clinical visits within 30 days of an MS relapse. We also estimated the 9HPT and SDMT worsening as respectively a 20% increase[36 37] and 10% decrease[38 39] with respect to the baseline score.

We performed Cox regression analysis to determine whether the standardised loading of GM patterns at baseline could predict the clinical disability. We built one model for each independent variable (ie, ICA components, whole brain atrophy, DGM atrophy, lesion load and atrophy in smaller regions), adjusting for age, gender, trial arm and centre, and having the event and the time-to-event as dependent variables. We used the date of the baseline clinical visit, and the date of the clinical visit at which participants had an event or the last available clinical visit (in case no event was detected), to estimate the time-to-event.

To determine whether data-driven patterns provide added value above to regional MRI volumes and lesion loads, we performed post hoc analysis with multivariate Cox proportional

Table 1 Characteristics of participants

	N=988
Gender (M/F)	366/622
Age (mean±SD)	46.71±7.70
Trial arm (DMT/placebo)	420/422
EDSS (median, range)	6 (3–7.5)
SDMT (mean±SD)	39.86±14.20
9HPT (mean±SD)	35.81±19.62
EDSS progression confirmed at 3 months (no. progressed/not progressed)	197/643
9HPT worsening (no. worsened/not worsened)	177/244
SDMT worsening (no. worsened/not worsened)	173/187

EDSS progression was defined as 1 point increase from a baseline EDSS score ≤5.5, or as 0.5 points from a baseline EDSS score >5.5, excluding all clinical visits within 30 days from an attack, and these scores were confirmed at 3 months.[36] We estimated the 9HPT worsening as a 20% increase with respect to the baseline score (Lublin *et al*[37]; Tur *et al*[36]). We calculated the SDMT worsening as a 10% decrease with respect to the baseline score.[38 39]
DMT, disease-modifying treatment; EDSS, Expanded Disability Status Scale; F, females; 9HPT, Nine-Hole Peg Test; M, males; SDMT, Symbol Digit Modalities Test.

regression analysis. To identify the best predictive model for 9HPT and SDMT progression, we defined the following three models and obtained their concordance indexes (C-index):

1. 15 stable ICA components;
2. 15 ICA components together with conventionally used MRI measures (whole brain GM, DGM and lesion load);
3. conventionally used MRI measures

We estimated the C-index, which is a measure of the discrimination power of survival models, and represents the proportion of subjects with a progression on the clinical test and a worse outcome predicted by the model (concordant pairs) divided by the total number of possible evaluation pairs.[40] A C-index of 1 represents a perfect model prediction, while a value of 0.5 denotes random prediction. Age, gender, trial arm and centre were used as covariates for each model.

Data availability

Processed data and codes used in this study are available on request from qualified investigators.

RESULTS

Participants

For 15 subjects, their scans did not meet our inclusion criteria. Therefore, our final cohort comprised 988 patients with SPMS (366 men and 622 women with mean age of 46.71±7.70). Table 1 reports the demographic characteristics of these patients.

We performed a two-sample test for equality of proportion to determine whether the randomly selected subcohort was representative of the entire cohort for the strength of the acquisition field and gender. The proportion of participants scanned at 1.5 and 3 T did not differ in the two samples (respectively, χ^2=0.86, df=1, p=0.35, 95% CI −0.19 to 0.77 and χ^2=0.30, df=1, p=0.59, 95% CI −0.10 to 0.17). The proportion of females and males in the two samples did not differ (respectively, χ^2=2.10, df=1, p=0.15, 95% CI −0.03 to 0.25 and χ^2=2.10, df=1, p=0.15, 95% CI −0.25 to 0.03]).

Spatial maps of ICA components overlap with previously known networks

While allowing for up to 20 components, spatial cross-correlation showed that 15 (figure 2) were stable (online supplemental table s1). As reported in online supplemental table S2 and S3, components were stable regardless of the geographical region or the scanner vendor.

Nine of the identified structural GM patterns partially overlapped well-known functional systems, however ICA components spanned more areas compared with the functional networks. For example, component 5 is a sensorimotor-like network, encompassing the precentral gyrus, postcentral gyrus and supramarginal gyrus (action-execution network). Component 8 is a cortico-basal ganglia-like network, spanning the brain stem, pons, thalamus, nucleus accumbens, insula, putamen, caudate, pallidum, frontal and temporal lobe. Component 20 resamples a default mode-like network (DMN-like), spanning mainly the precuneus, posterior cingulate and middle frontal gyrus. For a detailed description of the remaining networks and of regions associated with each component, see figure 2, online supplemental table S4 and table 2.

Patterns represent brain volume loss or preservation

We identified ICA components representing a mixture of relative brain preservation and brain volume loss. Two representative examples are:

- Component 20 was positively correlated with whole brain GM volumes (r=0.28, 95% CI 0.22 to 0.33, p<0.001). Higher component loading was associated with higher GM volume, therefore this component represents a pattern of relatively greater regional volume at baseline.
- Component 13, instead, was inversely correlated with whole brain GM (r=−0.38, 95% CI −0.43 to −0.33, p<0.001). Higher loading on component 13 was associated with lower whole brain GM volumes, thus this pattern represents brain volume loss. Online supplemental table S5 shows correlations between the loading of each ICA component and whole brain GM volumes.

We identified which brain region in each GM pattern had brain volume loss and which was preserved (see table 2 and online supplemental table S4).

Baseline GM patterns correlate with clinical measures

Among all ICA components, component 8 (in which higher values corresponded to lower basal ganglia volumes) was significantly correlated with the SDMT and inverse 9HPT (respectively, r=−0.44, 95% CI −0.52 to −0.36], p<0.001 and r=−0.32, 95% CI −0.38 to −0.25, p<0.001). Component 6 (in which higher values corresponded to higher cerebellar volumes) was correlated with EDSS (rho=−0.11, p<0.05) (figure 3). Overall, SDMT and 9HPT showed higher correlation coefficients with some ICA component (especially component 8) than with conventional MRI measures (online supplemental table S5 and S6).

Predicting disability progression with survival modelling

Predicting the risk of 12-week confirmed EDSS progression

Data were available for 840 participants (317 males, 523 females, 419 patients under disease-modifying treatment (DMT), 421 patients in the placebo group, mean time-to-event of 1.98 years). A total of 28.5% of subjects had 12-week confirmed EDSS progression (figure 1). None of the ICA patterns predicted EDSS progression.

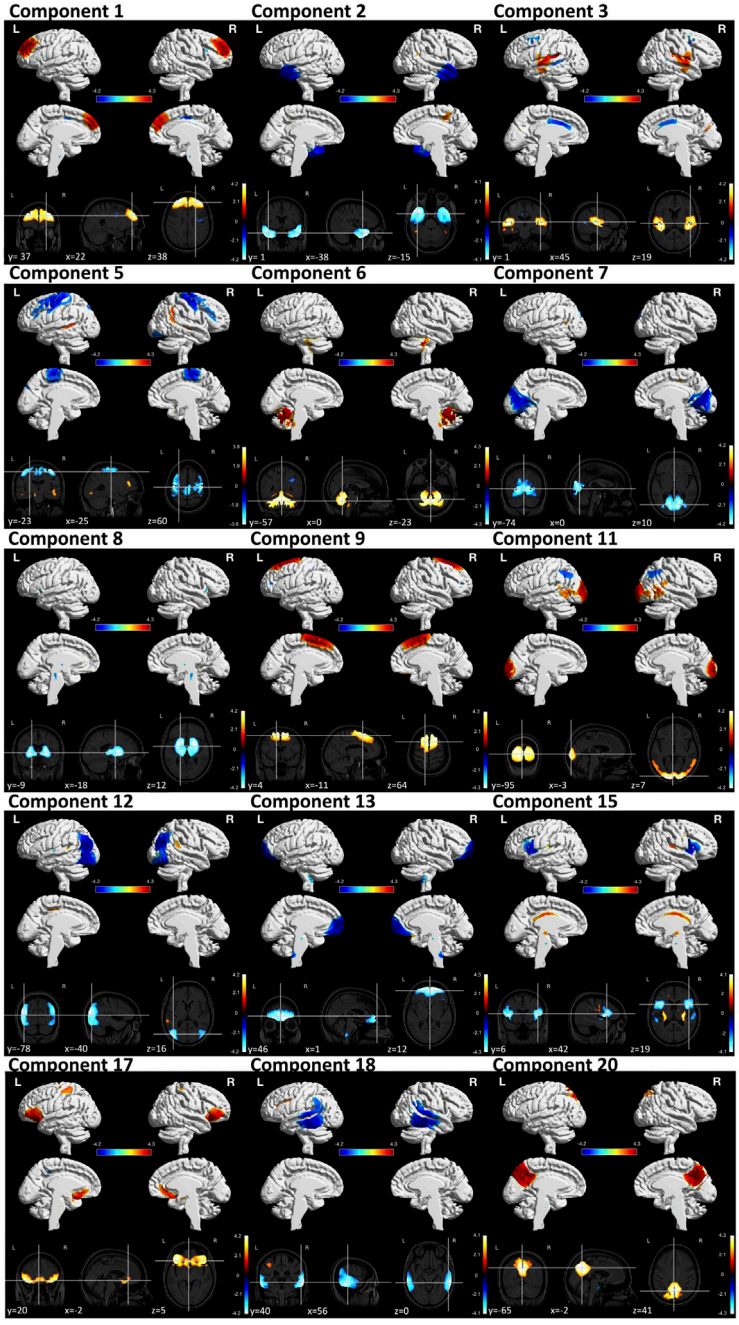


Figure 2 Stable independent component analysis (ICA) components. To determine the stability of the ICA components, we randomly split the sample into four folds and ran the ICA on each of them, as well as on the entire sample. While allowing for 20 components to be identified, cross-sectional correlations proved that only 15 out of the 20 ICA components were stable (emerged in all of the four folds and from the entire sample). The colour bar represents the loading of each component. Most of the identified networks resampled well-known functional systems. Component 3 represents an auditory-like network, spanning mainly the superior temporal gyrus, posterior insular and Heschl's gyrus (cognition-language-speech network). Component 5 is a sensorimotor-like network, encompassing the precentral gyrus, postcentral gyrus and supramarginal gyrus (action-execution network). Component 6 resamples a cerebellum-like network, involving mainly the cerebellum and fusiform gyrus, temporal and parietal lobe. Component 8 is a cortico-basal ganglia-like network, spanning the brain stem, pons, thalamus, nucleus accumbens, insula, putamen, caudate, pallidum, frontal and temporal lobe. Component 9 represents an executive control-like network, involving mainly medial frontal areas (action planning and inhibition). Component 11 is a visuo-like network, encompassing mainly several regions of the occipital pole and supramarginal, temporal and parietal areas. Component 15 resamples a salience-like network, involving the insula, thalamus and striatus (autonomic reaction to salient stimuli; goal-directed behaviour). Component 17 represents an affective and reward network, encompassing mainly the anterior cingulate, medial orbitofrontal cortex and prefrontal cortex. Component 20 resamples a default mode-like network (DMN-like), spanning mainly the precuneus, posterior cingulate and middle frontal gyrus. The remaining identified networks did not correspond to any major brain functional network, but can be labelled by their predominantly involved brain areas. Component 1 is a superior frontal network, encompassing mainly superior and medial frontal brain areas. Component 2 is a temporal-like network, involving mainly temporal brain regions. Component 7 is a precuneus-like network. Component 12 is an occipito-temporal-like network, spanning mainly the temporal and occipital pole. Component 13 represented a prefrontal cortex-like network, involving mainly frontal and orbitofrontal brain areas. Component 18 is a parieto-temporal-like network, involving mainly temporal and parietal brain areas.

Table 2 List of the clinically significant components with their corresponding involved brain regions

Components	Regions
1	▲ Superior frontal gyrus, ▲ Middle frontal gyrus, ▲ Superior frontal gyrus medial segment, ▲ Anterior cingulate gyrus, ▲ Opercular part of the inferior frontal gyrus
2	▼ Temporal pole, ▼ Inferior temporal gyrus, ▼ Middle temporal gyrus, ▼ Middle cingulate gyrus, ▼ Parahippocampal gyrus, ▼ Precentral gyrus medial segment, ▼ Posterior cingulate gyrus, ▼ Entorhinal area, ▼ Parietal lobule, ▼ Fusiform gyrus
6	▲ Cerebellum, ▲ Brain stem, ▲ Pons, ▼ Lingual gyrus, ▼ Fusiform gyrus, ▼ Temporal lobe ▼ Parietal lobe
7	▼ Superior occipital gyrus, ▼ Occipital lobe, ▼ Lingual gyrus, ▼ Calcalcarine cortex, ▼ Precuneus, ▼ Parietal lobe, ▼ Temporal lobe, ▲ Middle temporal gyrus, ▼ Frontal lobe, ▲ Precentral gyrus, ▼ Supramarginal gyrus
8	▼ Brain stem, ▼ Pons, ▼ Ventral DC, ▼ Thalamus, ▼ Insula, ▼ Accumbens, ▼ Caudate, ▼ Putamen, ▼ Pallidum, ▼ Frontal lobe ▲ Temporal lobe
11	▲ Occipital pole, ▲ Calcarine cortex, ▲ Cuneus, ▼ Middle temporal gyrus, ▼ Inferior temporal gyrus, ▲ Inferior occipital gyrus, ▼ Angular gyrus, ▼ Superior parietal lobule, ▼ Supramarginal gyrus
13	▼ Lateral orbital gyrus, ▼ Middle frontal gyrus, ▼ Superior frontal gyrus, ▼ Superior frontal gyrus medial segment, ▼ Anterior orbital gyrus, ▼ Medial frontal cortex, ▼ Gyrus rectus, ▼ Frontal pole, ▼ Medial orbital gyrus, ▼ Anterior cingulate gyrus, ▼ Brain stem, ▼ Lingual gyrus, ▼ Temporal pole
15	▲ Thalamus, ▲ Caudate, ▼ Anterior insula, ▼ Posterior insula, ▼ Planum polare, ▲ Putamen, ▼ Frontal operculum, ▼ Planum temporale, ▼ Claustrum, ▼ Triangular part of the inferior frontal gyrus, ▼ Opercular part of the inferior frontal gyrus, ▲ Precentral gyrus, ▼ Central operculum, ▼ Parietal operculum, ▼ Frontal lobe ▼ Temporal pole
17	▲ Hippocampus, ▲ Pons, ▲ Middle temporal gyrus, ▲ Superior temporal gyrus, ▲ Postcentral gyrus, ▲ Triangular part of the inferior frontal gyrus, ▲ Temporal pole, ▲ Posterior orbital gyrus, ▲ Medial orbital gyrus, ▲ Anterior insula, Claustrum, ▲ Basal forebrain, ▲ Putamen, ▲ Subcallosal area, ▲ Medial orbital gyrus, ▲ Gyrus rectus, ▲ Medial frontal cortex, ▲ Lateral orbital gyrus, ▲ Orbital part of the inferior frontal gyrus, ▲ Medial frontal cortex, ▲ Anterior cingulate gyrus, ▲ Anterior orbital gyrus, ▲ Posterior cingulate gyrus, ▲ Postcentral gyrus, ▲ Frontal operculum, ▲ Inferior temporal gyrus
18	▼ Middle occipital gyrus, ▼ Postcentral gyrus, ▼ Precentral gyrus, ▼ Opercular part of the inferior frontal gyrus, ▼ Fusiform gyrus, ▼ Parahippocampal gyrus, ▼ Frontal, ▼ Occipital lobe ▼ Parietal lobe, ▼ Inferior temporal gyrus, ▼ Middle temporal gyrus, ▼ Superior temporal gyrus, ▼ Supramarginal gyrus, ▼ Middle temporal gyrus
20	▲ Superior occipital gyrus, ▲ Superior parietal lobule, ▲ Precuneus, ▲ Posterior cingulate gyrus, ▲ Superior frontal gyrus, ▲ Middle frontal gyrus, ▲ Angular gyrus, ▲ Occipital lobule

We overlaid a whole brain parcellation mask with the identified ICA components in order to retrieve and label brain regions involved in each network. We correlated the loading of ICA components with the baseline volume of the areas involved in each network to identify which brain area in each network was atrophic (negative correlation between network loading and baseline volume) and which represented relative brain preservation (negative correlation between those volumes and ICA loadings).
▲ Relative preserved brain region.
▼ Atrophic brain region.
ICA, independent component analysis.

GM patterns predicted 9HPT worsening

Data for 361 subjects were available (134 males, 227 females, 191 patients under DMT, 170 patients in the placebo group). By the last available visit, 42% of participants experienced a worsening in the 9HPT after a mean time-to-conversion of 1.76 years (online supplemental figure S2). Component 2 (HR=1.30, 95%CI 1.06 to 1.60, p<0.01), component 20 (HR=1.21, 95%CI 1.01 to 1.45, p<0.05) and DGM (HR=0.72, 95%CI 0.52 to 0.99, p=0.05) predicted the worsening of the 9HPT (figure 4 and online supplemental table S7).

GM patterns predicted SDMT worsening

SDMT was available for 360 (140 males, 220 females; 185 under DMT, 175 in the placebo group) subjects. By the last available visit, 51% of participants had a 10% worsening[39] in SDMT score after a mean time-to-conversion of 1.36 years (online supplemental figure S3). SDMT worsening could be predicted by six of ICA components (component 7, component 8, component 13, component 15, component 17, component 18), lesion load and thalamus (figure 5 and online supplemental table S7).

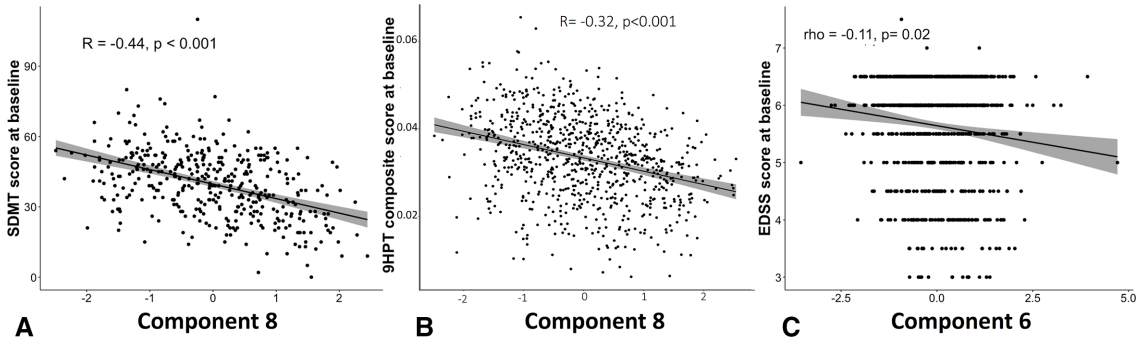


Figure 3 Correlations between baseline ICA components and baseline EDSS, 9HPT and SDM. Among the 15 stable ICA component, baseline SDMT score was more strongly associated with a mainly basal ganglia component (component 8). Among the three clinical tests, (A) SDMT had the highest correlations with ICA networks (mainly with component 8). (B) 9HPT was associated with the factor loading of component 8. 9HPT and SDMT correlated better with some ICA networks rather than with any other regional or whole brain MRI measure. (C) Among all the 15 networks, component 6 (ie, cerebellum, brain stem, pons) had the highest correlation with EDSS. We used the Bonferroni correction to correct for multiple comparisons. CI band is added to the figure. EDSS, Expanded Disability Status Scale; ICA. independent component analysis; SDMT, Symbol Digit Modalities Test; 9HPT, Nine-Hole Peg Test.

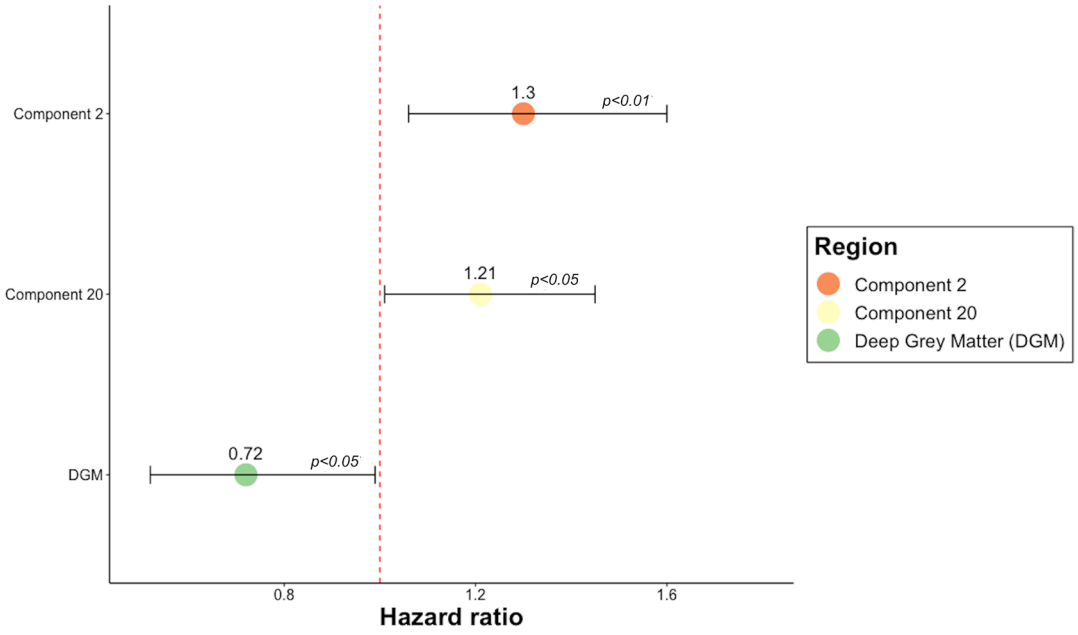


Figure 4 Cox regression models predictive of 9HPT worsening. HR of the statistically significant predictors of 9HPT worsening. The figure shows that two GM networks and the volume of the DGM can predict the 9HPT progression. HR >1 indicates that for each SD increase in the corresponding variable there is a higher risk of developing the event. HR <1 indicates that for each SD decrease in the corresponding variable, there is a higher risk of progressing on 9HPT. Error bars represent the CI. P values <0.05 represent a statistically significant relative risk of developing a 9HPT progression comparing subjects for each independent variable shown on the vertical axis. Component 2 encompasses the temporal lobe, middle cingulate gyrus, precentral gyrus medial segment, posterior cingulate gyrus, parietal lobule, inferior and middle temporal gyrus, parahippocampal gyrus, fusiform gyrus and entorhinal area. Component 20 consisted of precuneus, posterior cingulate gyrus, middle and superior frontal gyrus, angular gyrus, superior occipital and superior parietal lobule. 9HPT, Nine-Hole Peg Test; GM, grey matter.

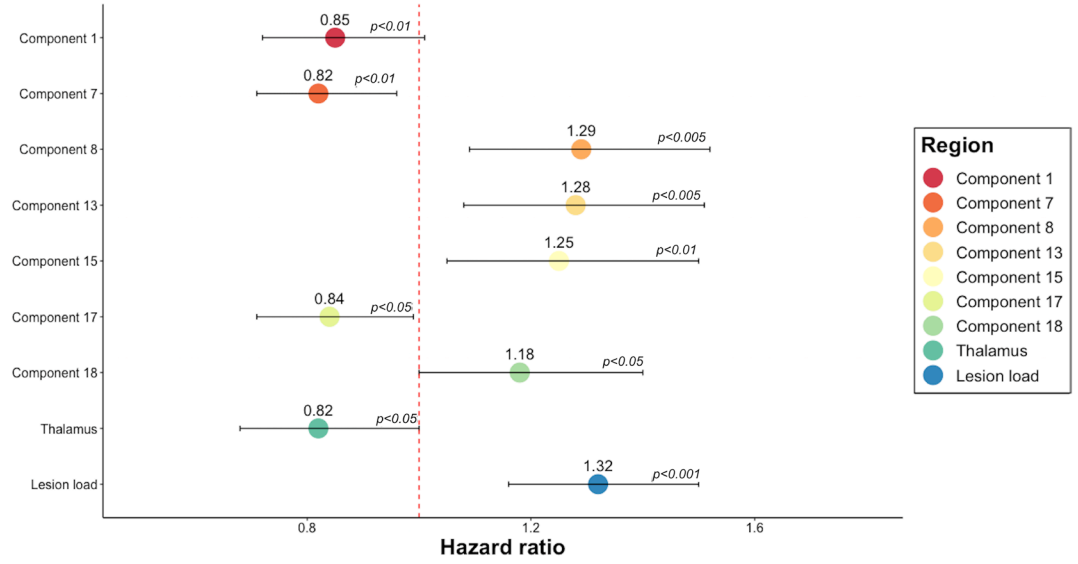


Figure 5 Cox regression models predictive of SDMT worsening. HR of the statistically significant predictors of SDMT worsening in separate Cox regression models. The figure shows that six ICA components, lesion load and the volumes of the thalamus could predict the SDMT progression. HR >1 indicates that for each SD increase in the corresponding variable, there is a higher risk of developing the event. HR >1 indicates that for each SD decrease in the corresponding variable, there is a higher risk of progressing on SDMT. For each SD increase in component 8 (encompassing mainly basal ganglia regions), which is inversely related to GM volumes, there was a 29% higher risk of developing SDMT progression. For each SD decrease in the volume of the thalamus, there is a 18% increased risk of worsening in SDMT. Error bars represent the CI of HR. P values <0.05 represent a statistically significant relative risk of developing an SDMT progression for each independent variable shown on the vertical axis. GM, grey matter; ICA. independent component analysis; SDMT, Symbol Digit Modalities Test.

Table 3 Comparison between different predictive models for 9HPT and SDMT progression

	20% 9HPT worsening			10% SDMT worsening		
Predictors	C-index	SE	Likelihood ratio test (p value)	C-index	SE	Likelihood ratio test (p value)
15 ICA components+DGM+whole brain DGM+lesion load	0.69	0.025	9e-05	0.71	0.021	4e-06
15 ICA components	0.68	0.025	2e-04	0.72	0.021	5e-06
Whole GM+DGM+lesion load	0.65	0.025	8e-05	0.69	0.022	3e-05

Models including ICA components and conventionally assessed MRI measures, or considering both DGM and the predictive ICA components identified by Cox models, have a stronger predictive value when compared with models including just measures of whole brain and DGM volumes and lesion load. Models including ICA components and conventionally assessed MRI measures have a stronger predictive value when compared with models including just measures of whole brain and DGM volumes and lesion load. C-index is generally used to validate the predictive ability of survival models. The likelihood ratio test represents the predictive statistically significance value of each model.
*Sex, age, trial arm and centre were used as covariates in each model.
C-index, concordance index; DGM, deep grey matter; GM, grey matter; 9HPT, Nine-Hole Peg Test; ICA, independent component analysis; SDMT, Symbol Digit Modalities Test.

GM patterns add value to conventional MRI measures: *Cox proportional regression analysis*

Models with ICA components had a higher C-index (C-index=0.69 (SE=0.025) compared with models including only conventional MRI measures (C-index=0.65, SE=0.025) (table 3).

The highest C-index belonged to a model with all ICA components (C-index=0.72, SE=0.021). When compared with models including just conventional MRI measures (C-index=0.69, SE=0.022), models that also include ICA components had higher C-index >0.71 (SE=0.021) (table 3).

DISCUSSION

Our main finding is that data-driven patterns of GM volume change predict physical and cognitive disability in a large cohort of people with SPMS. Furthermore, some of the GM ICA components correlated more closely with concurrent disability than regional or whole brain GM volumes, and some pattern-derived measures better predicted disability progression: two of these patterns predicted the 9HPT worsening and six components predicted cognitive disability better than any other assessed MRI measures. Post hoc analysis showed that pattern-based measures added value to conventional MRI measures. Interestingly, the components that correlated with concurrent disability were not necessarily the same as those associated with progression.

We found that some ICA components showed higher correlation coefficients with disability than regional and whole brain measures. For each disability measure, different ICA components dominated, encompassing both cortical and subcortical areas. For EDSS, it was component 6, which included regions in the cerebellum, brain stem, pons, lingual gyrus, fusiform gyrus, temporal and parietal lobe. For 9HPT, components 6 and 8 (thalamus, brain stem, pons, ventral diencephalon, insula, accumbens, caudate, putamen, pallidum, frontal and temporal lobe) were significant, and for the cognitive dysfunction (SDMT) five components spanning mainly the precuneus, posterior cingulate, temporal and frontal brain regions dominated. While the highest correlation coefficient for SDMT and 9HPT were reported with ICA component 8 (respectively, r=−0.44, 95% CI −0.52 to −0.36, p<0.001 and r=−0.32, 95% CI −0.38 to −0.25, p<0.001), lesion load and thalamus volume had higher correlation coefficients than other GM patterns. EDSS at baseline was more strongly associated with component 6 than DGM, whole GM atrophy and lesion load measures, but the volume of the thalamus in isolation had a higher correlation with EDSS. While whole brain GM and DGM measures span the whole brain, considering several regions not associated with the lower limb functions, component 6 comprised primarily (but not only) areas related to motor functions. Nonetheless, the involvement in this pattern of brain regions not related to motor functions might have decrease the strength of the correlation with EDSS score when compared with the volume of the thalamus taken in isolation. Thalamus is a neuralgic site for motor control, which has already been reported to be associated with EDSS.[41]

We found baseline ICA components correlated with baseline and longitudinal 9HPT and SDMT measures, and baseline EDSS, but did not predict EDSS progression. Correlation coefficients for EDSS and ICA components were mild to negligible and may not be clinically relevant. While correlations were statistically significant after applying Bonferroni correction and these values are in line with previous reports,[9 42–44] the statistical significance might have emerged because of the large sample size. We used correlations as a starting point to determine whether the identified components were associated with the current clinical status of patients with SPMS. We then applied survival models to determine the clinical relevance of GM patterns measures. Moreover, while ICA components may have greater relative clinical effects earlier and later in the course of MS, the limitations of disability measures, which are well-recognised for EDSS, might play a role. The EDSS was designed as a composite score, but is heavily weighted towards walking impairment, particularly affecting mid to higher score ranges (such as patients with progressive MS in our cohort, in which lower limb functions were already heavily impaired).[45] In contrast, the 9HPT and SDMT were designed as more specific measures and are more likely to reflect the effects of pathology. Finally, in patients with SPMS the preservation of upper limbs and cognitive functions—poorly represented by the EDSS score—are paramount. Thus, we can conclude that ICA components are marker of disease worsening in SPMS, but only as reflected by 9HPT and SDMT.

While the ICA components were identified without prior knowledge of functionally relevant brain regions, their correlations with disability reinforce their usage in predictive models. These components include regions linked with specific neurological and cognitive functions and also those that were both functionally and structurally related. Component 8, which was mainly a basal ganglia-fronto-temporal pattern, correlated with 9HPT and SDMT at baseline. It includes regions of DGM and cortical areas known to be involved in motor control, memory and learning.[46] These regions are also part of the cortico-thalamic, cortico-basal ganglia-thalamo-cortical and thalamo-cortical pathways that control both sensory and motor information coming from and going to the cortex.[47] Basal ganglia represent a series of interconnected subcortical nuclei which are known to be involved in selecting and implementing purposeful actions, facilitating voluntary movements and inhibiting the competing

or interfering ones, and controlling non-motor behaviours (eg, working memory, procedural learning, decision making, higher-order process of movement initiation).[47 48] Moreover, atrophy in the cortex, caudate, putamen, globus pallidus, thalamus and nucleus accumbens have been reported to be associated with lower performance in the SDMT.[49] Component six was mainly a cerebellar pattern. It encompassed brain regions in close proximity and functionally related (cerebellum, brain stem, pons, parietal lobe), already known to be involved in motor functions.

Consistent with previous work in predominantly RRMS populations, we found covarying and clinically relevant patterns of GM atrophy. Previous studies using ICA have identified 8[17] and 10[9] GM components. In the present study, we looked for 20 components, a practical maximum given available computational power, but found 15 could be consistently identified. Our ICA components only partially overlapped with previously reported GM patterns. For instance, component 5 resembled pattern 2 in the study by Steenwijk *et al* (they both include the middle temporal gyrus, superior temporal gyrus, supramarginal gyrus, postcentral gyrus and parietal lobule). However, in addition our ICA component includes other brain regions (cuneus and frontal gyrus) not reported by the previous study. Pattern 8 reported by Bergsland *et al* encompasses similar brain areas as in the ICA component 7 presented here (eg, calcarine cortex, precuneus, occipital and frontal lobe). However, a perfect match is never present and overall GM patterns presented in this study encompass a higher number of brain regions when compared with the above-mentioned studies. Most patterns identified in other studies were not replicated here, nor between studies. There are several potential reasons for this. First, cohort difference: when compared with the study by Bergsland *et al*, differences in our results may be related to more severe atrophy in SPMS compared with RRMS. Then there are methodological differences, for example, we used GM volumes as input to the ICA instead of cortical thickness, and we allowed for more components to be extracted. Further work is required to resolve these inconsistencies, but a clear overarching finding is that ICA-based analyses identify overlapping components which could otherwise be lost in whole brain and regional atrophy measures, and that these components are clinically relevant.

ICA components represent brain GM volumes that covary, not necessarily in the same direction. Six components showed both positive and negative loading values. This is consistent with previously reported studies on structural ICA where authors identified 5 components out of 10 that encompassed both relatively preserved and relatively atrophied brain regions.[9]

We can speculate that the interplay of several factors and mechanisms might explain the observed patterns.[15 50] For instance, DGM (in particular the putamen and caudate) is known to present several connections with motor and associative cortices, but appear to be susceptible sites for extensive demyelination and iron deposition.[51] Other brain regions involved in ICA components are known to be more susceptible to neurodegeneration due to CSF exposure (deep sulci in the temporal pole) and hypoxia (pallidum, precuneus and posterior cingulate). For example, the precuneus and posterior cingulate present extensive connections with several other brain regions and are part of the default brain functional system, known to present under normal condition the highest level of energy consumption.[52] Because neurons require a higher amount of energy to adapt to demyelination,[53] this could make highly connected brain regions more susceptible to neurodegeneration. Therefore, several mechanisms can cause the observed patterns of volume changes. Future work with longitudinal ICA studies will investigate this further.

Our study has some limitations. In this study, we re-analysed data from a large, negative, multicentre study, where MRI data were acquired with different scanners. External validation was not performed because data from another SPMS cohort were not available. However, we performed cross-validation for the different vendors and acquisition regions to determine whether they could affect the identified patterns. We found that components were stable among cross-validation folds. To take into account the effect of centre had on the association and predictive ability of MRI measures on clinical outputs, we used the centre as covariate in our regression models. Non-isotropic 2D T1-weighted scans were acquired, and while we were still able to measure cortical volumes, and identify multiple ICA components, isotropic 3D scans may enable future studies to identify additional components. Although high-resolution 3D T1-weighted scans would enable cortical thickness rather than volumes to be measured, and so may reveal different ICA patterns that we have been unable to detect and may improve on the predictive power of ICA measures, with baseline 2D T1-weighted scans we detected patterns of GM that predicted both motor and cognitive disability. Considering the broader availability of low resolution 2D scans in phase III clinical trials,[54] these findings reinforce the applicability of ICA patterns to both 2D and 3D images, and underline that 2D MRI can be analysed in retrospective imaging studies.

We preferred GM volumes over cortical thickness measures as input for ICA because we were also interested in changes of subcortical brain regions. Similar to previous works, we smoothed the probability maps to account for intersubject variability, but this will have reduced sensitivity to small regional effects, although offset by the large size of the cohort.

This study focused on GM patterns, however MS is a generalised disorder, and so while our ICA components often complemented whole or regional brain GM measures, future work will determine whether WM regions could increase the predictive accuracy of pattern-based measures. Because we used data from a phase III clinical trial, no data for healthy controls were available thus we cannot exclude whether the same patterns would be identified and how they would differ among healthy controls. Using MRI for healthy controls from the HCP, nine GM patterns identified from patients with SPMS were uniquely correlated and present also in healthy controls, suggesting that they are not disease specific. Two GM patterns identified from participants with SPMS (component 5 and 9) were spatially correlated with and represented respectively the sensorimotor element and the frontal element of a single GM pattern detected from HC. The same was true for components 7 and 11 identified from subjects with SPMS that represented different regions and correlated with a single HC GM pattern. A disconnection might explain the inability of ICA to detect component 5 and 9, and 7 and 11 as single GM patterns in participants with SPMS. However, data from HC and patients with MS acquired with the same protocol are needed to clarify this further. Although slightly overlapping with HC patterns, two SPMS ICA components (component 12 and 18) were detected in patients with MS but not from HCs, which might be disease specific. These GM patterns spanned mainly temporal regions that are known to be affected in patients with SPMS.[9] These regions are subserved by WM tracts impaired in SPMS. Particularly, a previous ICA study, based on DTI data from people with SPMS, has shown that among all WM tracts, the anterior commissure and corpus callosum show the most extensive WM abnormalities and the strongest correlations

with cognitive deficits.[55] While our analysis of HCs allows us to assess whether or not patterns are more or less MS-specific, we cannot directly compare ICA measures between the MS and HCs because the MRI scans were acquired with different protocols. While these components provide predictive value, further work is required to determine which components represent general structural variability that may be seen in general population, and those more specific to MS.

Because of the cross-sectional nature of this study, baseline MRI were used to predict longitudinal clinical disability. This represents a limitation of the study because it does not allow to determine causal relationships in brain network changes nor the dynamics of networks. Previous longitudinal studies have determined longitudinal evolution of structural GM networks in RRMS.[56] Similarly, a longitudinal study on structural brain networks obtained from cortical thickness in patients with primary progressive MS reported changes in network dynamics (increased connectivity and efficiency) from baseline to year 5 of follow-up mainly in patients with a fast disease progression.[57] Nonetheless, the cross-sectional nature of the study does not undermine the ability of baseline MRI data to predict cognitive and motor worsening and their relevance for clinical applicability. Recognising the importance and compelling utility of longitudinal studies, future longitudinal studies should investigate longitudinal changes and reorganisation of GM patterns.

C-indexes were close, but all statistical models were statistically significant and models that considered ICA components had higher accuracy and smaller SE. It therefore means that by adding ICA, we were able to predict outcomes with higher accuracy but the effect sizes were small. ICA can identify clinically relevant relationships in covarying brain regions and might provide new insights into the disease pathomechanisms. Therefore, although the accuracy of predictions is good but not excellent, the little improvement provided by ICA components is rewarding.

We used conventional norms and a raw score change of 4 point (10% change in magnitude) as a clinically meaningful measure[39] to define SDMT progression. However, future studies should investigate whether regression-based norms for SDMT might better detect abnormal performance in patients with MS[58–60] and therefore improve the performance of ICA components.

In conclusion, we have shown that ICA identifies regional patterns of GM volume loss, several of which are relevant to concurrent disability and some predict future progression. Several of the ICA-derived GM patterns were more closely linked with disability, and better able to predict disability progression, than simple MRI measures. Since the source data for this study was a phase III clinical trial, the ICA pipeline we have developed can readily be deployed in future clinical trials. Given the ability of some components to predict future progression, they could be used to stratify those who are more likely to progress.

Author affiliations
[1]NMR Research Unit, Queen Square Multiple Sclerosis Centre, Department of Neuroinflammation, UCL Queen Square Institute of Neurology, Faculty of Brain Sciences, University College London, London, UK
[2]McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada
[3]Department of Brain & Behavioural Sciences, University of Pavia, Pavia, Italy
[4]Brain Connectivity Centre, IRCCS Mondino Foundation, Pavia, Italy
[5]Centre for Medical Image Computing (CMIC), Department of Computer Science, University College London, London, UK
[6]Department of Radiology and Nuclear Medicine, Amsterdam University Medical Centers, Amsterdam, NL
[7]Institute for Health Research (NIHR), University College London Hospitals (UCLH) Biomedical Research Centre (BRC), London, UK

Twitter Elisa Colato @elisacolato and Arman Eshaghi @es_arman

Acknowledgements This study was supported by the International Progressive MSAlliance (IPMSA, award reference number PA-1603-08175). The authors are grateful to all the IPMSA investigators who have contributed trial data to this study as part of EPITOME: Enhancing Power of Intervention Trials Through Optimised MRI Endpoints network.DC,FB and OC are supported by the NIHR biomedical research centreat UCLH. Data collection and sharing for part of this project was provided by the Human Connectome Project (HCP; PrincipalInvestigators: Bruce Rosen, M.D., Ph.D., Arthur W. Toga, Ph.D.,Van J. Weeden, MD). HCP funding was provided by the NationalInstitute of Dental and Craniofacial Research (NIDCR), the National Institute of Mental Health (NIMH), and the NationalInstitute of Neurological Disorders and Stroke (NINDS). HCP data are disseminated by the Laboratory of Neuro Imaging at theUniversity of Southern California.

Contributors AE, DTC, OC, EC, SN, DLA and FB designed the study. EC, JS and AE carried out the image analysis. EC and CT performed the statistical analysis. EC, AE, DTC, OC, FB, CW-K, SN, DLA and JS interpreted the data and wrote the manuscript

Funding This study was supported by the International Progressive MS Alliance (IPMSA, award reference number PA-1603–08175). This study was supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre. DC, FB and OC are supported by the NIHR biomedical research centre at UCLH. OC is NIHR Research Professor (RP-2017-08-ST2-004). She also receives funding from MRC Research Grant (MR/S026088/1). Data collection and sharing for part of this project was provided by the Human Connectome Project (HCP; Principal Investigators: Bruce Rosen, MD, PhD, Arthur W. Toga, PhD, Van J. Weeden, MD). HCP funding was provided by the National Institute of Dental and Craniofacial Research (NIDCR), the National Institute of Mental Health (NIMH) and the National Institute of Neurological Disorders and Stroke (NINDS). HCP data are disseminated by the Laboratory of Neuro Imaging at the University of Southern California.

Disclaimer The authors have no competing interest with respect to this research. The full disclosure statement is as follows: CT has received an ECTRIMS Post-doctoral Research Fellowship in 2015. She has been a consultant for Roche in the last 12 months. DLA has received research grant funding and/or personal compensation for consulting from Acorda, Adelphi, Alkermes, Biogen, Celgene, Frequency Therapeutics, Genentech, Genzyme, Hoffman-La Roche, Immuene Tolerance Network, Immunotec, MedDay, EMD Serono, Novartis, Pfizer, Receptos, Roche, Sanofi-Aventis, Canadian Institutes of Health Research, MS Society of Canada, and International Progressive MS Alliance; and holds an equity interest in NeuroRx Research.S.N. has received research funding from the Canadian Institutes of Health Research, the International Progressive MS Alliance, the Myelin Repair Foundation and Immunotec. He has received honoraria/travel support from Genentech and MedDay, and personal compensation from NeuroRx Research. C.G.WK. receives funding from the MS Society (#77), Wings for Life (#169111), Horizon2020 (CDS-QUAMRI, #634541), BRC (#BRC704/CAP/CGW), UCL Global Challenges Research Fund (GCRF), MRC (#MR/S026088/1), Ataxia UK. C.G.WK. is a shareholder in Queen Square Analytics Ltd. F.B has received compensation for consulting services and/or speaking activities from Bayer Schering Pharma, Biogen Idec, Merck Serono, Novartis, Genzyme,Synthon BV, Roche, Teva, Jansen research and IXICO and is supported by the NIHR Biomedical Research Centre at UCLH. OC has received research grants from the MS Society of Great Britain & Northern Ireland, National Institute for Health Research (NIHR) University College London Hospitals Biomedical Research Centre, EUH2020, Spinal Cord Research Foundation, and Rosetrees Trust. She serves as a consultant for Novartis, Teva, and Roche and has received an honorarium from the American Academy of Neurology as Associate Editor of Neurology and serves on the Editorial Board of Multiple Sclerosis Journal. In the last 3 years DC is a consultant for Biogen and Hoffmann-La Roche. He has received research funding from the International Progressive MS Alliance, the MS Society, and the National Institute for Health Research (NIHR) University College London Hospitals (UCLH) Biomedical Research Centre. AE serves on the Editorial Board of Neurology. He has received speaker's honoraria from Biogen and At The Limits educational programme, travel support from the National Multiple Sclerosis Society and honorarium from the Journal of Neurology, Neurosurgery and Psychiatry for Editorial Commentaries. AE and FB have equity stake in Queen Square Analytics. EC and JS have nothing to disclose.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval The Institutional Review Board at the Montreal Neurological Institute (MNI), Quebec, Canada approved this study under the auspices of International Progressive MS Alliance (reference number: IRB00010120).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Processed data and codes used in this study are available on request from qualified investigators.

ORCID iDs
Elisa Colato http://orcid.org/0000-0002-4101-7058
Jonathan Stutters http://orcid.org/0000-0002-9151-0844
Carmen Tur http://orcid.org/0000-0003-1849-3184
Sridar Narayanan http://orcid.org/0000-0002-6283-150X
Douglas L Arnold http://orcid.org/0000-0003-4266-0106
Claudia A M Gandini Wheeler-Kingshott http://orcid.org/0000-0002-4832-1300
Frederik Barkhof http://orcid.org/0000-0003-3543-3706
Olga Ciccarelli http://orcid.org/0000-0001-7485-1367
Declan T Chard http://orcid.org/0000-0003-3076-2682
Arman Eshaghi http://orcid.org/0000-0002-6652-3512

REFERENCES

1 Filippi M, Bar-Or A, Piehl F, *et al*. Multiple sclerosis. *Nat Rev Dis Primers* 2018;4:1–27.
2 Filippi M, Preziosa P, Banwell BL, *et al*. Assessment of lesions on magnetic resonance imaging in multiple sclerosis: practical guidelines. *Brain* 2019;142:1858–75.
3 Lassmann H. Multiple sclerosis pathology. *Cold Spring Harb Perspect Med* 2018;8:a028936.
4 Roosendaal SD, Bendfeldt K, Vrenken H, *et al*. Grey matter volume in a large cohort of MS patients: relation to MRI parameters and disability. *Mult Scler* 2011;17:1098–106.
5 Fisher E, Lee J-C, Nakamura K, *et al*. Gray matter atrophy in multiple sclerosis: a longitudinal study. *Ann Neurol* 2008;64:255–65.
6 Montalban X, Hauser SL, Kappos L, *et al*. Ocrelizumab versus placebo in primary progressive multiple sclerosis. *N Engl J Med* 2017;376:209–20.
7 Kappos L, Bar-Or A, Cree BAC, *et al*. Siponimod versus placebo in secondary progressive multiple sclerosis (expand): a double-blind, randomised, phase 3 study. *Lancet* 2018;391:1263–73.
8 De Stefano N, Giorgio A, Battaglini M, *et al*. Assessing brain atrophy rates in a large population of untreated multiple sclerosis subtypes. *Neurology* 2010;74:1868–76.
9 Steenwijk MD, Geurts JJG, Daams M, *et al*. Cortical atrophy patterns in multiple sclerosis are non-random and clinically relevant. *Brain* 2016;139:115–26.
10 Azevedo CJ, Cen SY, Khadka S, *et al*. Thalamic atrophy in multiple sclerosis: a magnetic resonance imaging marker of neurodegeneration throughout disease. *Ann Neurol* 2018;83:223–34.
11 Dekker I, Eijlers AJC, Popescu V, *et al*. Predicting clinical progression in multiple sclerosis after 6 and 12 years. *Eur J Neurol* 2019;26:893–902.
12 Pardini M, Yaldizli Özgür, Sethi V, *et al*. Motor network efficiency and disability in multiple sclerosis. *Neurology* 2015;85:1115–22.
13 Pichet Binette A, Gonneaud J, Vogel JW, *et al*. Morphometric network differences in ageing versus Alzheimer's disease dementia. *Brain* 2020;143:635–49.
14 Langlois D, Chartier S, Gosselin D. An introduction to independent component analysis: InfoMax and FastICA algorithms. *Tutor Quant Methods Psychol* 2016;6:31–8.
15 Chard DT, Miller DH. What lies beneath grey matter atrophy in multiple sclerosis? *Brain* 2016;139:7–10.
16 Gupta CN, Turner JA, Calhoun VD. Source-based morphometry: a decade of covarying structural brain patterns. *Brain Struct Funct* 2019;224:3031–44.
17 Bergsland N, Horakova D, Dwyer MG, *et al*. Gray matter atrophy patterns in multiple sclerosis: a 10-year source-based morphometry study. *Neuroimage Clin* 2018;17:444–51.
18 Pagani E, Rocca MA, Gallo A, *et al*. Regional brain atrophy evolves differently in patients with multiple sclerosis according to clinical phenotype. *AJNR Am J Neuroradiol* 2005;26:341–6.
19 Honce JM. Gray matter pathology in MS: neuroimaging and clinical correlations. *Mult Scler Int* 2013;2013:1–16.
20 Kapoor R, Ho P-R, Campbell N, *et al*. Effect of natalizumab on disease progression in secondary progressive multiple sclerosis (ASCEND): a phase 3, randomised, double-blind, placebo-controlled trial with an open-label extension. *Lancet Neurol* 2018;17:405–15.
21 Eshaghi A, Kievit RA, Prados F. *Application of mechanistic methods to clinical trials in multiple sclerosis: the simvastatin case*, 2018: 1–42.
22 Tustison NJ, Avants BB, Cook PA. *N4ITK : Improved N3 Bias Correction*, 2010: 1310–20.
23 Prados F, Cardoso MJ, Kanber B, *et al*. A multi-time-point modality-agnostic patch-based method for lesion filling in multiple sclerosis. *Neuroimage* 2016;139:376–84.
24 Cardoso MJ, Modat M, Wolz R, *et al*. Geodesic information flows: Spatially-Variant graphs and their application to segmentation and fusion. *IEEE Trans Med Imaging* 2015;34:1976–88.
25 Klein A, Tourville J. 101 labeled brain images and a consistent human cortical labeling protocol. *Front Neurosci* 2012;6:171.
26 Prinster A, Quarantelli M, Orefice G, *et al*. Grey matter loss in relapsing-remitting multiple sclerosis: a voxel-based morphometry study. *Neuroimage* 2006;29:859–67.
27 Tustison NJ, Avants BB. Explicit B-spline regularization in diffeomorphic image registration. *Front Neuroinform* 2013;7:1–13.
28 Ashburner J, Friston KJ. Voxel-based morphometry--the methods. *Neuroimage* 2000;11:805–21.
29 Hyvärinen A, Oja E. *Independent component analysis: algorithms and applications*, 2000.
30 Esposito F, Scarabino T, Hyvarinen A, *et al*. Independent component analysis of fMRI group studies by self-organizing clustering. *Neuroimage* 2005;25:193–205.
31 Prestel M, Steinfath TP, Tremmel M, *et al*. fMRI BOLD correlates of EEG independent components: spatial correspondence with the default mode network. *Front Hum Neurosci* 2018;12:1–13.
32 Smith SM, Fox PT, Miller KL, *et al*. Correspondence of the brain's functional architecture during activation and rest. *Proc Natl Acad Sci U S A* 2009;106:13040–5.
33 Dunlop K, Talishinsky A, Liston C. Intrinsic brain network biomarkers of antidepressant response: a review. *Curr Psychiatry Rep* 2019;21:87.
34 Schoonheim MM, Geurts JJG. What causes deep gray matter atrophy in multiple sclerosis? *AJNR Am J Neuroradiol* 2019;40:107–8.
35 Eshaghi A, Marinescu RV, Young AL, *et al*. Progression of regional grey matter atrophy in multiple sclerosis. *Brain* 2018;141:1665–77.
36 Tur C, Moccia M, Barkhof F, *et al*. Assessing treatment outcomes in multiple sclerosis trials and in the clinical setting. *Nat Rev Neurol* 2018;14:75–93.
37 Lublin F, Miller DH, Freedman MS, *et al*. Oral fingolimod in primary progressive multiple sclerosis (informs): a phase 3, randomised, double-blind, placebo-controlled trial. *Lancet* 2016;387:1075–84.
38 D'hooghe MB, Gielen J, Van Remoortel A, *et al*. Single MRI-based volumetric assessment in clinical practice is associated with MS-related disability. *J Magn Reson Imaging* 2019;49:1312–21.
39 Benedict RH, DeLuca J, Phillips G, *et al*. Validity of the symbol digit modalities test as a cognition performance outcome measure for multiple sclerosis. *Mult Scler* 2017;23:721–33.
40 Uno H, Cai T, Pencina MJ, *et al*. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Stat Med* 2011;30:1105–17.
41 Magon S, Tsagkas C, Gaetano L, *et al*. Volume loss in the deep gray matter and thalamic subnuclei: a longitudinal study on disability progression in multiple sclerosis. *J Neurol* 2020;267:1536–46.
42 Biberacher V, Boucard CC, Schmidt P, *et al*. Atrophy and structural variability of the upper cervical cord in early multiple sclerosis. *Mult Scler* 2015;21:875–84.
43 Poonawalla AH, Datta S, Juneja V, *et al*. Composite MRI scores improve correlation with EDSS in multiple sclerosis. *Mult Scler* 2010;16:1117–25.
44 Zivadinov R, Leist TP. Clinical-magnetic resonance imaging correlations in multiple sclerosis. *J Neuroimaging* 2005;15:10S–21.
45 Barkhof F. The clinico-radiological paradox in multiple sclerosis revisited. *Curr Opin Neurol* 2002;15:239–45.
46 Wolff M, Vann SD. The cognitive thalamus as a gateway to mental representations. *J Neurosci* 2019;39:3–14.
47 Herrero M-T, Barcia C, Navarro J. Functional anatomy of thalamus and basal ganglia. *Child's Nervous System* 2002;18:386–404.
48 Simonyan K. Recent advances in understanding the role of the basal ganglia. *F1000Res* 2019;8:122.
49 Batista S, Zivadinov R, Hoogs M, *et al*. Basal ganglia, thalamus and neocortical atrophy predicting slowed cognitive processing in multiple sclerosis. *J Neurol* 2012;259:139–46.
50 Calabrese M, Magliozzi R, Ciccarelli O, *et al*. Exploring the origins of grey matter damage in multiple sclerosis. *Nat Rev Neurosci* 2015;16:147–58.
51 Hametner S, Wimmer I, Haider L, *et al*. Iron and neurodegeneration in the multiple sclerosis brain. *Ann Neurol* 2013;74:848–61.
52 Zhang D, Raichle ME. Disease and the brain's dark energy. *Nat Rev Neurol* 2010;6:15–28.
53 Trapp BD, Stys PK. Virtual hypoxia and chronic necrosis of demyelinated axons in multiple sclerosis. *Lancet Neurol* 2009;8:280–91.
54 Wolinsky JS, Narayana PA, O'Connor P, *et al*. Glatiramer acetate in primary progressive multiple sclerosis: results of a multinational, multicenter, double-blind, placebo-controlled trial. *Ann Neurol* 2007;61:14–24.
55 Meijer KA, Cercignani M, Muhlert N, *et al*. Patterns of white matter damage are non-random and associated with cognitive function in secondary progressive multiple sclerosis. *Neuroimage Clin* 2016;12:123–31.

56 Fleischer V, Koirala N, Droby A, *et al*. Longitudinal cortical network reorganization in early relapsing-remitting multiple sclerosis. *Ther Adv Neurol Disord* 2019;12:1–15.
57 Tur C, Kanber B, Eshaghi A, *et al*. Clinical relevance of cortical network dynamics in early primary progressive MS. *Mult Scler* 2020;26:442–56.
58 Berrigan LI, Fisk JD, Walker LAS, *et al*. Reliability of regression-based normative data for the oral symbol digit modalities test: an evaluation of demographic influences, construct validity, and impairment classification rates in multiple sclerosis samples. *Clin Neuropsychol* 2014;28:281–99.
59 Parmenter BA, Testa SM, Schretlen DJ, *et al*. The utility of regression-based norms in interpreting the minimal assessment of cognitive function in multiple sclerosis (MACFIMS). *J Int Neuropsychol Soc* 2010;16:6–16.
60 Marrie RA, Whitehouse CE, Patel R, *et al*. Performance of regression-based norms for cognitive functioning of persons with multiple sclerosis in an independent sample. *Front Neurol* 2021;11:1–13.

Supplementary material

Image acquisition

A representative MRI protocol from one centre included the following image acquisitions:

(1) 2D T1-weighted: TR= 30 milliseconds, TE= 11.0 milliseconds, slice thickness= 3.0mm, imaging frequency= 63.43, acquisition matrix= 0x256x192x0, flip angle= 30.0, number of phase encoding steps= 192, FoV= 187*250mm;

(2) 2D fast fluid- attenuated inversion recovery (2D FLAIR): with voxel size= 0.98 × 0.98, TR= 2980.0 milliseconds, TE = 12 milliseconds, slice thickness= 3.0mm, imaging frequency= 63.643, number of phase encoding steps= 258, matrix size = 0x256x256x0, flip angle= 180.0, FoV= 250*250mm;

(3) 2D T2-weighted: with voxel size= 0.98 × 0.98, TR= 5750.0 milliseconds, TE = 86 milliseconds, slice thickness = 3.0mm, imaging frequency= 63.643, number of phase encoding steps= 259, matrix size = 0x256x256x0, flip angle= 180.0, FoV= 250*250mm;

(4) 2D proton density (PD): with voxel size= 0.98 × 0.98, TR= 2980.0 milliseconds, TE = 12 milliseconds, slice thickness = 3.0mm, imaging frequency= 63.643, number of phase encoding steps= 258, matrix size = 0x256x256x0, flip angle= 180.0, FoV= 250*250mm;

Participants underwent either 1.5 or 3.0 Tesla.

MANUFACTURER	**MAGNETIC FIELD**			
	1.5 T	3.0 T	Missing	**TOTAL**
GE MEDICAL SYSTEMS	293	46	-	339
Marconi Medical Systems, Inc.	5	-	-	5
Philips Medical Systems	153	70	-	223
SIEMENS	366	51	-	417
TOSHIBA_MEC	-	-	3	3
Missing	-	1	-	1
TOTAL	817	168	3	988

Study-specific template construction

We randomly selected 39 subjects to create a study-specific template in ANTs. The number of subject was chosen to be as large as feasible, based on computation time, and as significantly larger than previous cohorts used for this purpose (e.g., N= 19 (Eshaghi et al., 2014), N= 20 (Whitwell *et al.*, 2007)).

Gender (M/F)	29/ 10
Age (mean±SD)	46.78±6.45
Trial Arm (DMT/Placebo)	420/422
EDSS (median, range)	6 (3-7)
SDMT (mean±SD)	40.04±15.70
9HPT (mean±SD)	32.17±13.29

We interpolated the T1w lesion filled scans to 1x1x1mm, and registered them using a rigid body transformation to the MNI152 space (http://www.bic.mni.mcgill.ca/ServicesAtlases/ICBM152NLin2009).

MANUFACTURER	MAGNETIC FIELD		
	1.5 T	3.0 T	TOTAL
GE MEDICAL SYSTEMS	8	1	9
Marconi Medical Systems, Inc.	2	0	2
Philips Medical Systems	9	5	14
SIEMENS	11	3	14
TOSHIBA_MEC	-	-	-
Missing	-	-	-
TOTAL	30	9	39

To determine whether the randomly selected sub-sample was representative of the cohort, we performed two-sample Kolmogorov-Smirnov test for continuous variables, and two-sample test for equality of proportions for the categorical ones. The randomly selected sub-sample did not significantly differ from the entire cohort when considering age (D= 0.11, p= 0.72), baseline mean whole brain GM volume (D= 0.10, p= 0.85), baseline mean DGM volume measures (D= 0.13, p= 0.56), mean 9HPT score (D= 0.17, p= 0.28), mean SDMT score (D=

0.18, p= 0.75), and EDSS (D= 0.11, p= 0.80). The proportion of participants scanned at 1.5 and 3T did not differ in the two samples (respectively, X^2 = 0.86, df = 1, p= 0.35, CI [-0.19:0.077], and X^2 =0.97, df = 1, p= 0.33, CI [-0.07: 0.20]). The proportion of females and males in the two samples did not differ (respectively, X^2 = 2.10, df = 1, p= 0.15, CI [-0.03: 0.25], and X^2 = 2.10, df = 1, p= 0.15, CI [-0.25: 0.03]).

Whole brain mask

We created a whole brain mask to: (1) to constrain the ICA to the parenchymal brain, and (2) identify and label brain regions involved in each ICA-component. We used the majority voting algorithm implemented in antsJointLabelFusion to fuse the 39 parcellation maps of subjects that had contributed to the template and obtain a single parcellation map where the likelihood that each region corresponds to the labelled one is higher (H. Wang et al., 2013; H. Wang & Yushkevich, 2013). We used this atlas-mask to label brain regions involved in each component. We excluded brain structures from the brain mask (e.g., ventricles, meninges, etc.), binarised the output and used the obtained mask to mask the processed GM probability maps and reduce the computational expense of the ICA.

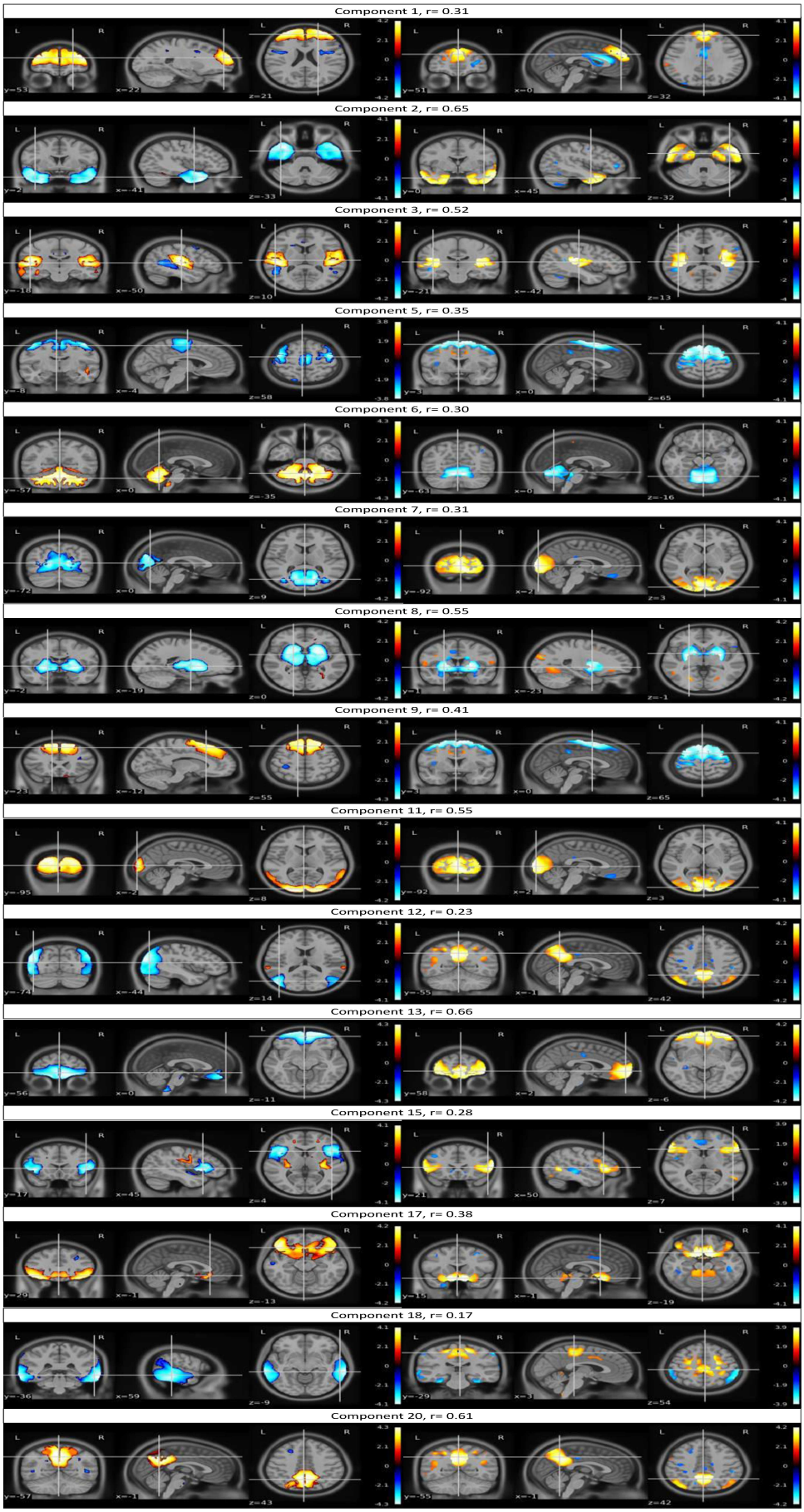
Association across baseline regional volumes and ICA loadings

We determined the association of each ICA-components with baseline volume of the encompassed brain regions as follows. From the pre-processing stage, for each participant we obtained a parcellation map, and estimated the baseline volume measure of whole brain GM and of single brain regions (higher values correspond to bigger volumes). By overlaying the identified ICA-components with the whole brain mask created from the parcellation maps (as described above), we determined which brain regions were involved in each component. We correlated the loading of ICA-components (generated from baseline scans) with the volume of each brain region encompassed by the GM pattern. In this way, we could determine which brain regions in each component presented a positive correlation with the baseline volume of that region (higher baseline volume higher loading) and which presented a negative correlation (higher baseline volume, lower loading).

ICA in healthy controls

To determine whether similar components to the ones identified in patients with SPMS were present in HCs, we randomly selected a sample of 400 participants from the Human Connectome Project. We applied the same pipeline to analyse T1 and T2 scans. To increase the accuracy of the registrations, we substituted the customised study specific template (created from participants with MS and reconstructed 2D scans), with the MNI template (created from 3D images of healthy controls). To determine whether the same components were present in healthy controls and in patients with MS, we visually inspected the resulting ICA components and calculated spatial cross-correlations.

Components	Correlation
1	0.31
2	0.65
3	0.52
5	0.35
6	0.30
7	0.31
8	0.55
9	0.41
11	0.55
12	0.23
13	0.66
15	0.28
17	0.38
18	0.17
20	0.61


Component 1, r= 0.31
Component 2, r= 0.65
Component 3, r= 0.52
Component 5, r= 0.35
Component 6, r= 0.30
Component 7, r= 0.31
Component 8, r= 0.55
Component 9, r= 0.41
Component 11, r= 0.55
Component 12, r= 0.23
Component 13, r= 0.66
Component 15, r= 0.28
Component 17, r= 0.38
Component 18, r= 0.17
Component 20, r= 0.61

Nine GM patterns identified from patients with SPMS were uniquely correlated and present also in HCs, suggesting that they are not disease specific. Although slightly overlapping with HC patterns, two SPMS ICA-components (component 12 and 18) were detected in patients with MS but not in HCs, which might be disease specific.

Supplementary Tables

Table s1. *Results of the spatial cross-correlations between the 20 components obtained in each of the four sub-sample and from the entire cohort*

Network	Sub-sample 1 vs. 2	Sub-sample 1 vs. 3	Sub-sample 1 vs.4	Entire cohort vs. sub-sample 1	Sub-sample 2 vs. 3	Sub-sample 2 vs. 4	Entire cohort vs. sub-sample 2	Sub-sample 3 vs. 4	Entire cohort vs. sub-sample 3	Entire cohort vs. sub-sample 4
1	0.86	0.80	0.84	0.89	0.82	0.88	0.91	0.82	0.84	0.88
2	0.88	0.88	0.87	0.90	0.92	0.90	0.95	0.89	0.93	0.92
3	0.41	0.30	0.52	0.35	0.66	0.36	0.54	0.48	0.63	0.73
5	0.37	0.61	0.50	0.43	0.28	0.16	0.41	0.40	0.42	0.14
6	0.94	0.94	0.93	0.94	0.95	0.95	0.97	0.94	0.95	0.96
7	0.10	0.27	0.20	0.22	0.55	0.51	0.63	0.60	0.73	0.75
8	0.93	0.90	0.88	0.95	0.88	0.93	0.95	0.82	0.90	0.91
9	0.62	0.61	0.64	0.75	0.56	0.69	0.77	0.62	0.72	0.76
11	0.78	0.62	0.41	0.78	0.53	0.35	0.74	0.53	0.81	0.52
12	0.48	0.69	0.47	0.82	0.51	0.20	0.55	0.24	0.83	0.39
13	0.85	0.78	0.86	0.92	0.76	0.83	0.89	0.77	0.82	0.90
15	0.41	0.42	0.52	0.52	0.20	0.33	0.18	0.43	0.69	0.59
17	0.40	0.63	0.60	0.69	0.52	0.46	0.51	0.66	0.81	0.71
18	0.63	0.51	0.44	0.53	0.55	0.61	0.61	0.33	0.82	0.36
20	0.69	0.67	0.62	0.72	0.80	0.70	0.81	0.75	0.90	0.80

Table legend: The table reports the spatial cross-correlations results for the stable ICA-components (i.e., showed a statistically significant correlations ($p<0.05$) in all 4 folds and in the entire cohort).

Table s2. *Results of the spatial cross-correlations across the 15 stable components obtained from the entire cohort and those obtained in each of the five cross-validation folds (based on acquisition region).*

COMPONENTS	NORTHERN AMERICA	NORTHERN EUROPE	SOUTHERN EUROPE	EASTERN EUROPE	WESTERN EUROPE
V1	0.68	0.91	0.68	0.68	0.81
V2	0.89	0.93	0.80	0.88	0.89
V3	0.68	0.72	0.19	0.66	0.53
V5	0.22	0.24	0.20	0.51	0.21
V6	0.95	0.90	0.76	0.96	0.93
V7	0.71	0.68	0.54	0.84	0.41
V8	0.93	0.84	0.86	0.94	0.84
V9	0.35	0.73	0.39	0.29	0.11
V11	0.83	0.76	0.79	0.83	0.82
V12	0.58	0.63	0.40	0.79	0.70
V13	0.54	0.74	0.69	0.82	0.70
V15	0.46	0.17	0.24	0.21	0.24
V17	0.42	0.42	0.27	0.85	0.64
V18	0.11	0.62	0.21	0.77	0.44
V20	0.35	0.80	0.52	0.84	0.71

Table s3. *Results of the spatial cross-correlations across the 15 stable components obtained from the entire cohort and those obtained in each of the three cross-validation folds (acquisition manufacturer).*

COMPONENTS	SIEMENS	GE	PHILIPS
V1	0.85	0.85	0.78
V2	0.96	0.92	0.91
V3	0.61	0.71	0.63
V5	0.36	0.43	0.35
V6	0.98	0.97	0.88
V7	0.86	0.75	0.69
V8	0.92	0.94	0.64
V9	0.36	0.82	0.19
V11	0.66	0.74	0.75
V12	0.46	0.53	0.60
V13	0.96	0.90	0.83
V15	0.52	0.68	0.40
V17	0.83	0.76	0.37
V18	0.59	0.66	0.62
V20	0.83	0.90	0.69

Table legend: The spatial cross-correlations results for the stable ICA-components (i.e., showed a statistically significant correlations ($p<0.05$) between the entire cohort and the 5 cross-validation folds representing the acquisition manufacturer).

Table s4. *List of the 15 stable components with their corresponding involved brain regions*

Networks	Regions
Component 1	▲Superior frontal gyrus, ▲Middle frontal gyrus, ▲Superior frontal gyrus medial segment, ▲Anterior cingulate gyrus, ▲Opercular part of the inferior frontal gyrus
Component 2	▼Temporal pole, ▼Inferior temporal gyrus, ▼Middle temporal gyrus, ▼Middle cingulate gyrus, ▼Parahippocampal gyrus, ▼Precentral gyrus medial segment, ▼Posterior cingulate gyrus, ▼Entorhinal area, ▼Parietal lobule, ▼ Fusiform gyrus
Component 3	▲Parietal and ▲Temporal pole, ▲Middle temporal gyrus, ▲Superior temporal gyrus, ▲Planum temporale, ▲ Parietal operculum, ▲Planum polare, ▲Central operculum, ▲Middle cingulate gyrus, ▲Cingulate, ▲Middle frontal gyrus, ▲Posterior Insula, ▲Transverse temporal gyrus
Component 5	▼Middle temporal gyrus, ▼Parietal lobule, ▼Supramarginal gyrus, ▼Postcentral gyrus, ▼Precentral gyrus, ▼ Triangular part of the inferior frontal gyrus, ▼Middle frontal gyrus, ▲Superior temporal gyrus, ▼Angular gyrus, ▼ Cuneus, ▼Superior frontal gyrus, ▼Precentral gyrus medial segment, ▼Inferior occipital gyrus
Component 6	▲Cerebellum, ▲Brain Stem, ▲Pons, ▼Lingual gyrus, ▼Fusiform gyrus, ▼Temporal and ▼Parietal lobe
Component 7	▼Superior occipital gyrus, ▼Occipital lobe, ▼Lingual gyrus, ▼Calcalcarine cortex, ▼Precuneus, ▼Parietal lobe, ▼Temporal lobe, ▲Middle temporal gyrus, ▼Frontal lobe, ▲Precentral gyrus, ▼Supramarginal gyrus
Component 8	▼Brain Stem, ▼Pons, ▼Ventral DC, ▼Thalamus, ▼Insula, ▼Accumbens, ▼Caudate, ▼Putamen, ▼Pallidum, ▼ Frontal and ▲Temporal lobe
Component 9	▲Angular gyrus, ▲Middle occipital gyrus, ▲Postcentral gyrus, ▲Medial orbital gyrus, ▲Triangular part of the inferior frontal gyrus, ▲Middle frontal gyrus, ▲Superior frontal gyrus medial segment, ▲Superior frontal gyrus, ▲ Superior parietal lobule, ▲Claustrum, ▲Occipital, ▲Frontal, and ▲Parietal pole
Component 11	▲Occipital pole, ▲Calcarine cortex, ▲Cuneus, ▼Middle temporal gyrus, ▼Inferior temporal gyrus, ▲Inferior occipital gyrus, ▼Angular gyrus, ▼Superior parietal lobule, ▼Supramarginal gyrus
Component 12	▼Angular gyrus, ▼Inferior temporal gyrus, ▼Middle temporal gyrus, ▼Inferior occipital gyrus, ▼Middle occipital gyrus, ▼Superior occipital gyrus, ▼Fusiform gyrus, ▼Occipital fusiform gyrus, ▼Precentral gyrus, ▼Middle frontal gyrus, ▼Claustrum, ▼Superior frontal gyrus
Component 13	▼Lateral orbital gyrus, ▼Middle frontal gyrus, ▼Superior frontal gyrus, ▼Superior frontal gyrus medial segment, ▼ Anterior orbital gyrus, ▼Medial frontal cortex, ▼Gyrus rectus, ▼Frontal pole, ▼Medial orbital gyrus, ▼Anterior cingulate gyrus, ▼Brain Stem, ▼Lingual gyrus, ▼Temporal pole
Component 15	▲Thalamus, ▲Caudate, ▼Anterior insula, ▼Posterior insula, ▼Planum polare, ▲Putamen, ▼Frontal operculum, ▼Planum temporale, ▼Claustrum, ▼Triangular part of the inferior frontal gyrus, ▼Opercular part of the inferior frontal gyrus, ▲Precentral gyrus, ▼Central operculum, ▼Parietal operculum, ▼Frontal and ▼Temporal pole
Component 17	▲Hippocampus, ▲Pons, ▲Middle temporal gyrus, ▲Superior temporal gyrus, ▲Postcentral gyrus, ▲Triangular part of the inferior frontal gyrus, ▲Temporal pole, ▲Posterior orbital gyrus, ▲Medial orbital gyrus, ▲Anterior insula, Claustrum, ▲Basal Forebrain, ▲Putamen, ▲Subcallosal area, ▲Medial orbital gyrus, ▲Gyrus rectus, ▲Medial frontal cortex, ▲Lateral orbital gyrus, ▲Orbital part of the inferior frontal gyrus, ▲Medial frontal cortex, ▲Anterior

	cingulate gyrus, ▲Anterior orbital gyrus, ▲Posterior cingulate gyrus, ▲Postcentral gyrus, ▲Frontal operculum, ▲ Inferior temporal gyrus
Component 18	▼Middle occipital gyrus, ▼Postcentral gyrus, ▼Precentral gyrus, ▼Opercular part of the inferior frontal gyrus, ▼ Fusiform gyrus, ▼Parahippocampal gyrus, ▼Frontal, ▼Occipital and ▼Parietal lobe, ▼Inferior temporal gyrus, ▼ Middle temporal gyrus, ▼Superior temporal gyrus, ▼Supramarginal gyrus, ▼Middle temporal gyrus
Component 20	▲Superior occipital gyrus, ▲Superior parietal lobule, ▲Precuneus, ▲Posterior cingulate gyrus, ▲Superior frontal gyrus, ▲Middle frontal gyrus, ▲Angular gyrus, ▲Occipital lobule

▲ Relative preserved brain region

▼ Atrophic brain region

Table legend: We overlaid a whole brain parcellation mask with the identified ICA-components in order to retrieve and label brain regions involved in each network. We correlated the loading of ICA-components with the baseline volume of the areas involved in each network to identify which brain area in each network showed volume loss (negative correlation between network loading and baseline volume) and which represented relative brain preservation (negative correlation between those volumes and ICA-loadings).

Table s5. *Correlations between the loading values of ICA-components and whole brain GM volumes, SDMT, EDSS, and 9HPT score*

GM NETWORKS	WHOLE BRAIN GM VOLUME	EDSS N= 830	9HPT N= 829	SDMT N= 391
Component 1	r= 0.08 95% CI [0.01:0.14] p= 0.28	rho= - 0.01 p= 0.79	r= 0.13 95% CI [0.06: 0.19] p< 0.01*	r= 0.22 95% CI [0.13:0.32] p<0.001*
Component 2	r= - 0.25 95%CI [-0.3:-0.19] p<0.001*	rho= 0.08 p= 0.15	r= -0.04 95% CI [-0.11: 0.03] p= 0.33	r= - 0.07 95% CI [- 0.17:0.03] p= 1
Component 3	r= 0.18 95%CI [0.11:0.23] p<0.001*	rho= - 0.05 p= 0.22	r= 0.05 95% CI [-0.02:0.11] p= 0.31	r= 0.04 95% CI [- 0.06:0.14] p= 1
Component 5	r= - 0.16 95%CI [-0.22:-0.09] p<0.001*	rho= - 0.03 p= 0.39	r= 0.05 95% CI [- 0.02: 0.11] p= 0.35	r= 0.07 95% CI [- 0.03:0.16] p= 1
Component 6	r= - 0.09 95% CI [-0.15:-0.02] p=0.10	rho= - 0.11 p= 0.02*	r= 0.19 95% CI [0.12: 0.25] p <0.001*	r= 0.10 95% CI [0:0.19] p= 1
Component 7	r= - 0.07 95% CI [-0.13:-0.01] p= 0.38	rho= 0.04 p= 0.36	r= 0.02 95% CI [- 0.05: 0.09] p= 0.63	r= - 0.06 95% CI [- 0.16:0.04] p= 1
Component 8	r= - 0.18 95% CI [-0.24:-0.12] p<0.001*	rho= 0.06 p= 0.22	r= -0.32 95% CI [-0.38:-0.25] p<0.001*	r= - 0.44 95% CI [- 0.52: - 0.36] p<0.001*
Component 9	r= 0.20 95% CI	rho= 0.06	r= -0.06 95% CI	r= - 0.01 95% CI

	[0.14:0.26] p<0.001*	p= 0.21	[-0.13:0.01] p= 0.21	[- 0.11:0.09] p= 1
Component 11	r= - 0.12 95% CI [-0.17:-0.05] p<0.01*	rho= 0.07 p= 0.12	r= 0.11 95% CI [0.04: 0.17] p< 0.01*	r= 0.16 95% CI [0.06:0.25] p<0.05*
Component 12	r= - 0.17 95% CI [-0.23:-0.11] p<0.001*	rho= - 0.04 p= 0.35	r= -0.03 95% CI [- 0.09:0.04] p= 0.56	r= - 0.12 95% CI [- 0.21: -0.02] p=0.39
Component 13	r= - 0.38 95% CI [-0.43:-0.33] p<0.001*	rho= 0.07 p= 0.16	r= -0.06 95% CI [- 0.12:0.01] p= 0.24	r= - 0.12 95% CI [- 0.21: - 0.02] p=0.39
Component **1**5	r= - 0.01 95% CI [-0.07:0.06] p= 1	rho= 0.07 p= 0.18	r= -0.04 95% CI [- 0.11:0.03] p= 0.36	r= - 0.13 95% CI [- 0.23: - 0.03] p= 0.21
Component **1**7	r= 0.19 95% CI [0.13:0.25] p<0.001*	rho= 0.02 p= 0.65	r= 0.02 95% CI [- 0.05:0.09] p= 0.61	r= 0.03 95% CI [- 0.07:0.13] p= 1
Component **1**8	r= - 0.28 95% CI [-0.33:-0.22] p<0.001*	rho= 0.08 p= 0.19	r= - 0.03 95% CI [- 0.09: 0.04] p= 0.53	r= - 0.15 95% CI [- 0.25: - 0.05] p<0.05*
Component 20	r= 0.28 95% CI [0.22:0.33] p<0.001*	rho= 0.05 p= 0.22	r= 0.08 95% CI [0.01: 0.15] p= 0.08	r= 0.18 95% CI [0.08: 0.27] p<0.01*

Acronyms: GM= grey matter, EDSS= expanded disability status scale; 9HPT= nine hole peg test; SDMT= symbol digit modalities test; CI= confidence interval

*Statistically significant after Bonferroni correction

We report the Bonferroni Adjusted P values

Table s6. *Correlations across the SDMT, EDSS, and 9HPT score at baseline and the lesion load, DGM, GM, and thalamus volumes.*

MRI MEASURES	EDSS N= 830	9HPT N= 829	SDMT N= 391
Whole brain GM	rho= - 0.09 p= 0.052	r= 0.08 95% CI [0.01: 0.14] p<0.05*	r= 0.22 95% CI [0.12:0.31] p= 0.18
Deep GM	rho= - 0.07 p<0.05*	r= 0.06 95% CI [- 0.01:0.13] p= 0.09	r= 0.13 95% CI [0.03:0.23] p<0.001**
Lesion load	rho= 0.07 p= 0.06	r= -0.29 95% CI [-0.35:-0.23] p< 0.001*	r= -0.43 95% CI [-0.51:-0.35] p<0.001*
Thalamus	rho= - 0.12 p< 0.005*	r= 0.27 95% CI [0.21: 0.33] p<0.001*	r= 0.41 95% CI [0.32:0.49] p<0.001*

Acronyms: GM= grey matter, EDSS= expanded disability status scale; 9HPT= nine-hole peg test; SDMT= symbol digit modalities test; CI= confidence interval

*Statistically significant after Bonferroni correction

We report the Bonferroni Adjusted P values

Table s7. *Survival analysis for the EDSS, 9HPT, and SDMT progression*

Predictors	EDSS progression confirmed at 3 months			20% 9HPT worsening			10% SDMT worsening		
Brain regions	**HR**	**95% CI**	**p-value**	**HR**	**95% CI**	**p-value**	**HR**	**95% CI**	**p-value**
Component 1	1.00	0.86-1.16	0.97	1.18	0.96-1.45	0.11	0.85	0.72-1.01	0.06
Component 2	1.10	0.94-1.27	0.24	1.30	1.06-1.60	0.01**	1.03	0.86-1.22	0.76
Component 3	1.02	0.88-1.17	0.82	1.03	0.86-1.25	0.73	0.97	0.82-1.15	0.69
Component 5	0.90	0.78-1.04	0.15	0.92	0.76-1.11	0.38	1.07	0.91-1.26	0.44
Component 6	1.04	0.90-1.20	0.64	0.88	0.72-1.08	0.21	0.97	0.81-1.16	0.74
Component 7	1.05	0.91-1.21	0.50	0.86	0.72-1.03	0.10	0.82	0.71-0.96	0.01*
Component 8	1.04	0.90-1.20	0.58	1.02	0.84-1.24	0.84	1.29	1.09-1.52	p<0.003***
Component 9	0.98	0.84-1.13	0.75	0.95	0.77-1.16	0.58	0.91	0.76-1.09	0.30
Component 11	0.98	0.85-1.14	0.82	1.03	0.85-1.24	0.76	0.93	0.78-1.12	0.46
Component 12	1.01	0.88-1.17	0.88	0.85	0.71-1.02	0.08	1.14	0.96-1.34	0.13
Component 13	1.01	0.87-1.16	0.95	0.94	0.78-1.13	0.49	1.28	1.08-1.51	p<0.005***
Component 15	0.99	0.86-1.14	0.87	1.07	0.89-1.28	0.50	1.25	1.05-1.50	0.01**
Component 17	0.97	0.84-1.12	0.68	0.89	0.74-1.07	0.22	0.84	0.71-0.99	0.04*
Component 18	1.08	0.94-1.25	0.30	1.14	0.95-1.36	0.17	1.18	1.00-1.40	0.046*
Component 20	1.05	0.91-1.21	0.55	1.21	1.01-1.45	0.04*	1.00	0.85-1.17	0.96
Whole GM	0.98	0.61-1.02	0.07	0.90	0.71-1.14	0.38	0.83	0.67-1.03	0.09
DGM	0.79	0.81-1.19	0.86	0.72	0.52-0.99	0.05*	0.95	0.71-1.26	0.72
Lesion load	0.99	0.87-1.12	0.83	0.98	0.82-1.16	0.79	1.32	1.16-1.50	p<0.001****
Thalamus	0.96	0.81-1.14	0.63	1.01	0.81-1.26	0.90	0.82	0.68-1.00	0.048*

*Caption: *p<0.05 ** p<0.01 ***p<0.005 ****p<0.001*

Table legend: We report here the results for the univariate Cox regression models for all the assessed independent variables (predictors). For each standard deviation unit decrease in the baseline caudate volume the risk of developing a confirmed disability progression increased by the 19%. Two ICA-components (component 2 and component 20), and the baseline volume of the DGM predicted the 20% worsening of the 9HPT. Six ICA-components, lesion load, and the volume of the thalamus predicted the 10% SDMT worsening.

Acronyms: EDSS= expanded disability status scale; 9HPT= nine-hole peg test; SDMT= symbol digit modalities test; HR= hazard ratio, CI= confidence interval; DGM= deep grey matter

Figure s1. *Cumulative hazard function for the EDSS progression confirmed at 3 months.*

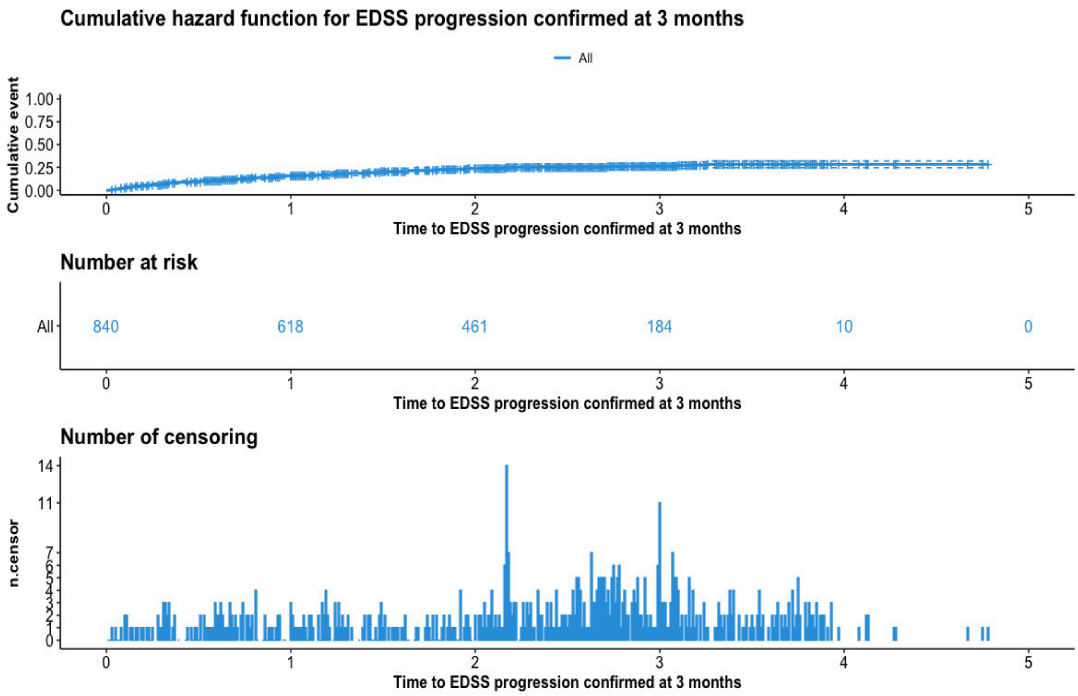


Figure legend: The figure shows the cumulative hazard function for the EDSS progression confirmed at 3 months. By the end of the study, the 28.5% of patients reported a 12-week confirmed EDSS progression. The number of censored and the number of subjects at risk of developing the EDSS progression at each timepoint are also reported

Figure s2. *Cumulative hazard function for the 20% 9HPT worsening*

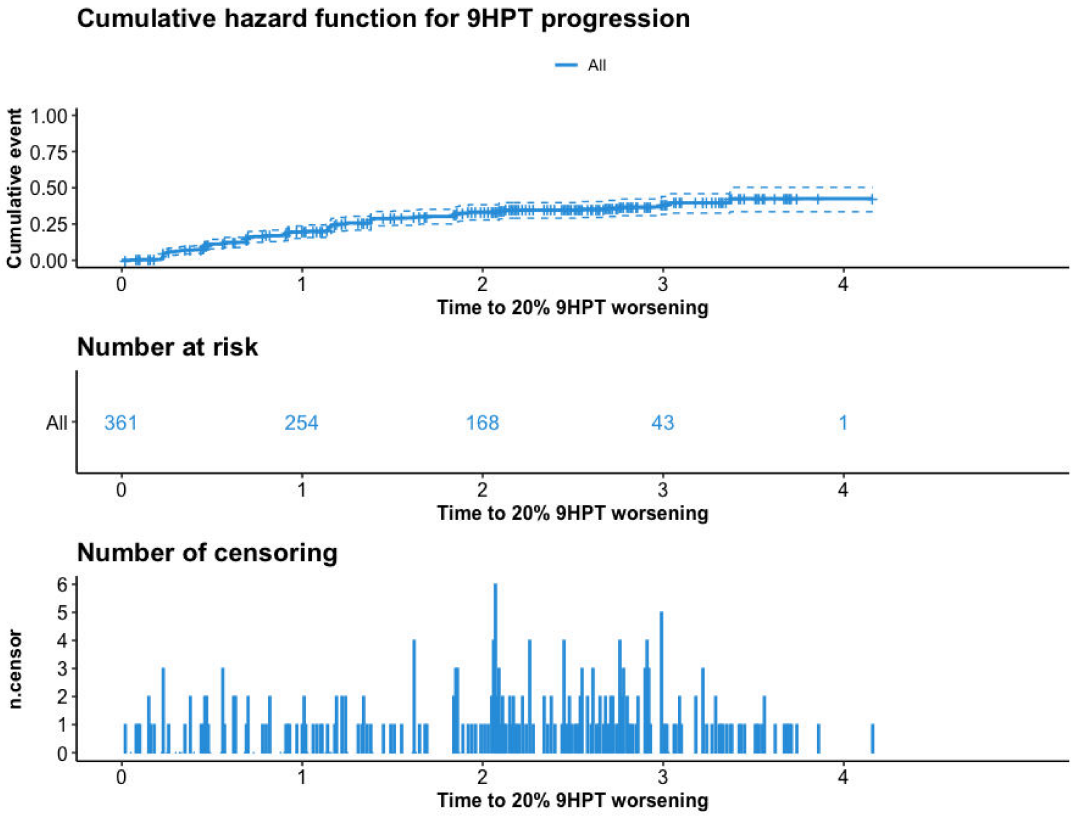


Figure legend: The figure shows the cumulative hazard function for the 20% 9HPT worsening. By the end of the study, the 42% of patients reported a 9HPT progression. The number of censored and the number of subjects at risk of developing the 9HPT progression at each timepoint are also reported

Figure s3. *Cumulative hazard function for the 10% SDMT worsening*

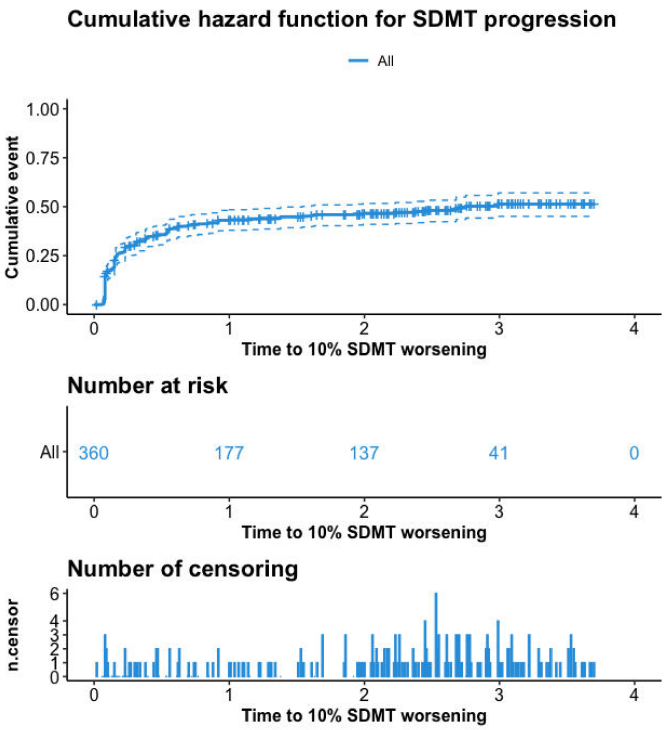


Figure legend: The figure shows the cumulative hazard function for the 10% SDMT worsening. By the end of the study, the 51% of subjects reported a SDMT progression. The number of censored and the number of subjects at risk of developing the SDMT progression at each timepoint are also reported