Brain morphology predicts long-term placebo response in multiple sclerosis patients

Mariya V. Cherkasova[1*], Jessie Fanglu Fu[3], Michael Jarrett[5], Poljanka Johnson[1], Shawna Abel[1], Roger Tam[2,6], Alexander Rauscher[4], Vesna Sossi[3], Shannon Kolind[1], David Li[1,2], A. Dessa Sadovnick[1,7], Lindsay Machan[2], J. Marc Girard[8], Francois Emond[9], Reza Vosoughi[10], Anthony Traboulsee[1], A. Jon Stoessl[1*]

[1] Department of Medicine (Division of Neurology) and Djavad Mowafaghian Centre for Brain Health, University of British Columbia, Vancouver, British Columbia, Canada
[2] Department of Radiology, University of British Columbia, Vancouver, British Columbia, Canada
[3] Department of Physics and Astronomy, University of British Columbia, Vancouver, British Columbia, Canada
[4]Depatment of Pediatrics (Division of Neurology), University of British Columbia, Vancouver, British Columbia, Canada
[5]Population Data BC, University of British Columbia, Vancouver, British Columbia, Canada
[6]School of Biomedical Engineering, University of British Columbia, Vancouver, British Columbia, Canada
[7]Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia, Canada
[8]Centre Hospitalier de l'Université de Montréal, Montréal, Québec, Canada
[9]CHU de Québec-Université Laval, Hôpital de l'Enfant-Jésus, Québec, Canada
[10]Department of Internal Medicine (Neurology), University of Manitoba, Winnipeg, Canada

*Correspondence:
A. Jon Stoessl, CM, MD, FRCPC, FCAHS Professor & Head, Neurology;
Director, Djavad Mowafaghian Centre for Brain Health;
Canada Research Chair in Parkinson's;
Pacific Parkinson's Research Centre & Parkinson's Foundation Centre of Excellence;
University of British Columbia;
T: 604-822-7967
jstoessl@mail.ubc.ca

Mariya V. Cherkasova
University of British Columbia
2215 Westbrook Mall,
Vancouver, BC V6T 2B5
phone: 604-822-3428
cherkaso@mail.ubc.ca

Conflict of Interest: The authors have declared that no conflict of interest exists.

Keywords: placebo effect, liberation procedure, multiple sclerosis, cortical thickness, graph theory, structural connectivity

Manuscript word count (including figure legends, tables and references): 10,740

ABSTRACT

Significant insights have been gained into the neural mechanisms of acute placebo responses. However, little is known about the mechanisms of long-term placebo responses, such as those seen in clinical trials, or the interactions between these mechanisms and brain disease. We examined morphological brain correlates of placebo responses in a randomized clinical trial of the “liberation procedure”, a controversial endovascular treatment for multiple sclerosis (MS). Patients were randomized to receive either balloon or sham extracranial venoplasty and followed for 48 weeks. The trial did not support therapeutic efficacy of venoplasty, but a subset of both venoplasty and sham patients reported a transient improvement in health-related quality of life, indicating a placebo response. Graph theoretical analysis of cortical thickness (CT) covariance showed that placebo non-responders had a more homogenous regional CT pattern with denser clustering and a more small-world-like architecture. In placebo non-responders, lesion load inversely predicted CT in primary somatosensory and motor areas, association areas, precuneus and insula, primarily in the right hemisphere. In placebo responders, lesion load was unrelated to CT, but lesion activity was higher. Lesions associated with cortical atrophy in MS may result in a cortical configuration that cannot support a sustained placebo response.

INTRODUCTION

A placebo response is a therapeutic response to an inert intervention believed to be potent. Placebo responses are commonly observed in randomized controlled trials (RCTs), sometimes for as long as two years (1). The neurobiological mechanisms underlying placebo responses have been primarily studied in the context of acute interventions usually performed in laboratory settings. The findings have pointed to the neural circuits involved in reward expectancy and emotion regulation, as well as the 'effector pathways' specific to the type of response, such as the pain matrix in the case of placebo analgesia (2)(3)(4)(5). The neural mechanisms of long-term placebo responses, such as those observed in clinical trials, have been less studied. Several studies have observed patterns of fMRI-derived functional and structural connectivity predicting subsequent placebo responses in clinical trials of chronic pain, as well as modulation of functional connectivity patterns associated with the experienced placebo analgesia. However, the specific patterns have differed between studies (6)(7)(8)(9) (10). Others have observed modulation of metabolic brain networks involving limbic regions in response to antidepressant placebo in depressed males (11) and in response to a sham surgery in a trial of gene therapy in Parkinson's disease (12)(13).

In RCTs examining treatments of neuropsychiatric or neurodegenerative disorders, neural mechanisms producing placebo responses may interact with the neuropathological processes of the disease. For example, Alzheimer's patients show a reduced capacity for placebo analgesia in an experimental context; this impaired capacity to mount a placebo response is associated with and may result from a reduced capacity for neural coding of expectancy associated with disrupted connectivity of the prefrontal cortex with the rest of the brain (14). Such effects of brain disease on the capacity of the patient to benefit from placebo and thereby, to some degree, from real therapies, have important implications for interpreting the results of clinical trials. However, little is known about the interactions between placebo responses and brain disease.

To address this question, we examined structural neural predictors of placebo responses of multiple sclerosis (MS) participants undergoing an RCT of a controversial extracranial venoplasty procedure dubbed the "liberation therapy". The treatment was based on the hypothesis that chronic cerebrospinal venous insufficiency (CCSVI), or deficient venous drainage caused by extracranial venous stenoses, is a major contributor to the pathogenesis of MS. Despite initially promising effects of the procedure on symptoms and disease course

reported in uncontrolled, unblinded studies (15)(16)(17)(18)(19)(20), CCSVI has been found to not be specifically related to MS (21), and the procedure has now been demonstrated to be ineffective in two separate double blind sham-controlled RCTs, one of these conducted by the pioneers of the procedure (22), and the other by our group (23). In the latter, while venoplasty was not superior to sham venoplasty on any outcome measure, a subset of both venoplasty and sham treated patients experienced a significant transient improvement in self-reported health-related quality of life (MSQOL-54) over the trial's 48 week follow-up, peaking at 12 weeks following treatment and subsequently subsiding. This suggested a placebo response in both groups, with substantial individual variation (Figure 1A). We examined MRI-based structural neural predictors of this placebo response by comparing placebo responders versus non-responders in terms of brain volume, white matter lesions and cortical thickness at the onset of the trial (Figure 1).

In addition to examining global and localized measures of brain morphology, which do not take into account the global structure and connectivity of the brain, we performed a graph theoretical analysis of regional cortical thickness in placebo responders vs. non-responders. Graph theory is a modality invariant framework used to study brain connectivity by representing the brain as a complex network of "nodes" (brain regions) connected via links or "edges". In structural brain networks, the edges represent actual anatomical connections, i.e. white matter fibers, whereas in functional networks, they represent temporal covariation in neural activity. Graph theoretical analysis of regional cortical thickness has been previously used as a measure of structural connectivity in both healthy controls and MS patients (24)(25), as well as in other neurological disorders (26). However, rather than representing anatomical links, the edges represent inter-regional covariation in cortical thickness (CT), considered a proxy of functional connectivity: CT in a given region tends to covary with CT in functionally connected regions, and these covariance patterns partly recapitulate functional networks (26). Graph theory describes networks using measures of global and local connectivity. Studies using the graph theoretical approach to studying CT covariance patterns in MS have suggested increased network segregation in MS patients with an enhancement of local network properties early in the disease, progressing to a shift in both local and global network properties with advancing disease. Relative to controls, early remitting-relapsing (RRMS) patients (< 5 years) displayed stronger modular decomposition and increased local network properties with an increasing strengthening

of these properties over time (27)(28). A parallel strengthening of local properties was seen in diffusion tensor imaging (DTI) based white matter networks in a similar group of early RRMS patients (29). White matter lesion load in RRMS patients was proportional to decreases in both local and global network efficiency (25); similar decreases in global and local efficiency covarying with disease severity and lesion burden were also observed in DTI-based white matter networks (30). Finally, a study in advanced MS patients pointed to a more regular pattern of cortical thickness organization with altered local and global properties, which was correlated with MEG-based functional connectivity (24). However, it has remained unclear whether these changes in cortical architecture are related to placebo or active treatment responses in MS.

RESULTS

Placebo responders and non-responders were identified based on the area under the curve (AUC) measure of change in the physical health composite of the MSQOL-54 adjusted for baseline MSQOL-54 scores, as higher baseline scores predicted less change in on this measure ($b = -0.91$, $SE = 0.32$, $t = -2.88$, 95% CI = -1.53 - -0.28, $p = 0.005$). This was done for only those patients whose T1-weighted MRIs were of sufficient quality to be included in the volumetric and cortical thickness analyses: 88 out of the total 104 participants in the trial (for the remaining scans, signal intensity at the lateral extremes was too low for successful surface extraction). The patients whose AUC z-scores were > 0 were designated as placebo responders; those with scores ≤ 0 were designated as non-responders. This was done for both venoplasty and sham participants under the assumption that with respect to benefit, both treatments are effectively sham. Analyses of only those participants who received the sham intervention (responders: n=25, non-responders: n=21) yielded qualitatively similar trends (see Supplementary Table S4, Figure S1, S2). The responder and non-responder groups did not differ on any demographic or clinical characteristics, although there were non-significant trends for the non-responders to have higher Extended Disability Status Scale (EDSS) scores and more difficulty with ambulation at baseline (Table 1). The proportion of placebo responders vs. non-responders did not differ as a function of site: this was a multisite study, with participants from 4 different sites across Canada (see methods). As expected, the two groups had markedly different treatment responses in terms of self-reported health-related quality of life (Figure 1A).

Placebo Responders Have Higher Lesion Activity

Relative to the non-responders, placebo responders had significantly larger normalized brain volumes at baseline, prior to the intervention, without adjustment for covariates (responders: 0.77 ± 0.03; non-responders: 0.75 ±0.03; b=0.01, CI: 0.00 – 0.03, p=0.05). However, this was no longer true after age, gender, age x gender interaction, disease duration and site were included in the model (b=0, CI: -0.01 – 0.02, p=0.39). Percent brain volume change over the 48 weeks of trial also did not differ between the groups (p=0.34), though there was progressive atrophy in both groups (b = -0.31, SE = 0.09, t = -3.45, p = 0.0009). No significant differences between the groups in regional grey or white matter volume were revealed by voxel-based morphometry, and there were no differences in regional CT, adjusting for total brain volume, age, gender, the age x gender interaction, disease duration, site and lesion load. There was, however, a significant dimensional association between the magnitude of placebo response and cortical thickness of a left precuneus region (x = -3.16, y = -70.12, z = 37.10; Figure 2 A), controlling for the same covariates, as well as health-related quality of life at baseline.

There were no significant differences in lesion load, computed as the total volume occupied by lesions on fluid-attenuated inversion recovery (FLAIR) scans, either at baseline or at 24-week and 48-week follow-up points ($p_s \geq 0.78$), and there was no significant increase in lesion load over the 48 weeks (p = 0.11). There were also no significant differences between placebo responders and non-responders in terms of FLAIR lesion locations at baseline, based on a voxel-wise t-statistic comparing the two groups' lesion probability maps, thresholded using non-parametric permutation inference (31) (Figure 2B; see Supplement for a 3D movie showing the probability of each voxel being lesional as a function of group). Although the majority of participants across both groups did not have gadolinium enhanced lesions on their T1 gadolinium scans, there were significant differences in lesion activity between the groups (Table 2). Placebo responders were more likely to display gadolinium enhanced lesions at baseline (b = 1.26, SE = 0.46, z = 2.73, p = 0.006) and to have newly enhancing lesions (that were not enhanced on the previous scan) at 24 weeks (b = 1.26, SE = 0.42, z = 3.01, p = 0.003) and at 48 weeks (b = 1.77, SE = 0.51, z = 3.43, p = 0.0006). A similar pattern was evident for new lesions on FLAIR scans both at 24 (b = 1.11, SE = 0.33, z = 3.34, p = 0.0008) and at 48 weeks (b = 1.57, SE = 0.40, z = 3.97, p = 0.00007) and for new active lesions that were not present on previous gadolinium T1 scans, both at 24 (b = 1.24, SE = 0.27, z = 4.52, p=0.000006) and at 48 weeks (b = 1.54, SD =

0.36, z=4.25, p = 0.00002). The incidence of enlarging lesions was also higher in the placebo responders at 24 weeks (b = 2.61, SE = 0.75, z = 3.49, p = 0.0005), though not at 48 (p = 0.14) weeks.

Placebo Responders Have a More Homogenous and Small-World-Like CT Covariance Pattern

To examine the cortical organization in the two groups we conducted a graph theoretical analysis on CT of 78 cortical areas derived using automated anatomical labeling (AAL). We used the following common measures of network segregation and integration (32). 1) A network's clustering coefficient, which describes a fraction of a node's neighbors that are also neighbours of each other, is a measure of clustered connectivity around individual nodes. In the context of CT networks, it reflects uniformity of CT with respect to individual nodes. 2) Modularity, a measure of how clearly a network may be subdivided into segregated non-overlapping communities of highly correlated groups of regions. 3) Characteristic path length is a measure of network integration representing the number of edges typically required to connect pairs of nodes in the network. In the context of CT networks, path length does not represent the number of anatomical links required to connect two nodes, but rather the number of required indirect correlations surpassing the sparsity threshold (see below). 4) Global efficiency is a related measure computed as the average inverse shortest path length in the network. 5) Small-worldness, computed as the ratio of clustering coefficient and characteristic path length, describes a topology featuring numerous short-range connections with an admixture of few long-range connections, which balances specialized and distributed processing while minimizing wiring costs. Like other anatomical networks, cortical thickness covariance networks have been reported to have small world topologies (26).

Mean regional CT values for each participant, adjusted for total brain volume, age, gender, age by gender interaction, disease duration, site and lesion load were used to construct pairwise correlation matrices for the placebo responder group and the non-responder group (Figure 1F, 5A). In addition, scans from gender and age-matched healthy participants (13 males, 30 females, age: 52.98 ± 8.93) were used to construct a control correlation matrix. Six of these scans were obtained at the University of British Columbia site, and the rest were taken from Open Access Series of Imaging Studies (OASIS-3). This was done to provide a benchmark for comparison of placebo responders and non-responders and help determine which CT covariance

pattern was more normative. Graph theory metrics were computed and compared between the groups using a range of linearly spaced sparsity thresholds from 10% to 50%. Error estimates from leave-one-out cross validations were used to examine the robustness of graph theory metrics and visualize group differences in the estimated metrics, and statistical significance of pairwise group differences in these metrics was evaluated against the null distributions of group differences on the metrics based on 1000 random permutations of participants in the sample. Group differences in the metrics presented below were consistent over the range of sparsity thresholds. As input matrices are not stable at low sparsity, which results in high error estimates, we chose a relatively high sparsity threshold of 43% yielding stable input matrices for presenting the results of random permutation tests. P-values for the entire range of sparsity thresholds are given in the Supplement.

Cortical thickness of placebo non-responders was more regionally homogeneous (range r: .1 to .95, median r: .66) relative to that of the placebo responders (range r_s: -.15 to .95, median r: .49) and controls (range r_s: -.22 to .92, median r: .41, Figure 3A, B). Hence, the CT network of the non-responders had a more homogenous degree distribution relative to both responders and controls. While placebo responders did not differ significantly from controls on any of the computed graph theory metrics ($p_s \geq 0.29$), placebo non-responders differed significantly from both responders and controls on several metrics. Relative to placebo responders and controls, non-responders had a more segregated network topology with higher mean and normalized clustering coefficients (non-responders vs. responders: $p_s = 0.02$; non-responders vs. controls: $p_s = 0.01$). Non-responders also had marginally shorter normalized and non-normalized characteristic pathlengths relative to the responders ($p_s = 0.06$) and significantly shorter pathlengths relative to the controls ($p_s < 0.04$). A related metric of global efficiency was also increased in the non-responders relative to the responders ($p = 0.05$) and the controls ($p = 0.03$). Overall, this resulted in stronger small-world attributes for the non-responders relative to the responders and controls ($p_s = 0.01$). Given that all groups had CT networks with a small world topology, a higher small world index in placebo non-responders indicated a shift towards a more regular and less random topology relative to the responders and controls.

Lesion Load Inversely Predicts CT Only in Placebo Non-Responders

One interpretation of more regular CT graphs in neurological populations has been that they arise from coordinated tissue loss leading to more correlated CT patterns. As white matter lesions in MS may result in coordinated tissue loss in anatomically connected regions and have been associated with impaired network efficiency (25), we examined the associations between white matter lesion load and cortical thickness in the placebo responders and non-responders. Although lesion load and location did not differ significantly between the two groups, there was a significant difference between the groups in the association between cortical thickness and lesion load, controlling for total brain volume, age, disease duration and site. While there was no relationship between lesion load and cortical thickness in the placebo responders, in the non-responders, greater lesion load was associated with cortical thinning in 9 clusters identified using random field theory to correct for cluster significance (Figure 4, Table 3). The clusters covered a substantial portion of the right hemisphere including portions of the primary motor and sensory cortices as well as frontal, somatosensory, visual and auditory association areas. The major clusters included: the primary motor cortex, comprising both the paracentral lobule and the precentral gyrus and extending into the premotor cortex (middle frontal gyrus, Brodmann area (BA 6) and the insula; primary somatosensory cortex (postcentral gyrus, BA 3) extending to the superior parietal lobule and the precuneus; superior occipital gyrus (BA 19) extending to middle temporal gyrus (BA 39); middle temporal gyrus (BA 21) extending to inferior temporal gyrus (BA 20). There were also smaller primary motor and superior occipital clusters in the left hemisphere. Thus, lesions may have played a differential role in the loss cortical tissue resultant cortical organization in the two groups.

DISCUSSION

Our findings revealed significant structural brain differences between the MS patients who experienced a placebo response to an ineffective invasive intervention and those who did not. These findings advance our understanding of both the neurobiology of placebo responses and of how the propensity to experience them is impacted by the neuropathological changes in MS.

While placebo responders and non-responders did not differ significantly in terms of clinical characteristics, lesion load or lesion location, brain volume or regional cortical thickness,

there were significant differences between the groups in terms of their CT covariance patterns. Relative to the placebo responders, the non-responders had more uniform CT across different brain regions. In graph theoretical terms, the non-responders had a more segregated and clustered topology. Coupled with marginally shorter characteristic pathlengths (also reflected in an increased global efficiency), this resulted in stronger small world attributes, which indicated a shift towards a more regular and less random topology in the placebo non-responders relative to the responders and the controls. While a more regular network may be associated with smaller wiring costs, which could be imposed by axonal loss in MS, this type of network may be less capable of distributed processing and functional integration. Differences in network topology between the non-responders and controls, in the absence of differences between the responders and controls, suggest that the more segregated and regular topology of the non-responders is anomalous. Previous studies have suggested that CT covariance networks in MS are characterized by increased segregation with an enhancement of local properties, as well as a shift towards more regular networks in advanced disease (27)(28)(25)(24). Although CT networks do not reflect true structural connectivity, convergent findings have emerged from studies of DTI-based structural networks in MS patients (29) (30) and MEG-based functional connectivity (24). Together, the evidence points to a shift towards more segregated and regular networks in MS, and our findings suggest that the MS patients who experience these shifts may also lose their capacity to experience placebo effects. It follows that placebo responses may require a cortical organization whose topology favours distributed processing and functional integration.

One interpretation of our findings is that in the placebo non-responders, the disease process may have resulted in a more synchronized cortical tissue loss across different brain regions leading to increasingly correlated cortical thickness values. This pattern of neurodegeneration might have resulted in the loss of capacity to mount a placebo response. This interpretation is supported by our post-hoc analysis relating lesion load to regional cortical thickness. In the placebo responders, there was no association between lesion load and cortical thickness, whereas in the placebo non-responders, lesion load significantly predicted cortical thinning in a substantial portion of the right hemisphere, including the primary sensory and motor areas and frontal, somatosensory, visual and auditory association areas. Association areas, as well as the precuneus and the insula have been identified as hubs by graph theoretical studies of structural connectivity (33). Synchronous loss of neurons in these regions and of projections

between them could impair the associative processes that may play an important role in placebo response, such as the ability to integrate one's appraisals of one's somatic functioning in different modalities with expectancy of therapeutic benefit. The precuneus and the insula, which are key structures involved in self-referential thinking and interoceptive awareness, may play a central role in generating such expectancy-biased appraisals and have also been identified as likely functional connectivity hubs (33), but see (34). Notably, placebo responders and non-responders did not differ in the degree of cortical thinning in these regions, although left precuneus thickness predicted placebo response dimensionally. Rather, they differed in the association between lesion load and CT, potentially pointing to the role of these regions' connectivity in addition to their integrity. Insults to these structural hubs and their connections are likely to result in changes in network organization with major implications for functional integration of neural activity (35)(36)(37).

Our data do not permit us to determine definitively the cause of the differential associations between cortical thickness and lesion load in placebo responders versus non-responders. However, it is likely that lesion characteristics play a role. Cortical grey matter loss in MS may arise from a combination of primary pathological processes and secondary effects of white matter damage. With regard to the latter, chronic inactive lesions are more likely to be associated with axonal degeneration (38). Although both placebo responders and non-responders had relatively advanced disease (10-20 years, median EDSS of 4.0), at which point most lesions are inactive (39), and gadolinium enhancing lesions were observed in only 22 out of the 88 patients over the 48 weeks of the trial, placebo responders had a significantly higher incidence of these active lesions. They were also more likely to have lesions that became enlarged at 24 or at 48 weeks relative to baseline (in the same subset of participants who displayed enhancing lesions), which may indicate either expanding inflammatory activity or "smoldering" or slowly expanding lesions (39), which have some inflammatory activity but can only be definitively distinguished from inactive lesions using MRI with gradient echo sequences to detect a rim with iron-laden activated microglia and macrophages surrounding an inactive centre (40)(41). Based on this, and given equivalent lesion load in the two groups, it is plausible that placebo non-responders may have had a higher proportion of inactive lesions. These may be more likely to be chronic lesions with axonal loss, which could in turn drive synchronized loss cortical tissue, although our data do not speak to that. In addition, the absence of evidence of active

inflammation in non-enhanced lesions may have caused their volume to be reduced compared to that of active lesions, resulting in a smaller lesion load. A supplementary analysis showed that our finding of unique associations between lesion load and cortical thickness in the placebo non-responders remained after removing all participants with enhancing lesions at baseline (Supplementary figure S3). Nonetheless, lesion activity could have partially contributed to our findings.

Associations of cortical thinning with lesion load and other measures of white matter integrity in MS have been previously reported and were in some cases modulated by lesion location and disease characteristics (42)(43)(44)(45)(46). For example, the volume of periventricular and not of other white matter lesions were found to predict cortical thickness in RRMS (45), consistent with the role of CSF-mediated immune pathology in cortical atrophy. We explored the possibility that the differential associations between lesion load and CT could be driven by lesion location but did not find significant differences in lesion location between placebo responders and non-responders. The relationship between white matter pathology and CT has also been found to depend on disease subtype: white matter pathology predicted cortical thinning more in RRMS than in SPMS, which was interpreted to suggest that cortical grey matter atrophy occurs largely independent of white matter pathology in SPMS (42)(43). Roughly half of our cohort were RRMS patients, with SPMS and PPMS patients making up the rest. Thus, restricting the analysis to a single patient subtype would severely limit our power and could lead to unjustified conclusions. However, as can be seen in Supplementary Figure S4, the association does not appear to be driven by the patients with a specific MS subtype.

In sum, given the associations we observed between lesion load and cortical thinning in areas of high spatial overlap with network hubs, it is plausible that lesions driving synchronized cortical tissue loss in these regions resulted in the CT network reorganization we observed, although the mechanisms whereby this occurred remain unclear.

It remains unclear whether network topology characteristics associated with the absence of placebo response revealed by our study would generalize to other patient groups or other types of placebo response. This could be tested in future studies investigating structural or functional neural correlates of placebo responses; graph theoretical analysis can could also be applied to existing published datasets. Thus far, graph theoretical network analysis has seldom been used to study neural correlates of placebo responses. We are aware of only one study using graph theory

metrics of DTI-based structural networks to predict placebo response in migraine patients (10). In that study, increased global and local efficiency at baseline inversely predicted placebo analgesic response to sham acupuncture, which is consistent with our findings.

Our study had several limitations. First, to maximize sample size, participants receiving sham and active venoplasty interventions were combined under the reasonable assumption that both interventions were effectively sham, considering that venoplasty has been demonstrated to be ineffective in two separate studies. Indeed, our placebo non-responder group included non-significantly more sham than active venoplasty participants (Table 1). Moreover, our supplementary analysis focusing only on those participants who received the sham intervention yielded qualitatively similar trends (Supplementary Table S4, Figures S1 and S2). We report the findings from the full sample, as they are likely more reliable. Second, we also had to exclude some patients from our analysis due to poor MRI quality. Third, our findings are limited by our definition of placebo response. Placebo responders and non-responders were identified based on subjective self-report regarding physical health-related quality of life over the course of the trial. As such, our findings may not generalize to placebo responses manifesting as more objective measures of clinical improvement, which we did not observe in the trial (23). However, even subjective placebo responses can be quite compelling for the patients: in the case of venoplasty, they may have contributed to fueling the efforts of patient advocacy groups to legitimize the procedure in the face of skepticism from the scientific community (47). In addition, in the context of a clinical trial with a 48-week follow-up, placebo response is necessarily confounded with the natural course of the disease and could in part reflect changes in health-related quality of life due to relapses, remissions and regression to the mean. Although such contributions are possible, we consider them unlikely to be substantial for the following reasons. 1) While we did see evidence of regression to the mean, with participants with poorer health-related quality of life at baseline more likely to show an improvement over the course of the trial, we adjusted our measure of placebo response for baseline scores. 2) While we did observe a higher incidence of active lesions in the placebo responders (though only in a minority of patients), the incidence of newly enhancing lesions did not change over the course of the trial, indicating that lesion activity did not transiently diminish in the placebo responders (Figure S5). Therefore, it is unlikely that the transient increase in health-related quality of life in the responder group was driven by remissions. This is further supported by the absence of such transient improvement on more

objective clinician-rated measures, such as the EDSS. Finally, as mentioned earlier, our graph theoretical analysis was based on cortical thickness covariance patterns, which do not represent true structural connectivity. Rather, they are thought of as an indirect reflection of functional connectivity between brain regions. Resting state fMRI data could have provided valuable direct information regarding functional connectivity differences between placebo responders and non-responders. Unfortunately, no resting state sequences were collected as part of this trial.

In conclusion, our findings demonstrate that the absence of placebo response in MS is associated with 1) a more regular and segregated topology of the cortical thickness covariance network, 2) cortical tissue loss in regions spatially overlapping hubs that is associated with and may be secondary to white matter pathology, and 3) lower lesion activity. Grey matter atrophy and dysfunction of cortical networks are important predictors for both physical disability and cognitive dysfunction in MS (48)(49)(50)(51)(52)(53)(54). Although previous research has pointed to relationships between cortical thickness covariance patterns and disease characteristics, these studies have not examined the relationships between CT network characteristics and treatment response. Our data are the first to show that brain network characteristics favoring integration are associated with placebo response in MS patients. Considering that placebo response is a constituent of active therapeutic response, these network characteristics may by extension predict response to active therapies. Our findings also underscore the importance of brain networks in the generation of the placebo response and highlight graph theory network analysis as a tool that can elucidate the neural substrates of placebo responses when applied to neuroimaging data.

METHODS

Detailed description of the trial's methods is given in (23).

Participants

Participants with relapsing remitting (RRMS), secondary (SPMS) and primary progressive (PPMS) MS were recruited between May 29, 2013 and Aug 19, 2015 from four Canadian academic centers: 1) University of British Columbia Hospital, Vancouver; 2) Health Sciences Centre, Winnipeg; 3) CHUM, Hôpital Notre-Dame, Montreal; 4) CHU de Québec-Université Laval, Hôpital de l'Enfant-Jésus, Québec. Inclusion criteria were: age 18-65 years, diagnosis of definite MS by the 2010 McDonald criteria (55), an Expanded Disability Status

Score (EDSS) (56) between 0 (i.e. minimal disability) and 6.5 (i.e. using bilateral aids to walk), neurologically stable disease within the 30 days prior to screening, and fulfillment of at least two ultrasound criteria for CCSVI defined by the Canadian Institute for Health Research Imaging Expert Panel with confirmation of >50% narrowing affecting at least one of 3 extracranial veins (either internal jugular and/or azygos) on catheter venography. Participants on standard disease-modifying therapies were permitted to continue on the medication, and changes were allowed for on study relapses after randomization. Exclusion criteria included treatment with vasodilators, parasympathomimetics, sympathicolytics, calcium channel blockers, previous venoplasty and/or stenting, previous jugular or subclavian central line or major neck surgery or radiation, previous contrast allergy, inability to undergo MRI, inadequate medical records confirming diagnosis and disease course, and inability to complete all study visits. The total of 104 participants meeting the above criteria were randomized and entered into the trial, but only 88 of them had T1-weighted MRIs of sufficient quality for volumetric and cortical thickness analyses.

Study approval

The clinical research ethics boards at the four participating centers approved the study protocol, and participants gave written informed consent.

Procedure

Eligible participants were randomized 1:1 to either sham or active balloon venoplasty of all narrowed veins under study. Stratified randomization (RRMS versus progressive MS course) at each site was completed by a permuted-block size of six. Venography was performed under conscious sedation and the duration of time within the angiography suite was uniform for both venoplasty and sham treated participants. A 5-French diagnostic catheter was introduced through the common femoral vein to selectively catheterize the right and left internal jugular veins as well as the azygos vein. The venoplasty participants were treated with an angioplasty balloon 2mm greater than the nominal vein diameter which was inflated for 60 seconds. The participants randomized to sham had a catheter that was advanced across the stenosis and left for 60 seconds.

After randomization and intervention, participants were followed for 48 weeks, with Adverse Event assessments, standardized Patient-Reported Outcome scales including MSQOL-54 (72 hours and weeks 2, 12, 24, 36 and 48), Multiple Sclerosis Functional Composite score

(MSFC) (57), EDSS (weeks 2, 12, 24 and 48), MRI and ultrasound (weeks 24 and 48) assessments.

MRI data acquisition

The MS patient MRIs were carried out on the following scanners: Site 1: Philips Intera 3T; Site 2: Siemens Verio 3T; Site 3: Philips Achieva 3T; Site 4: Philips Achieva 3T. The scanning protocol was standardized across the four sites. All scans were acquired in sagittal orientation, and the parameters for each site were as follows. Sites 1 and 3: resolution = 320 x 320 x 200 and voxel size = 0.8 mm for all scans; 3D T1 time repetition (TR) = 6.2 ms and time echo (TE) = 3.0 ms; 3D T2 TR = 2500 ms and TE = 363 ms; 3D FLAIR TR = 8000 ms, TE = 337 ms and inversion time (TI) = 2400 ms. Site 2: 3D T1 resolution =320 x 320 x 176, voxel size = 0.78 x 0.78 x 0.78 mm, TE =3.4 ms, TR=1900 ms, TI = 900; 3D T2 resolution = 512 x 512 x 160, voxel size = 0.488 x 0.488 x 1 mm, TE = 409 ms, TR = 3200 ms; FLAIR: resolution = 320 x 320 x 160, voxel size = 0.83 x 0.83 x 0.83 mm, TE = 337 ms, TR = 8000 ms, TI = 2400. Site 4: 3D T1 resolution = 320 x 320 x 112, slice size = 0.8 x 0.8 x 1.6 mm, TE = 3.2 ms, TR = 6.5 ms; 3D T2 resolution = 320 x 320 x 112, slice size = 0.8 x 0.8 x 1.6 mm, TE = 255, TR = 2500; FLAIR resolution = 336 x 336 x 112, slice size = 0.76 x 0.76 x 1.6 mm, TE = 330 ms, TR = 4800 ms, TI = 1650. For the control group, six 3D T1 scans were acquired on the Philips 3T Achieva scanner at Site 1 in sagittal orientation using the following parameters: resolution = 256 x 250 x 165, voxel size = 1.0 x 1.0 x 1.0, TR = 8.1 ms, TE = 3.5 ms, SENSE factor 1.8. The rest of the control 3D T1 scans were obtained from the Open Access Series of Imaging Studies repository (OASIS-3, www.oasis-brains.org, RRID:SCR_007385). These scans were acquired on Siemens TIM Trio 3T MRI scanner using the following parameters: orientation = transverse, resolution = 256 x 256 x 176, slice size = 1 mm, TR = 2400 ms, TE = 3.16 ms.

Quantification and Statistical Analysis

Placebo response

Change in MSQOL-54 Physical Health Composite was used as the measure of placebo response, as this measure showed a transient significant improvement in the trial (23). Scores at baseline and the subsequent 6 assessment points were used to compute the area under the curve scores for each participant using the Matlab function *trapz*. These values were then adjusted for

the baseline score using linear regression, and standardized residuals were then used to subdivide participants into placebo responders (those with scores > 0) and placebo non-responders (those with scores ≤ 0).

Lesion analysis

For lesion segmentation, a trained radiologist blinded to treatment assignment placed one or more seeds points on each lesion. The FLAIR images were thresholded to produce binary masks containing only hyperintense regions representing candidate lesions. These regions were subsequently processed using the manual seed points and connected component analysis to remove false positives, leaving only true lesions (58). Lesion load was calculated as the sum of volumes of all T2-FLAIR lesion masks (mm^3) for a given patient. Lesion activity analysis was carried out by the same radiologist; gadolinium enhanced lesions indicating blood brain barrier disruption were identified on T1 scans, newly enhancing lesions were counted as the lesions that were enhanced on the current scan but were not enhancing in previous scans during the trial; new and newly enlarging lesions on the T2-FLAIR scans were counted as lesions that were respectively absent (new) or stable and smaller in volume (newly enlarging) in previous assessment points in the trial; unique newly active lesions were newly enhancing lesions and non-enhancing new or newly enlarging lesions on a current scan without double counting. Lesion load and its change over the trial's duration was compared between placebo responders and non-responders with a linear mixed-effects model using the lme4 R package (RRID:SCR_015654) (59), *lmer* function, including age, gender, age by gender interaction, disease duration and site as covariates. Group differences in lesion counts at each assessment point during the trial were analyzed with poisson regressions using the lme4 *glm* function and using the same covariates.

The analysis of lesion location was performed using the Functional MRI of the Brain's Software Library (FSL, http://www.fmrib.ox.ac.uk/fsl, RRID:SCR_002823). Only the lesions identified on the T2-FLAIR scans were considered for this analysis, as there were too few participants with gadolinium enhanced lesions on the baseline scans (4 non-responders; 6 responders). First, all T2 scans were first affinely registered to the MNI-152 template using FSL's Linear Image Registration Tool (FLIRT). The resulting transformation parameters were then applied to the lesion mask files to move them into the MNI-152 space. The correspondence of lesion masks to the locations of hyperintense lesions on the T2-FLAIR scans in the MNI-152

space was verified by visual inspection. To create and compare the lesion probability maps for the placebo responders and non-responders, the registered lesion masks from all the patients in each group were merged into a single 4D file. Each 4D file was then averaged across the patient dimension at each voxel using the *fslmaths* utility, yielding a lesion probability map per each group representing the probability of each voxel being lesional (see Supplement for lesion probability map movie). For the lesion location comparison, the two 4D files were analyzed using the *Randomise* algorithm (31), which uses non-parametric permutation inference to threshold a voxel-wise statistical map produced, in this case, by voxel-wise unpaired t-tests performed on the two groups.

MRI processing for brain morphometric and cortical thickness analyses

Normalized brain volumes were computed on the 3D T1w MRIs using a segmentation-based approach that computed the parenchymal volume, normalized by intracranial volume (60). Percentage brain volume change was computed using an automated in-house registration-based method, which estimates volumetric change from parenchymal edge displacement between scans. This method was designed to be particularly robust to scanner-induced variations in voxel size and intensity across time (61).

Prior to further volumetric analyses, white matter lesions were filled with intensities similar to those of neighboring white matter voxels using the Lesion Filling function in FSL (62). This reduces intensity contrast within lesion areas and can improve registration and segmentation of MS brains and resulting morphometric measurements (63). Native T1 weighted MRIs were processed through CIVET automated pipeline (version 2.1, in-house software developed in the lab of Alan C. Evans, Montreal Neurological Institute) (64) housed on CBRAIN (RRID:SCR_005513), a web-based platform for analyzing neuroimaging data developed by the McGill Centre for Integrative Neuroscience (MCIN) (65). The pipeline included the following steps. 1) The native MRIs were registered to standard space using a 9-parameter nonlinear transformation (66). For the controls, the MNI-152 template was used, whereas for the MS patients, a high-resolution Alzheimer's disease template was used, provided by CIVET: because of paraventricular brain atrophy, the AD template produced superior registration. Simultaneous non-uniformity correction was performed using N3 algorithms (67). 3) The registered corrected images were classified into white matter, gray matter, CSF, and background, using an artificial

neural network classifier (64). 4) Binary volumes consisting of gray matter voxels and white matter voxels were then extracted from classified images, smoothed using an 8-mm FWHM smoothing kernel and used in the subsequent voxel-based morphometry analyses (see below). 5) Cortical surfaces were extracted from the classified images using the Constrained Laplacian Anatomic Segmentation Using Proximities (CLASP) surface extraction procedure (68)(69). This process generates a triangulated mesh at the interface of gray matter and white matter and expands the mesh outwards toward the pial surface. 6) Cortical thickness was measured in native space as the distance between corresponding vertices on the inner and outer surfaces of the mesh across 40,962 vertices in each hemisphere (70). The cortical thickness maps were then blurred using a 20 mm surface based blurring kernel to improve sensitivity (71). Quality control was performed using a combination of the CIVET QC tool and visual inspection of the outputs.

Voxel-based morphometry

Voxel-based morphometry was conducted on the 8-mm smoothed gray and white matter volumes using the PET and VBM module of Statistical Parametric Mapping (SPM 12, RRID:SCR_007037). The grey and white matter volumes were compared using voxel-wise two sample t-tests with total brain volume, age, gender, interaction of age and gender, disease duration and site. Continuous associations between volumes and the degree of placebo response, represented by the AUC change in MSQOL-54 scores adjusted for baseline scores, were also explored using multiple regression models with the same covariates. The resulting statistical maps were thresholded using family-wise error corrected threshold of $p = 0.05$.

Cortical thickness analysis

Native-space CT at 40,962 vertices was analyzed statistically using SurfStat (RRID:SCR_007081) (http://www.math.mcgill.ca/keith/surfstat/), a Matlab toolbox for the statistical analysis of surface and data with linear mixed effects models and random field theory to correct for multiple comparisons in determining vertex and cluster significance (72). The analyses were performed using Matlab17a. We first considered models predicting CT as a function of group membership (placebo responder vs. non-responder), adjusting for total brain volume, age, gender, the interaction of age x gender, disease duration, site and lesion load at baseline:

$$CT = \beta 1 + \beta 2\ Volume + \beta 3\ Age + \beta 4 Gender + \beta 5 Age*Gender + \beta 6 Disease_Duration + \beta 7\ Site + \beta 8 Lesion_Load + \beta 9\ Group$$

In light of the differential associations of lesion load with cortical thickness in placebo non-responders and responders, which may be viewed as violating the assumption of homogeneity of regression slopes (73), the analysis was performed using models with and without lesion load as a covariate, as well as those substituting baseline EDSS as a measure of disease burden. As this did not qualitatively change the results, we report the findings from the models with lesion load as a covariate. We additionally considered whether placebo response as a continuous variable was associated with CT. We analyzed this with a model predicting vertex-wise CT from the AUC measure of change in MSQOL-54, adjusting for total brain volume, age, disease duration, site and baseline MSQOL-54 scores:

$$CT = \beta 1 + \beta 2\ Volume + \beta 3\ Age + \beta 4\ Gender + \beta 5\ Age*Gender + \beta 6 Disease_Duration + \beta 7\ Site + \beta 8 Lesion_Load + \beta 9\ Baseline + \beta 10\ Response$$

As these analyses yielded no significant vertices or clusters, the results are not shown.

To analyze Associations between cortical thickness and lesion load in placebo responders versus non-responders, we first used a model testing the significance of the interaction between lesion load and group in predicting vertex-wise CT, adjusting for total brain volume, age, disease duration and site.

$$CT = \beta 1 + \beta 2\ Volume + \beta 3\ Age + \beta 4\ Gender + \beta 5\ Age*Gender + \beta 6\ Disease_Duration + \beta 7 Site + \beta 8\ Responder*Lesion_Load$$

Once the significance of this interaction was established, we analyzed the associations between CT and lesion load separately in each group.

$$CT = \beta 1 + \beta 2\ Volume + \beta 3\ Age + \beta 4\ Gender + \beta 5\ Age*Gender + \beta 6\ Disease_Duration + \beta 7 Site + \beta 8\ Lesion_Load$$

Once we established that CT was significantly predicted by lesion load only in the placebo non-responders, we used a mask of the clusters where this association was evident to extract the mean cortical thickness values across all vertices in these clusters for every participant. We then used a scatter plot (Figure 3B) to visualize the associations between lesion load and mean CT of these regions in each group.

Graph theoretical cortical thickness analysis

Cortical surfaces were parcellated into 78 regions based on AAL (74), and mean CT values for each participant were extracted for each of the 78 regions.

A linear regression analysis was then performed to adjust these CT values for total brain volume, age, gender, interaction of age and gender, disease duration, site, total brain volume and lesion load. Again, given differential associations of lesion load with cortical thickness in placebo non-responders and responders, the graph theoretical analyses were performed on the data both with and without lesion load covaried out, as well as data adjusting for baseline EDSS in place of lesion load as a measure of disease burden. As this did not qualitatively change the results, we report the results adjusted for lesion load. In addition, given the significant group difference in normalized brain volume in placebo responders vs. non-responders, we also performed the analysis with and without adjustment for total brain volume. As these analyses yielded similar results, we report the analysis controlling for both brain volume and lesion load. For controls, CT values were adjusted for total brain volume, age, gender, and interaction of age and gender. The resulting residuals were then substituted for the raw cortical thickness values to construct an inter-regional correlation matrices for placebo responders, non-responders and controls: Rij (i, j =1, 2 ... n, where n is the number of regions). These correlation matrices were then used to construct binarized structural networks and compute graph metrics at 30 linearly spaced sparsity thresholds ranging from 0.1 to 0.5. This thresholding approach normalizes each group-level graph to have the same number of edges. The following graph metrics were computed using the Brain Connectivity Toolbox (RRID:SCR_004841) (32) in Matlab2017a. Using the formulas given in (32), we computed the following graph metrics. 1) Clustering coefficient was computed as a fraction of a node's neighbors that are also neighbours of each other:

$$C = \frac{1}{n}\sum_{i\in n} C_i = \frac{1}{n}\sum_{i\in n}\frac{2t_i}{k_i(k_i-1)}$$

, where n is the total number of nodes, C_i is the clustering coefficient of node i and ki is the degree of node i; $C_i = 0$ *for* $k_i < 2$ (75). Normalized clustering coefficient was computed as the ratio of the mean clustering coefficient C and normalization factor C_{rnad} computed as the mean clustering coefficient of 10 random networks with the same number of nodes and edges as the tested input network. 2) Modularity was computed as:

$$Q = \sum_{u \in M} \left[e_{uu} - \left(\sum_{v \in M} e_{uv} \right)^2 \right]$$

, where the network is fully subdivided into a set of non-overlapping modules M; e_{uv} is the proportion of all the edges that connect the nodes in module v (76). 3) Characteristic path length was computed as:

$$L = \frac{1}{n} \sum_{i \in N} L_i = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j \in N. j \neq i} {d_{ij}}^{-1}}{n-1}$$

, where Li is the average distance between node i and all other nodes and d_{ij} is the distance from node i to node j (75). Normalized path length was computed as the ratio of characteristic path length L and a normalization factor L_{rnad} based on 10 random networks, as described above. 4) Global efficiency of the network was computed as:

$$E = \frac{1}{n} \sum_{i \in N} E_i = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j \in N. j \neq i} {d_{ij}}^{-1}}{n-1}$$

, where E_i is the efficiency of node i (77). 5) Small-worldness was computed as the ratio of normalized clustering coefficient and normalized characteristic path length:

$$S = \frac{C / C_{rand}}{L / L_{rand}}$$

, where C and C_{rand} are clustering coefficients for the tested and random networks, respectively, and L and L_{rand} are the characteristic path lengths for these networks (78). Leave-one-out cross-validation was performed to estimate the stability of the graph metrics, and error estimates from these cross validations were used to visualize group differences (Figure 3). Statistical significance associated with these differences was evaluated against null distributions of group differences on the metrics based on 1000 random permutations of participants in the input matrix.

AUTHOR CONTRIBUTIONS

MVC: conceptualization, methodology, formal analysis, visualization, writing – original draft, writing – review and editing
JFF: formal analysis, software, visualization, writing – review and editing
MJ: data curation
PJ: data curation, investigation
SA: data curation, investigation
RT: investigation, formal analysis, resources, writing – review and editing
AR: investigation, formal analysis, resources, writing – review and editing
VS: formal analysis, writing – review and editing
SK: conceptualization, investigation, writing – review and editing
DL: investigation, writing – review and editing
ADS: funding acquisition, investigation
LM: investigation
MG: investigation
FE: investigation
RV: investigation
AT: conceptualization, funding acquisition, investigation, methodology, supervision, resources, writing – review and editing
AJS: conceptualization, methodology, supervision, resources, writing – review and editing

ACKNOWLEGMENTS

This work was primarily supported by cooperative agreements from the Canadian Institutes of Health Research (CIHR), MS Society of Canada, Michael Smith Foundation for Health Research, Research Manitoba, and Ministère de la Santé et des Services Sociaux du Québec. Healthy control data were provided in part by OASIS-3: Principal Investigators: T. Benzinger, D. Marcus, J. Morris; NIH P50AG00561, P30NS09857781, P01AG026276, P01AG003991, R01AG043434, UL1TR000448, R01EB009352. We thank Andrew Riddehough, and Kevin Atkins and Kevin Lam at the UBC MS/MRI group for computing support and data handling.

REFERENCES

1. Olanow WC et al. Gene delivery of neurturin to putamen and substantia nigra in Parkinson disease: A double-blind, randomized, controlled trial. *Ann. Neurol.* 2015;78(2):248–257.

2. Scott DJ et al. Individual Differences in Reward Responding Explain Placebo-Induced Expectations and Effects. *Neuron* 2007;55(2):325–336.

3. Lidstone SC et al. Effects of Expectation on Placebo-Induced Dopamine Release in Parkinson Disease. *Arch. Gen. Psychiatry* 2010;67(8):857.

4. Atlas LY, Wager TD. A meta-analysis of brain emchansims of placebo analgesia: consistent findings and unanswered questions. In: Benedetti F ed. *Placebo, Handbook of Experimental Psychology*. Berlin: Springer-Verlag; 2014:39–64

5. Price DD, Finniss DG, Benedetti F. A comprehensive review of the placebo effect: recent advances and current thought. *Annu. Rev. Psychol.* 2008;59(1):565–590.

6. Wanigasekera V et al. Disambiguating pharmacological mechanisms from placebo in neuropathic pain using functional neuroimaging. *Br. J. Anaesth.* 2018;120(2):299–307.

7. Tétreault P et al. Brain connectivity predicts placebo response across chronic pain clinical trials. *PLoS Biol.* 2016;14(10):1–22.

8. Hashmi JA et al. Brain networks predicting placebo analgesia in a clinical trial for chronic back pain. *Pain* 2012;153(12):2393–2402.

9. Vachon-Presseau E et al. Brain and psychological determinants of placebo pill response in chronic pain patients. *Nat. Commun.* 2018;9(1):3397:.

10. Liu J et al. Integration of white matter network is associated with interindividual differences in psychologically mediated placebo response in migraine patients. *Hum. Brain Mapp.* 2017;38(10):5250–5259.

11. Mayberg HS et al. The functional neuroanatomy of the placebo effect. *Am. J. Psychiatry* 2002;159(5):728–737.

12. Ko JH et al. Network modulation following sham surgery in Parkinson's disease. *J. Clin. Invest.* 2014;124(8):3656–3666.

13. Niethammer M, Eidelberg D. Metabolic brain networks in translational neurology: Concepts and applications. *Ann. Neurol.* 2012;72(5):635–647.

14. Benedetti F et al. Loss of expectation-related mechanisms in Alzheimer's disease makes analgesic therapies less effective. *Pain* 2006;121(1–2):133–144.

15. Dake MD, Dantzker N, Bennett WL, Cooke JP. Endovascular correction of cerebrovenous anomalies in multiple sclerosis : A retrospective review of an uncontrolled case series. *Vasc. Med.* 2012;17(3):131–137.
16. Hubbard D et al. Clinical improvement after extracranial venoplasty in multiple sclerosis. *J. Vasc. Interv. Radiol.* 2012;23(10):1302–1308.
17. Zamboni P et al. A prospective open-label study of endovascular treatment of chronic cerebrospinal venous insufficiency. *J. Vasc. Surg.* 2009;50(6):1348–1358.
18. Radak D et al. Percutaneous angioplasty of internal jugular and azygous veins in patients with chronic cerebrospinal venous insufficiency and multiple sclerosis : early and mid-term results. *Plebology* 2014;29(6):367–375.
19. Salvi F, Buccellato E, Galeotti R. Venous angioplasty in multiple sclerosis : neurological outcome at two years in a cohort of relapsing-remitting patients. *Funct. Neurol.* 2012;2007:55–60.
20. Zagaglia S, Balestrini S, Perticaroli E. Percutaneous transluminal angioplasty for chronic cerebrospinal venous insufficiency in multiple sclerosis : dichotomy between subjective and objective outcome scores. *Neurol. Sci.* 2013;34:2205–2210.
21. Traboulsee AL et al. Prevalence of extracranial venous narrowing on catheter venography in people with multiple sclerosis , their siblings , and unrelated healthy controls : a blinded , case-control study. *Lancet* 2013;6736(13):1–8.
22. Zamboni P et al. Efficacy and safety of extracranial vein angioplasty in multiple sclerosis. *JAMA Neurol 2018;* 75(1):35-43.
23. Traboulsee AL et al. Safety and efficacy of venoplasty in MS. *Neurology* 2018;91:e1660-e1668.
24. Tewarie P et al. Disruption of structural and functional networks in long-standing multiple sclerosis. *Hum. Brain Mapp.* 2014;35(12):5946–5961.
25. He Y et al. Impaired small-world efficiency in structural cortical networks in multiple sclerosis associated with white matter lesion load. *Brain* 2009;132(12):3366–3379.
26. Alexander-Bloch A, Giedd JN, Bullmore E. Imaging structural co-variance between human brain regions. *Nat. Rev. Neurosci.* 2013;14:322–336.
27. Fleischer V et al. Longitudinal cortical network reorganization in early relapsing–remitting multiple sclerosis. *Ther. Adv. Neurol. Disord.* 2019;12:1–15.

28. Muthuraman M et al. Structural brain network characteristics can differentiate CIS from early RRMS. *Front. Neurosci.* 2016;10:1–12.
29. Fleischer V et al. Increased structural white and grey matter network connectivity compensates for functional decline in early multiple sclerosis. *Mult. Scler.* 2017;23(3):432–441.
30. Shu N et al. Diffusion tensor tractography reveals disrupted topological efficiency in white matter structural networks in multiple sclerosis. *Cereb. Cortex* 2011;21(11):2565–2577.
31. Winkler AM, Ridgway GR, Webster MA, Smith SM, Nichols TE. NeuroImage Permutation inference for the general linear model. *Neuroimage* 2014;92:381–397.
32. Rubinov M, Sporns O. Complex network measures of brain connectivity: Uses and interpretations. *Neuroimage* 2010;52(3):1059–1069.
33. van den Heuvel MP, Sporns O. Network hubs in the human brain. *Trends Cogn. Sci.* 2013;17(12):683–696.
34. Warren DE et al. Network measures predict neuropsychological outcome after brain injury. *Proc. Natl. Acad. Sci.* 2014;111(39):14247–14252.
35. Gratton C, Nomura E, Pérez F, D'Esposito M. Focal brain lesions cause widespread disruption of the modular organization of the brain. *Hum. Brain Mapp.* 2012;24(6):1275–1285.
36. Crossley NA et al. The hubs of the human connectome are generally implicated in the anatomy of brain disorders. *Brain* 2014;137(8):2382–2395.
37. Pandit AS et al. Traumatic brain injury impairs small-world topology. *Neurology* 2013;80(20):1826–1833.
38. Mahad DH, Trapp BD, Lassmann H. Pathological mechanisms in progressive multiple sclerosis. *Lancet Neurol.* 2015;14(2):183–193.
39. Frischer JM et al. Clinical and pathological insights into the dynamic nature of the white matter multiple sclerosis plaque. *Ann. Neurol.* 2015;78(5):710–721.
40. Kaunzner UW et al. Quantitative susceptibility mapping identifies inflammation in a subset of chronic multiple sclerosis lesions. *Brain* 2019;142(1):133–145.
41. Dal-Bianco A et al. Slow expansion of multiple sclerosis iron rim lesions: pathology and 7 T magnetic resonance imaging. *Acta Neuropathol* 2017;3:25–42.
42. Steenwijk MD et al. Cortical atrophy patterns in multiple sclerosis are non-random and clinically relevant. *Brain* 2016;139(1):115–126.
43. Steenwijk MD et al. Unraveling the relationship between regional gray matter atrophy and

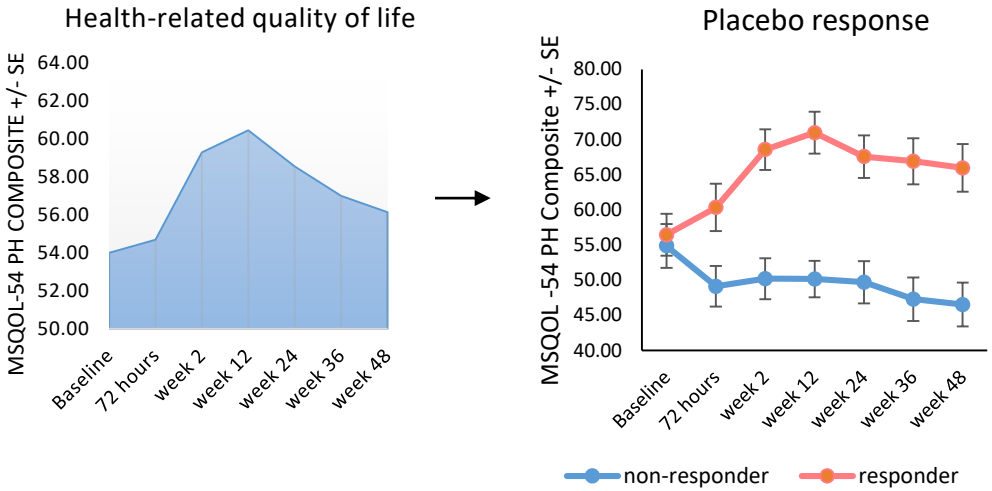
pathology in connected white matter tracts in long-standing multiple sclerosis. *Hum. Brain Mapp.* 2015;36(5):1796–1807.
44. Bergsland N et al. Corticospinal tract integrity is related to primary motor cortex thinning in relapsing-remitting multiple sclerosis. *Mult. Scler.* 2015;21(14):1771–1780.
45. Jehna M et al. Periventricular lesions correlate with cortical thinning in multiple sclerosis. *Ann. Neurol.* 2015;78(4):530–539.
46. Charil A et al. Focal cortical atrophy in multiple sclerosis: Relation to lesion load and disability. *Neuroimage* 2007;34(2):509–517.
47. Paylor B, Longstaff H, Rossi F, Illes J. Collision or convergence?: Beliefs and politics in neuroscience discovery, ethics, and intervention. *Trends Neurosci.* 2014;37(8):409–412.
48. Calabrese M et al. Widespread cortical thinning characterizes patients with MS with mild cognitive impairment. *Neurology* 2010;74(4):321–328.
49. Geurts JJG, Barkhof F. Grey matter pathology in multiple sclerosis. *Lancet Neurol.* 2008;7:841–851.
50. Calabrese M et al. Cortical atrophy is relevant in multiple sclerosis at clinical onset. *J. Neurol.* 2007;254(9):1212–1220.
51. Fisniku LK et al. Gray matter atrophy is related to long-term disability in multiple sclerosis. *Ann. Neurol.* 2008;64(3):247–254.
52. Chen JT et al. Relating neocortical pathology to disability progression in multiple sclerosis using MRI. *Neuroimage.* 2004;23(3):1168–1175.
53. Amato MP et al. Association of Neocortical Volume Changes With Cognitive Deterioration in Relapsing-Remitting Multiple Sclerosis. *Arch. Neurol.* 2007;64(8):1157–1161.
54. Sailer M et al. Focal thinning of the cerebral cortex in multiple sclerosis. *Brain* 2003;126(8):1734–1744.
55. Polman CH et al. Diagnostic Criteria for Multiple Sclerosis : 2010 Revisions to the McDonald Criteria. *Ann. Neurol.* 2011;69:292–302.
56. Kurtzke JF. Rating neurologic impairment in multiple sclerosis : An expanded disability status scale (EDSS). *Neurology* 1983;33:1444–1453.
57. Cutter GR et al. Development of a multiple sclerosis functional composite as a clinical trial outcome measure. *Brain* 1999;122:871–882.
58. McAusland J, Tam RC, Wong E, Riddehough A, Li DKB. Optimizing the Use of Radiologist

Seed Points for Improved Multiple Sclerosis Lesion Segmentation. *IEEE Trans. Med. Imaging* 2010;57(11):2689–2698.
59. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models using lme4. *J. Stat. Softw.* 2015;67(1):1–48.
60. Wicks E et al. Automatic computation of normalized brain volume on 3D T1-weighted MRI scans without registration to standard space. In: *Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM) Scientific Meeting*. 2015.
61. Smith SM et al. Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. *Neuroimage.* 2002; 17(1):479-489.
62. Battaglini M, Jenkinson M, De Stefano N. Evaluating and reducing the impact of white matter lesions on brain volume measurements. *Hum. Brain Mapp.* 2012;33(9):2062–2071.
63. Valverde S, Oliver A, Lladó X. A white matter lesion-filling approach to improve brain tissue volume measurements. *Neuroimage Clin.* 2014;6:86–92.
64. Zijdenbos AP, Forghani R, Evans AC. Automatic " Pipeline " Analysis of 3-D MRI Data for Clinical Trials : Application to Multiple Sclerosis. *IEEE Trans. Med. Imaging* 2002;21:1280–1291.
65. Sherif T et al. CBRAIN : a web-based , distributed computing platform for collaborative neuroimaging research. *Front. Integr. Neurosci.* 2014;8:1–13.
66. Collins L, Neelin P, Peters TM, Evans AC. Automatic 3D intersubject registration of MR volumetric data in standardized talairach space. *J. Comput. Assist. Tomogr.* 1994;18:192–205.
67. Sled JG, Zijdenbos AP, Evans AC. A Nonparametric method for automatic correction of intensity nonuniformity in MRI Data. *IEEE Trans. Med. Imaging* 1998;17(1):87–97.
68. Kim JS et al. Automated 3-D extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classification. *Neuroimage* 2005;27:210–221.
69. Macdonald D, Kabani N, Avis D, Evans AC. Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI. *Neuroimage* 2000;356:340–356.
70. Lerch JP, Evans AC. Cortical thickness analysis examined through power analysis and a population simulation. *Neuroimage* 2005;24:163–173.
71. Chung MK et al. Deformation-based surface morphometry applied to gray matter deformation. *Neuroimage* 2003;18:198–213.

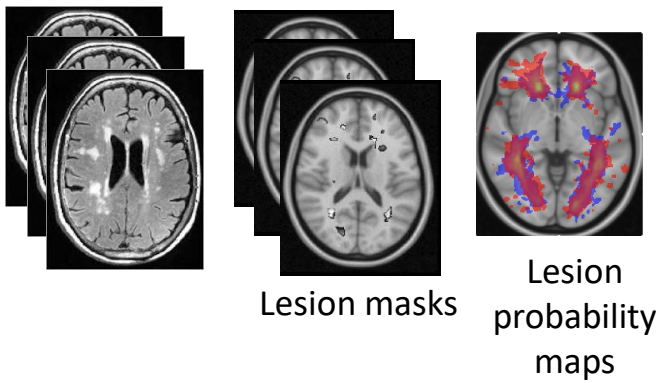
72. Worsley KJ, Andermann M, Koulis T, Macdonald D, Evans AC. Detecting changes in conisotropic images. *Hum. Brain Mapp.* 1999;101:98–101.
73. Miller GA, Chapman JP. Misunderstanding analysis of covariance. *J. Abnorm. Psychol.* 2001;110:40–48.
74. Tzourio-Mazoyer N et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 2002;15:273–289.
75. Watts DJ, Strogatz SH. Collective dynamics of 'small-world' networks. *Nature* 1998;393:440–442.
76. Newman MEJ. Modularity and community structure in networks. *Proc. Natl. Acad. Sci.* 2006;103(23):8577–8582.
77. Latora V, Marchiori M. Efficient behavior of small-world networks. *Phys. Rev. Lett.* 2008;87:198701.
78. Humphries MD, Gurney K. Network ' small-world-ness ': a quantitative method for determining canonical network equivalence. *PLoS One* 2008;3(4):e0002051.

Figure 1: Method overview

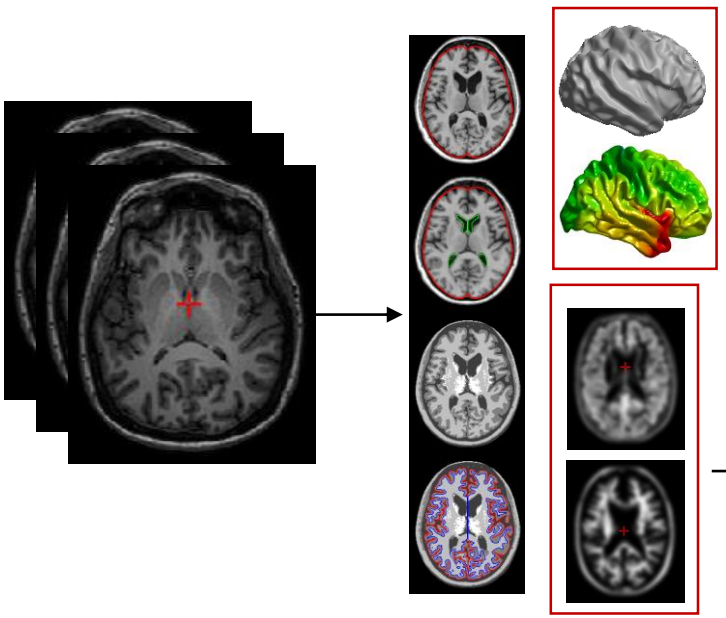
A: Placebo response classification



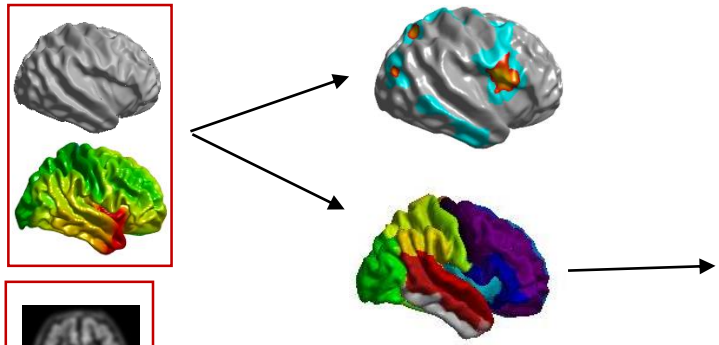
B: T2 FLAIR MRIs



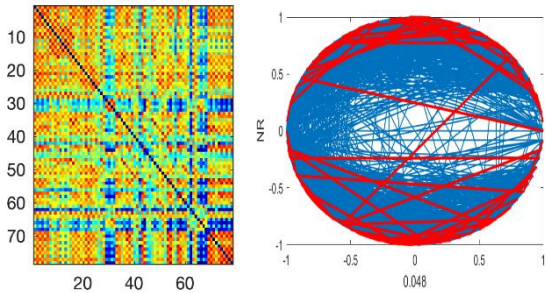
C: T1 MRIs, CIVET



D: SufStat



F: Graph Theory



E: SPM VBM

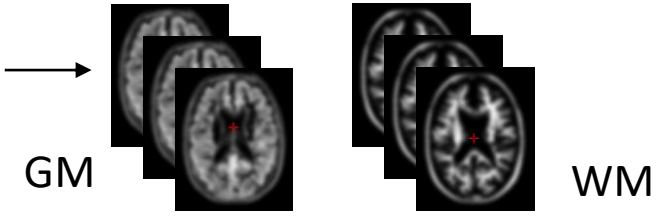
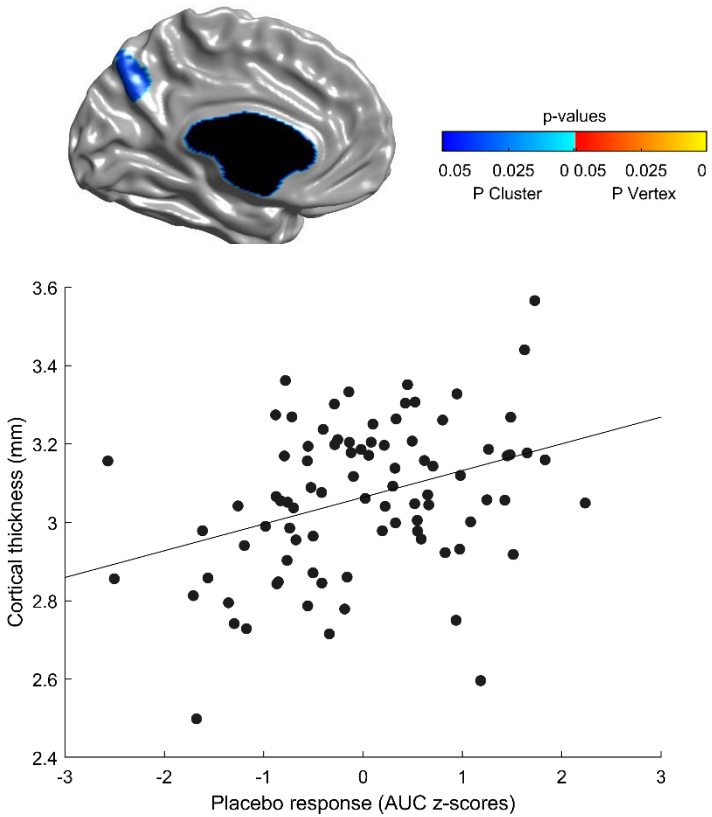
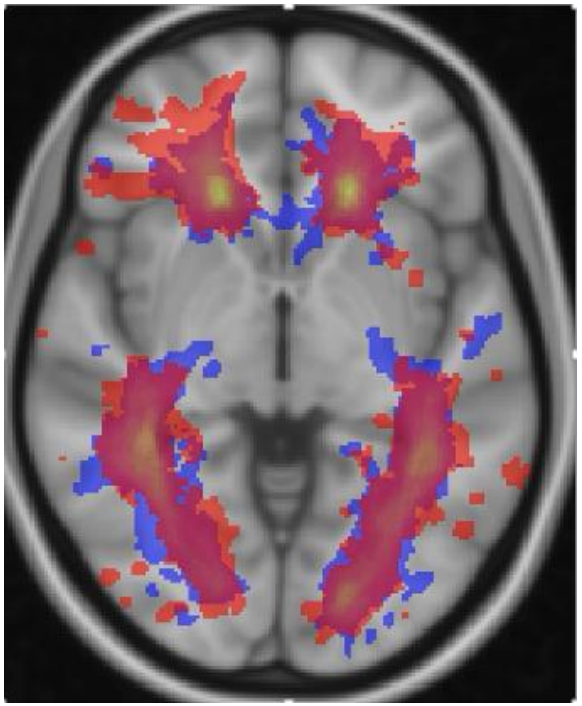


Figure 2. Brain structure of placebo responders and non-responders

A.



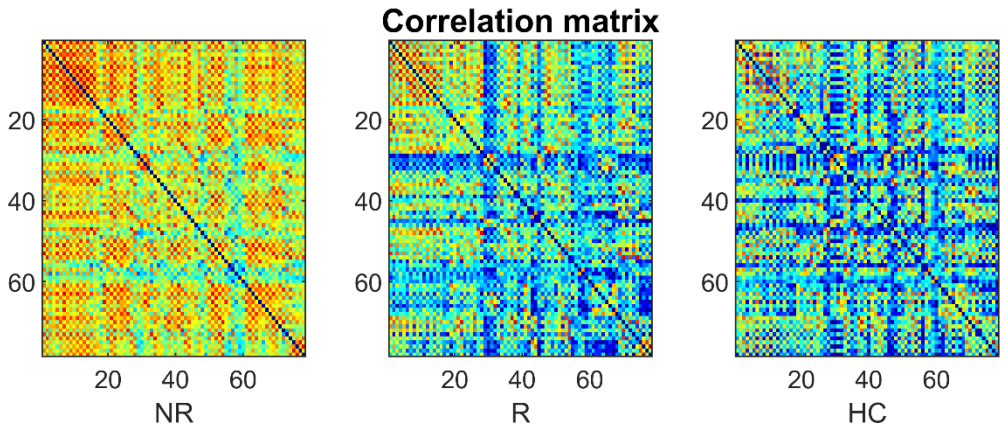
B.



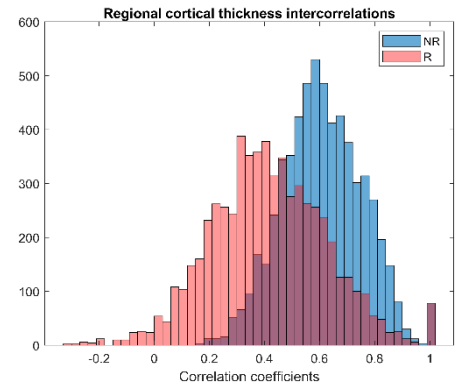
(A) Regional cortical thickness predicts a stronger placebo response. (B) Lesion probability maps in placebo responders (red) and non-responders (blue): no significant differences in lesion location between groups.

Figure 3. Structural covariance patterns of cortical thickness in placebo responders and non-responders: graph theoretical analysis

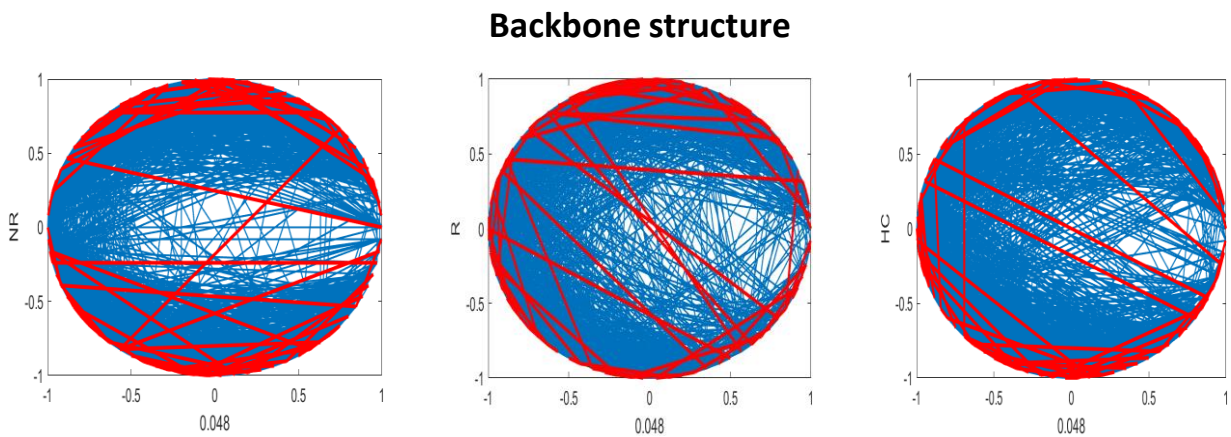
A.



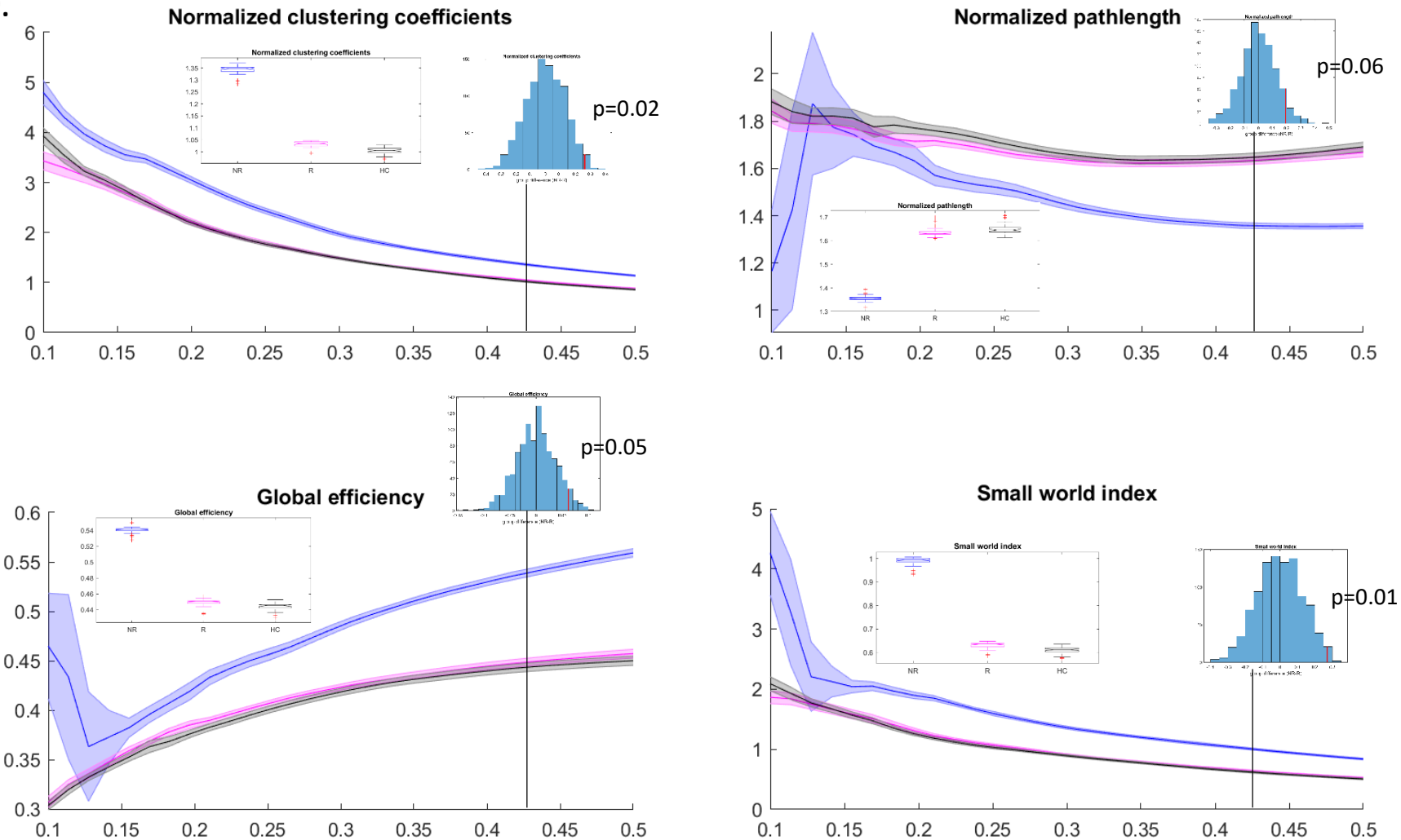
B.



C.



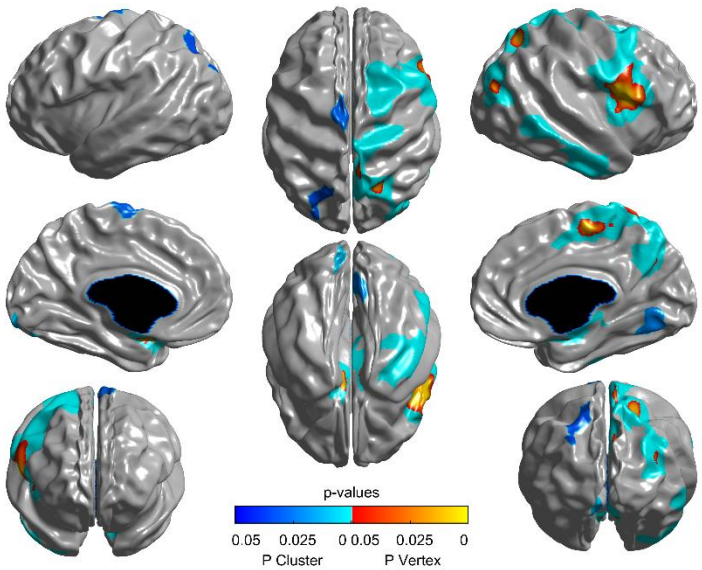
D.



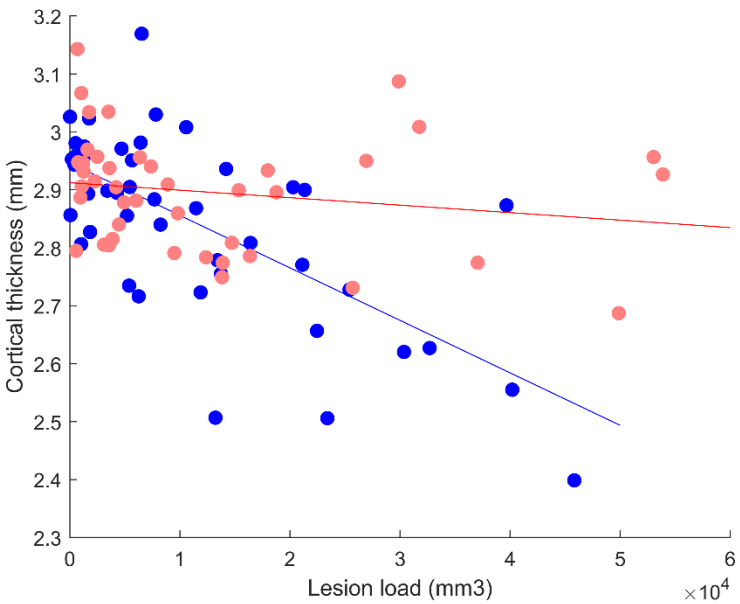
(A) Correlation matrices of cortical thickness values across 78 cortical areas delineated using Automated Anatomical Labeling (AAL) in placebo non-responders (blue), placebo responders (red), and a group of healthy age- and gender-matched controls (black). (B) A histogram depicting distributions of correlation coefficients in placebo responders and non-responders. (C) Backbone structure for these matrices at the sparsity threshold of 0.43. (D) Graph theoretical characteristics of these matrices across the range of sparsity thresholds from 0.1 to 0.5. Error ribbons represent standard deviation for parameter estimates from leave-one-out cross-validation. Box plots represent group comparisons at the sparsity threshold of 0.43 based on leave-one-out cross-validations; histograms represent the p-values based on permutation tests at the sparsity threshold of 0.43.

Figure 4. Associations between cortical thickness and T2 FLAIR lesion load as a function of placebo response

A.



B.



(A) Cortical areas whose thickness was significantly associated with lesion load in the placebo non-responders. (B) Mean cortical thickness of the regions shown in in relation to lesion load in the placebo responders (red) and non-responders (blue).

Table 1. Demographic and clinical characteristics of placebo responders and non-responders.

	Non-responder (n = 45)	**Responder (n = 34)**	**p**
Venoplasty, n	24	18	0.19
Sham, n	21	25	
Males, n	18	13	0.24
Females, n	27	30	
Age, mean (SD)	55.0 (7.4)	53.0 (8.5)	0.26
Disease duration, mean (SD)	18.8 (9.6)	16.7 (8.1)	0.29
MS type (n)			
RRMS	25	30	
PPMS	6	3	0.35
SPMS	14	10	
Baseline EDSS, median (range)	4.0 (0-6.5)	4.0 (0-6.5)	0.09†
Baseline MSFC, mean (SD)	7.2 (6.2)	8.9 (4.9)	0.17
25 foot walk, mean (SD)	15.80 (12.15)	11.87 (7.24)	0.07†
9 hole peg test, mean (SD)	108.50 (49.73)	94.41 (36.04)	0.13
PASAT, mean (SD)	37.53 (14.70)	38.63 (14.97)	0.73

RRMS = remitting relapsing MS; PPMS = primary progressive MS; SPMS = secondary progressive MS; EDSS = Expanded Disability Status Scale; MSFC = Multiple Sclerosis Function Composite; PASAT = Paced Auditory Serial Addition Test. For MSFC, raw scores rather than z-scores are given.

Table 2. Brain morphometric characteristics

Measure (SD)	**Non-responder (n = 45)**		**Responder (n = 43)**		**p**
Normalized brain parenchymal fraction (baseline)	0.75 (0.03)		0.77 (0.03)		0.39
% Change in brain volume from baseline					
24 weeks	-0.40 (0.50)		-0.43 (0.38)		0.78
48 weeks	-0.70 (0.60)		-0.61 (0.61)		0.65
WM lesion load (mm^3)					
Baseline	12031.63 (12002.73)		12269.53 (14538.12)		0.78
24 weeks	12681.97 (12185.17)		12581.44 (14527.14)		0.92
48 weeks	12792.65 (11875.89)		12907.21 (14698.09)		0.87
New T2 lesions (# of patients with lesion count)	W24	W48	W24	W48	
0	36	36	35	33	< 0.001**
1	4	5	3	4	
2	3	0	1	3	
3	0	2	0	1	
5	1	0	0	0	
6	0	0	1	0	
7	0	0	0	1	
8	0	0	2	0	

9	0		0	1		0	
11	0		0	0		1	
Unique newly active lesions (# of patients with lesion count)	BL	W24	W48	BL	W24	W48	
0	40	33	36	33	34	33	
1	1	6	5	3	2	3	
2	3	4	0	2	2	3	
3	0	0	1	2	1	1	
4	0	0	0	1	0	0	
5	0	0	1	0	0	0	< 0.0005***
6	0	1	0	0	1	1	
7	0	0	0	1	0	0	
8	0	0	0	0	1	1	
11	0	0	0	0	0	1	
15	0	0	0	0	1	0	
22	0	0	0	0	1	0	
Newly enhancing lesions (# of patients with lesion count)	BL	W24	W48	BL	W24	W48	
0	40	36	38	33	36	34	
1	1	6	4	3	4	5	
2	3	2	1	2	0	0	
3	0	0	0	2	1	2	
4	0	0	0	1	0	0	
5	0	0	0	0	0	2	< 0.01*
6	0	0	0	0	1	0	
7	0	0	0	1	0	0	
Newly enlarging T2 lesions (# of patients with lesion count)	W24		W48	W24		W48	
0	42		42	38		38	
1	2		0	2		4	W24: 0.0005***
2	0		1	1		1	W48: 0.14
7	0		0	1		0	
13	0		0	1		0	

BL = baseline; W24 = 24 weeks assessment point; W48 = 48 weeks assessment point; WM = white mater. P-values represent comparisons between placebo responders and non-responders for corresponding assessment points. The models include age, gender, age x gender interaction, disease duration and site as terms.

Table 3: Regions where cortical thickness was predicted by lesion load in placebo non-responders.

Cluster	# vertices	Peak	x	y	z	t	p
Right superior and medial frontoparietal p < 0.00005	4080	Medial frontal gyrus (R)	4.30	-24.62	57.64	5.22	0.008
		Superior parietal lobule (R)	28.69	-57.71	52.84	5.03	0.013
		Precuneus (R)	13.35	-66.54	32.83	4.96	0.015
		Middle occipital gyrus (R)	44.45	-77.33	22.46	4.83	0.021
		Postcentral gyrus (R)	11.82	-59.08	69.79	4.75	0.027
		Paracentral lobule (R)	13.69	-41.22	58.15	4.56	0.043
Right precentral gyrus / insula p < 0.00005	3026	Precentral gyrus (R)	52.78	-6.92	10.30	5.32	0.006
		Insula (R)	39.91	-0.08	16.26	5.21	0.008
Left parahippocampal gyrus p < 0.00005	383	Parahippocampal gyrus (L)	-12.13	-9.54	-12.02	5.41	0.005
Right inferior temporal gyrus p = 0.0002	1220	Inferior temporal gyrus (R)	59.56	-7.57	-28.66	4.51	0.05
Right parahippocampal gyrus p = 0.0005	335	Parahippocampal gyrus (R)	15.12	-12.79	-11.08	4.82	0.02
Left lingual gyrus p = 0.01	246	-	-16.26	-93.19	-12.51	3.14	-
Right lingual gyrus p = 0.02	311	-	10.08	-72.09	-4.17	4.03	-
Left precentral gyrus p = 0.03	385	-	-14.29	-19.43	70.91	4.12	-
Left superior parietal lobule p = 0.03	329	-	-31.33	-70.83	47.50	3.64	-

Significance thresholds for vertices and clusters of contiguous vertices and are determined using Random Field Theory; corrected p-values are given. Due to some of the clusters being extensive, the coordinates (in MNI space) are given for either peak vertices for clusters featuring such peaks (labeled in hot colors on Figure 3). For smaller clusters without peaks, average cluster coordinates are provided.