ABSTRACT

Objective: To define values of normalized brain volume (NBV) that can be categorized as low, medium or high, according to baseline characteristics of RRMS patients.

Methods: Expected-NBV (eNBV) was calculated for each patient based on age, disease duration, sex, baseline EDSS and T2-lesion volume, entering these variables into a multiple regression model run on 2342 RRMS patients (pooled FREEDOMS/FREEDOMS-II population).According to the difference between their observed NBV and their eNBV, patients were classified as having low-NBV, medium-NBV or high-NBV. We evaluated whether these NBV-categories were clinically meaningful by assessing correlation with disability worsening.

Results: The distribution of differences between observed NBV and eNBV was used to categorize patients as having low-, medium- or high-NBV. Taking the high-NBV-group as reference, the HRs for 2-year disability worsening, adjusted for treatment effect, were 1.23 (95%CI 0.92;1.63, p=0.16) for the medium-NBV and 1.75 (95%CI 1.26;2.44, p=0.001) for the low-NBV. The predictive value of NBV-groups was preserved over 4 years. Treatment effect appeared more evident in low-NBV patients (HR=0.58) than in medium-NBV (HR=0.72) and in high-NBV (HR=0.80) patients, but the difference was not significant (p=0.57).

Conclusions: RRMS patients can be categorized into disability risk-groups based on individual eNBV-values according to baseline demographics and clinical characteristics.

INTRODUCTION

Magnetic resonance imaging (MRI) derived measures of brain volume (BV) can be used to assess overall tissue damage in relapsing-remitting multiple sclerosis (RRMS).[1-4] In RRMS patients, brain volume loss (BVL) has been proven to occur throughout the disease course, including the early stages, and progresses at a rate considerably higher than that of healthy adults.[5-10] Recent work has defined relevant threshold values to discriminate pathological vs. physiological BVL rates in patients with RRMS,[11] and has demonstrated the relevance of BVL as a measure of damage,[12] with a proposal to include BVL in composite measures assessing RRMS disease activity.[13]

Many studies have shown that baseline BV and the rate of BVL correlate with disability worsening and are also predictive of future disability.[14-17] Also brain atrophy have been shown to be correlated in MS patients to cognitive impairment [18-19], employment status[20] and quality of life[21] . Whereas the correlation between low brain volumes and disability worsening in patients with RRMS has been demonstrated at a group level, it is more difficult to establish at the individual level whether the volume of the brain measured in a single patient can be considered low or high. In this post-hoc analysis, we evaluated the possibility of defining individualized values for BV normalized for head size (NBV) to classify an individual patient as having a low or a high volume of the brain according to the demographic characteristics and the disease status of the individual subject. We then evaluated this classification by testing its ability to predict -on-study disability worsening.

MATERIALS AND METHODS

Standard protocol approvals, registrations, and patient consents

The FREEDOMS and FREEDOMS II trials (ClinicalTrials.gov identifiers: NCT00289978 and NCT00355134, respectively) and their extensions were conducted in accordance with the International Conference on Harmonization Guidelines for Good Clinical Practice and the ethical principles of the Declaration of Helsinki.[18, 19] The ethics committees and institutional review boards of all participating centers approved the study protocols. All participants provided written informed consent.[20, 21]

Patients and study design

The study design and inclusion/exclusion criteria of the FREEDOMS and FREEDOMS II trials have been previously described.[20, 21] Briefly, the studies included RRMS patients (diagnosed according to the revised McDonald criteria[22]), aged 18–55 years who had a score of 0–5.5 on the Expanded Disability Status Scale (EDSS) and one or more documented relapses in the previous year or two or more relapses in the previous 2 years. The eligible patients were randomized (1:1:1) to receive fingolimod 0.5 or 1.25 mg/day, or placebo for 2 years.[20, 21] All patients who completed the 24-month FREEDOMS/FREEDOMS II trial were eligible for the extension. Following FDA approval of fingolimod 0.5 mg dose, all patients receiving fingolimod 1.25 mg in study extensions were switched to fingolimod 0.5 mg. The extension phase continued until the umbrella safety extension study (ClinicalTrials.gov identifier, NCT01201356) opened for enrollment for patients who participated in all completed fingolimod studies.[23, 24]

Assessments

Standardized neurological assessments, including determination of EDSS were conducted at baseline and every 3 months. MRI scans were obtained at baseline, months 6, 12, and 24 or at the end of the study, if the patient discontinued the study prematurely.[20, 21] MRI lesion activity and BVL were assessed by a central reading site (Medical Image Analysis Center, Basel, Switzerland) that remained blinded for clinical data and randomization. MRI protocols and analysis methods have been detailed elswhere[20,21]

Outcome measures

The NBV was assessed at baseline using Structural Image Evaluation using Normalization of Atrophy, Cross-sectional (SIENAX).[25] Disability worsening was defined as an increase in the EDSS score of ≥1 point sustained for ≥3 months (1.5 points if baseline EDSS = 0).

Statistical analysis

The analysis was run in two separate steps. In step 1, a cross-sectional analysis of baseline data was performed. Baseline factors that are correlated with baseline NBV were assessed for the pooled dataset from FREEDOMS and FREEDOMS II trials. A multivariate linear regression analysis, including age, sex, disease duration, EDSS, T2 lesion volume (T2LV) and trial indicator (reperesenting the studies), was performed and the partial correlation coefficient of each factor (adjusted for all the other variables) with NBV was estimated. Variables with a skewed distribution (T2LV and disease duration) were transformed in three ordered groups defined by the tertiles of their distribution.

Age and EDSS score were treated as continuous variables, whereas sex and trial were categorical variables. The “trial” effect was estimated and compared to the residual variability of NBV (after adjusting the regression model for all the other baseline characteristics) to understand the comparability of the NBV measure between the two trials, adjusting for the different population characteristics. The trial effect on NBV was evaluated by using the greatest standardized mean difference (Cohen’s d) between a model including and excluding the trial effect. The trial effect can be considered negligible if Cohen’s d<0.20. Then, for each individual patient, an expected NBV (eNBV) according to individual baseline characteristics was calculated using the coefficients from the regression model, according to the linear formula:

$$eNBV = a + b*T2LV + c*EDSS + d*Age + e*MS\ duration + f*Male\ sex$$

Where *a* is a constant and b, c, d, e, f are the partial coefficients estimated by the regression model.

The difference between the observed NBV and the expected (eNBV)-value from the statistical model for each patient indicates whether the patient has a NBV higher or lower than expected (according to demographic and disease status).

Based on the distribution of these differences, patients were categorized into three groups: Low-NBV for patients for whom the NBV was more than 1 SD below the eNBV, medium-NBV for patients for whom the NBV was within 1 SD from the eNBV, and high-NBV for patients for whom the NBV was more than 1 SD above the eNBV.

In step 2, the prognostic values of the three defined NBV groups on disability (EDSS) change over 2 years were evaluated in the placebo-treated patients and in the overall population, adjusting for treatment effect. In both the FREEDOMS and FREEDOMS II trials, baseline characteristics were similar between the fingolimod 0.5 mg and 1.25 mg groups and, the two doses of fingolimod showed similar results across all efficacy outcomes (relapse rates, clinical disability worsening, and focal MRI lesion activity and brain volume loss). Hence, the current analysis was simplified by pooling the fingolimod 0.5 mg and 1.25 mg groups (presented together as the fingolimod group). The prognostic value of the NBV was also evaluated on the whole cohort (patients originally randomized to placebo and fingolimod arms) over the 4-year extension period. Finally, a differential fingolimod effect as compared to placebo according to the different baseline NBV groups was assessed.

The impact of the baseline NBV category, on the cumulative risk of 3-month confirmed disability progression (CDP) was analyzed using a Cox model. The fingolimod treatment effect on CDP for each baseline NBV category was evaluated by a treatment by group interaction test. The difference in the probability of worsening according to baseline NBV in placebo- and fingolimod-treated patients was displayed using Kaplan–Meier plots.

RESULTS

NBV at baseline estimated from the patient's demographics and disease characteristics

Complete MRI and clinical baseline data were available for 2342 patients from the pooled intent-to-treat populations from FREEDOMS (N=1267), and FREEDOMS II (N=1075) studies. **Table 1** reports the baseline characteristics of the patients included in this analysis. The overall baseline average NBV (±SD) was 1518 (±84) cm^3 and was similar between the FREEDOMS (1514 [±85] cm^3) and FREEDOMS II (1521 [±83] cm^3) populations, and the distributions were nearly superimposable. T2LV, EDSS, age, disease duration, and sex showed significant correlations with NBV ($p<0.001$; for all; **Table 2**).

The eNBV according to each patient characteristic can be calculated as:

eNBV = 1763 - 34.6*T2LV - 8.2*EDSS - 3.3*Age - 10.8*MS duration - 23.4*Male sex

The numerical details for the eNBV calculation are reported (**Table 2)**, with examples of the calculation of eNBV for patients with different baseline and disease characteristics (**Figure 1)**. The R^2 value of the model was 36% indicating that the variables included were able to account for 36% of the variance in baseline NBV among patients.

The distribution of the differences between the observed NBV and eNBV (calculated for each patient) is shown in (**Figure 2a**). The SD of this distribution was 67 cm^3 and represented the between-patient variability in NBV not accounted for by the examined characteristics. The average NBV (adjusted for all the other baseline factors) was 14 cm^3 higher in FREEDOMS II than in FREEDOMS ($p<0.001$) indicating significant trial effect. However, since the greatest standardized mean difference of model residuals between individual studies (Cohen's d) was 0.13 (<0.20), the trial effect was considered

negligible as compared to the variability of NBV between patients and was not included in the calculation of the eNBV for each patient.

The distribution of NBV differences from expected values (using 1 SD [67cm^3] as the distance from the expected value) was used to categorize the patients in the low-NBV, medium-NBV and high-NBV groups.

- Low NBV: NBV< eNBV − 67cm^3
- Medium NBV: eNBV − 67cm^3 ≥ NBV ≤ eNBV + 67 cm^3
- High NBV: NBV > eNBV + 67 cm^3

According to this criterion, 365 patients (15.5%) were categorized as having low-NBV, 1610 (69%) as having medium-NBV, and 367 (15.5%) as having high-NBV (**Figure 2b**). As the classification is individualized according to each patient's demographic and disease characteristics at baseline (**Figure 1**), the same value of NBV can result in a different classification depending on the subject's characteristics. For example, an NBV of 1500 cm^3 would classify a young, non-disabled male patient with a low T2LV and a short disease duration in the low-NBV group, whereas the identical 1500 cm^3 NBV value in an elderly, highly disabled patient with a high T2LV and a long disease history, would fall into the high-NBV group (**Figure 1**).

Prognostic value of the proposed NBV categories for predicting disability worsening

The impact of the categorization according to baseline NBV, adjusted for all the other relevant characteristics in the model, on the cumulative risk of 3-month CDP over 2

years is shown in **Table 3 and Figure 3**. The probability of 2-year confirmed disability worsening, after adjusting for treatment effect, was associated with the NBV categorization (p=0.002). Taking the high-NBV group as a reference, the hazard ratios (HRs) for the medium-NBV and low-NBV groups were 1.23 (95% CI 0.92;1.63, p = 0.16) and 1.75 (95% CI 1.26;2.44, p = 0.001), respectively. In placebo-treated patients, the proportion of patients with disability worsening was 20% (standard error [SE] 4%) in the high-NBV group; 25% (2%) (HR = 1.30, 95% CI 0.83;2.05, p=0.25) in the medium-NBV group; and 37% (5%) (HR = 2.13, 95% CI 1.26;3.61, p=0.005) in the low-NBV group. The corresponding values in the fingolimod-treated patients were 16% (2%), 19% (1%), and 23% (3%), respectively (**Figure 3a**).

The treatment effect tended to be more evident in patients in the low-NBV group (HR = 0.58, 95% CI 0.38;0.88) than in patients in the medium-NBV (HR = 0.72, 95% CI 0.57;0.90) and in the high-NBV (HR=0.80, 95% CI 0.47;1.36) groups. However, the treatment effect was not significantly different across the three NBV groups since the p-value for interaction was not significant (p = 0.57).

A consistent pattern was also observed for cumulative risk of 6-month CDP over 2 years, although fewer events occurred (data not shown). Furthermore, the prognostic value of NBV in predicting disability worsening was confirmed over longer periods, up to 4 years, using data from the FREEDOMS and FREEDOMS II extension studies (**Figure 3b**).

DISCUSSION

MS is a complex disease of the central nervous system. Although macroscopic, inflammatory lesions are the most evident aspect of MS pathology, a widespread pathology also occurs.[1] Global brain atrophy, which starts at the earliest stage of MS and progresses through disease course, is the expression of such diffuse tissue pathology.[10] While quantification is routinely used for MRI-derived assessment of lesions in clinical practice to guide clinical decisions[22] this is much less common for MRI-derived cross-sectional and longitudinal measures of atrophy. Including BVL as a cumulative measure of widespread pathology of the CNS in RRMS, should provide a more comprehensive and balanced assessment of the patient disease course.

Brain volume in MS depends on many demographic (age, sex) and clinical characteristics (disease duration, disability level etc.).[17] In this study, we quantified the distribution of NBV in a large cohort of RRMS patients, adjusting for the relevant baseline characteristics in order to identify the NBV values of those subjects that are "far" from their expected values. By using this methodology, patients are classified as having a high, medium or low NBV relative to all other patients with similar baseline characteristics. We demonstrated that it is possible to define values for each individual patient of what we called "expected NBV" based on adjustments for disease-relevant covariates (T2LV, age, EDSS, disease duration, and sex) and proposed a formula which can be used to calculate the NBV cutoff for an individual patient. The putative clinical relevance of this was that, based on these individualized values of NBV, it is possible to identify patients who are at a higher risk for future disability worsening.

The present analysis represents the first attempt to define individual cutoff values of brain volume, acknowledging the patients baseline characteristics. Our results suggest that the classification of patients into low, medium, and high NBV categories allows for a better characterization of the patient risk of future disability worsening and may inform individual management through the course of the disease. It needs to be stressed, however, that the approach used here has several limitations that restricts its utility in clinical practice[26]. These include lack of standardization and varying quality in the MRI acquisition, as well as potential for inaccurate image post-processing. For example, data reported here might be different with different MRI acquisitions (e.g., field strength, voxel size, type of sequence, etc.) or different imaging analysis (e.g., software packages different from SIENAX, different imaging reading site, etc). These can be relevant sources of error that can affect the measurement estimation. Additionally, generic factors, such as lifestyle habits, genetic load, and concomitant pathologic conditions may also affect BV measurements and need to be taken into account when interpreting brain atrophy, particularly in the single patient assessment. Despite the above limitations, progress in computational technologies allows to expect a more convincing clinical use of MRI-based brain volumetry in a number of neurological disorders, including RRMS.[27] Further studies are needed on independent datasets to evaluate whether cutoffs need to be re-calibrating after correcting for different MRI acquisition protocols and analysis pipelines.

In the present study, the overall baseline average NBV was very similar between the two large, multi-center clinical datasets (FREEDOMS and FREEDOMS II) with distributions that were nearly superimposable. This is indicative of a great homogeneity

in the NBV values, which is of paramount importance for the validity of the results. To create common cut-offs, in fact, there must be no systematic differences in NBV when measured in different settings. For this reason, these results need to be validated in other datasets to understand whether the same methodology and the same individualized values of NBV can be extended to other patient populations, especially to those beyond controlled clinical trials. Moreover, it is important to re-emphasize that the present analysis was based on the NBV calculations using SIENAX and cannot be extended to other methods of BVL assessment. Finally, although SIENAX is a widely used, standard method in clinical trials, its use in clinical practice on individual patients might be challenging and requires experienced personnel to perform accurate measurements.[28]

Patients included in clinical trials are usually active inflammatory patients, due to the inclusion criteria aimed at enriching patients for activity at baseline. Whereas baseline inflammatory activity can affect the rate of brain volume change over time, the baseline value of NBV is much less affected by this and did not result associated to MRI activity assessed by gadolinium enhancement in this large dataset. More in general, however,, NBV is a measure of the brain volume loss occurred in the given patient until the time of the measurement, and for this is less intrinsically dependent of the inflammatory activity at the time, but rather dependent on the overall disease burden. In that respect, the formula used to calculate the expected NBV included the T2 lesion burden, as expression of the overall focal pathology that can be associated to the baseline brain volume. As a consequence, the findings reported here can be generalizable to a typical clinic population of MS patients, composed of a mixture of active and inactive patients.

When discussing brain volume, brain reserve and cognitive reserve also need to be considered. Several studies have indicated that both, brain reserve and cognitive reserve, may serve as a buffer against future cognitive impairment.[29-31] Our findings generally support this concept, as we showed that individuals with a higher-than-expected NBV had a better prognosis for disability worsening over 2 and 4 years, than those with a lower NBV. This suggests that a window of opportunity exists, wherein therapeutic interventions can be the most effective in delaying disease worsening implying that an early intervention with an effective disease-modifying therapy may be linked to a more favorable long-term outcome.

A number of studies have consistently shown a close correlation between BVL and disability worsening, particularly in the long-term.[14, 17, 32] A recent post hoc analysis of data from the fingolimod phase III trials provided evidence that NBV at baseline was the strongest predictor of disability in RRMS patients and the correlation of BVL with worsening disability was stronger over four years than in the first two years on study.[17] The prognostic value of brain atrophy rates in predicting long term disability has also been demonstrated in other studies.[14, 32] A large multicenter study reported that whole and central brain atrophy over one year can predict EDSS scores at 10 years in a heterogeneous group of MS patients.[14] In a follow-up study reassessing RRMS patients from a phase III trial of interferon β-1a, brain atrophy rate was found to be the best MRI predictor of disability status at the 8-year follow-up.[32] Consistent with these findings, we observed that the baseline NBV values in individual patients predicted disability worsening over 2 to 4 years.

In summary, we propose a methodology to calculate quantitative reference values for single-point NBV measurements, that allows to identify patients having a lower value of brain volume than that expected according to their clinical and demographic profile. This classification can be used to assign RRMS patients to distinct risk groups regarding future disability worsening and may help physicians to identify and appropriately treat patients at a higher risk of disability worsening.

Disclosures

Dr. Sormani has received compensation for consulting services and speaking activities from Biogen , Genzyme, Merck Serono, Novartis, Roche and Teva.

Dr. Kappos's Institution (University Hospital Basel) received in the last 3 years and used exclusively for research support: *steering committee, advisory board and consultancy fees* (Addex, Actelion, Alkermes, Almirall, Bayer Health Care Pharmaceuticals, Biogen Idec, CLC Behring, Excemed, Genentech, GeNeuro SA, Genzyme, Merck Serono, Mitsubishi Pharma, Novartis, Octopharma, Pfizer, Receptos, Roche, sanofi-aventis, Santhera, Teva and UCB.); *speaker fees* (Bayer Health Care, Biogen, Merck, Novartis, Sanofi-Aventis, Teva); *support of educational activities* (Bayer Health Care, Biogen, CSL Behring, Genzyme, Merck, Novartis, Sanofi, Teva); *royalties* (Neurostatus Systems GmbH); *grants* (Bayer Health Care, Biogen, Merck, Novartis, Roche, Swiss MS Society, the Swiss National Research Foundation, the European Union, Roche Research Foundations).

Dr. Radue has received research support from Biogen Idec, MerckSerono, Novartis and Actelion.

Dr. Cohen has received personal compensation for serving as a consultant or speaker from Biogen Idec, Lilly, Novartis, and Vaccinex and received research support through his institution from Biogen Idec, BioMS, Genzyme, Novartis, Synthon, and Teva.

Dr. Barkhof has received consulting fees from Bayer Schering Pharma, Sanofi-Aventis, Biogen Idec, Teva, Merck Serono, Novartis, Roche, SynthonBV, Janssen research and

Genzyme. FB has received fees for development of educational presentations including service on speakers' bureaus from Serono Symposia Foundation and Medscape and research grant from Dutch MS Society, EU-FP7.

Dr Sprenger´s previous institution, the University Hospital Basel, has received payments for consultation from Mitsubishi Pharma, Eli Lilly, Genzyme, Novartis, ATI, Actelion, Electrocore, Biogen Idec and Allergan. These payments were used exclusively for research support. TS has received research grants from the Swiss MS Society, Swiss National Research Foundation, EFIC-Grünenthal and Novartis Pharmaceuticals Switzerland. TS has received support for travel, accommodation and meeting expense from Teva, Genzyme, Novartis, Allergan and Biogen. Dr. Sprenger´s current institution has received payments for speaking/consultation from Novartis and Genzyme.

Dr. Piani Meier, **Dr Haering** and **Dr. Tomic** are employees of Novartis Pharma AG, Basel Switzerland.

Dr. De Stefano has served on scientific advisory boards and steering committees of clinical trials for Merck Serono SA, Novartis Pharma AG and Teva and has received support for congress participation or speaker honoraria from Bayer Schering AG, Biogen Idec, Merck Serono SA, Novartis Pharma AG, Sanofi-Aventis and Teva.

Study Funding (FREEDOMS/FREEDOMS II): Funded by Novartis Pharma AG, Basel, Switzerland

Acknowledgements

The authors would like to thank Rahul Birari and Rishard Salie (Novartis, Medical Communications) for editorial support and coordinating author reviews.

References

1. Filippi M, Rocca MA, Barkhof F, et al. Association between pathological and MRI findings in multiple sclerosis. *Lancet Neurol* 2012; 11: 349–360.

2. Kutzelnigg A and Lassmann H. Pathology of multiple sclerosis and related inflammatory demyelinating diseases. *Handb Clin Neurol* 2014; 122: 15–58.

3. Sormani MP, Arnold DL and De Stefano N. Treatment effect on brain atrophy correlates with treatment effect on disability in multiple sclerosis. *Ann Neurol* 2014; 75: 43–49.

4. Sormani MP, Bonzano L, Roccatagliata L, Cutter GR, Mancardi GL and Bruzzi P. Magnetic resonance imaging as a potential surrogate for relapses in multiple sclerosis: a meta-analytic approach. *Ann Neurol* 2009; 65: 268–275.

5. Takao H, Hayashi N and Ohtomo K. A longitudinal study of brain volume changes in normal aging. *Eur J Radiol* 2012; 81: 2801–2804.

6. Sidaros A, Skimminge A, Liptrot MG, et al. Long-term global and regional brain volume changes following severe traumatic brain injury: a longitudinal study with clinical correlates. *Neuroimage* 2009; 44: 1–8.

7. Scahill RI, Frost C, Jenkins R, Whitwell JL, Rossor MN and Fox NC. A longitudinal study of brain volume changes in normal aging using serial registered magnetic resonance imaging. *Arch Neurol* 2003; 60: 989–994.

8. Hedman AM, van Haren NE, Schnack HG, Kahn RS and Hulshoff Pol HE. Human brain changes across the life span: a review of 56 longitudinal magnetic resonance imaging studies. *Hum Brain Mapp* 2012; 33: 1987–2002.

9. 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: 1868–1876.

10. De Stefano N, Airas L, Grigoriadis N, et al. Clinical Relevance of Brain Volume Measures in Multiple Sclerosis. *CNS Drugs* 2014; 28: 147–156.

11. De Stefano N, Stromillo ML, Giorgio A, et al. Establishing pathological cut-offs of brain atrophy rates in multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2015: 93–99.

12. Sormani M, De Stefano N, Francis G, et al. Fingolimod effect on brain volume loss independently contributes to its effect on disability. *Mult Scler* 2015; 21: 916–924.

13. Kappos L, De Stefano N, Freedman MS, et al. Inclusion of brain volume loss in a revised measure of 'no evidence of disease activity' (NEDA-4) in relapsing-remitting multiple sclerosis. *Mult Scler* 2015: Epub ahead of print 19 Nov 2015. DOI: 10.1177/1352458515616701.

14. Popescu V, Agosta F, Hulst HE, et al. Brain atrophy and lesion load predict long term disability in multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2013; 84: 1082–1091.

15. Jacobsen C, Hagemeier J, Myhr KM, et al. Brain atrophy and disability progression in multiple sclerosis patients: a 10-year follow-up study. *J Neurol Neurosurg Psychiatry* 2014; 85: 1109–1115.

16. Bermel RA and Bakshi R. The measurement and clinical relevance of brain atrophy in multiple sclerosis. *Lancet Neurol* 2006; 5: 158–170.

17. Radue EW, Barkhof F, Kappos L, et al. Correlation between brain volume loss and clinical and MRI outcomes in multiple sclerosis. *Neurology* 2015; 84: 784–793.

18. Benedict RH, Weinstock-Guttman B, Fishman I, et al .Prediction of neuropsychological impairment in multiple sclerosis: comparison of conventional magnetic resonance imaging measures of atrophy and lesion burden. Arch Neurol. 2004;61:226-30.

19. Sacco R, Bisecco A, Corbo D, et al. Cognitive impairment and memory disorders in relapsing-remitting multiple sclerosis: the role of white matter, gray matter and hippocampus. J Neurol. 2015;262:1691-7.

20. Benedict RH, Weinstock-Guttman B, Fishman I, et al. Brain MRI lesions and atrophy are associated with employment status in patients with multiple sclerosis. J Neurol. 2015;262:2425-32.

21. Mowry EM, Beheshtian A, Waubant E, et al. Quality of life in multiple sclerosis is associated with lesion burden and brain volume measures. Neurology. 2009;72:1760-5.

18. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH harmonised tripartite guideline. Guideline for good clinical practice. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6/E6_R1_Guideline.pdf (1996, accessed 22 April 2016).

19. World Medical Association. Declaration of Helsinki: Ethical principles for medical research involving human subjects. http://www.wma.net/en/30publications/10policies/b3/index.html.pdf?print-media-type&footer-right=[page]/[toPage]; (accessed 22 April 2016).

20. Kappos L, Radue E-W, O'Connor P, et al. A placebo-controlled trial of oral fingolimod in relapsing multiple sclerosis. *N Engl J Med* 2010; 362: 387–401.

21. Calabresi PA, Radue EW, Goodin D, et al. Safety and efficacy of fingolimod in patients with relapsing-remitting multiple sclerosis (FREEDOMS II): a double-blind, randomised, placebo-controlled, phase 3 trial. *Lancet Neurol* 2014; 13: 545–556.

22. Polman CH, Reingold SC, Edan G, et al. Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". *Ann Neurol* 2005; 58: 840–846.

23. Cohen J, Von Rosenstiel P, Gottschalk R, Pradhan A, Zhang Y and Kappos L. Long-term Safety of Fingolimod: An Interim Analysis of the LONGTERMS Cohort (S4.006). *Neurology* 2015; 84.

24. Kappos L, O'Connor P, Radue EW, et al. Long-term effects of fingolimod in multiple sclerosis: the randomized FREEDOMS extension trial. *Neurology* 2015; 84: 1582–1591.

25. Smith SM, Zhang Y, Jenkinson M, et al. Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. *Neuroimage* 2002; 17: 479–489.

26. Giorgio A, Battaglini M, Smith SM and De Stefano N. Brain atrophy assessment in multiple sclerosis: importance and limitations. *Neuroimaging Clin N Am* 2008; 18: 675–686.

27. Giorgio A and De Stefano N. Clinical use of brain volumetry. *J Magn Reson Imaging* 2013; 37: 1–14.

28. Miller DH, Barkhof F, Frank JA, Parker GJ and Thompson AJ. Measurement of atrophy in multiple sclerosis: pathological basis, methodological aspects and clinical relevance. *Brain* 2002; 125: 1676–1695.

29. Amato MP, Razzolini L, Goretti B, et al. Cognitive reserve and cortical atrophy in multiple sclerosis: a longitudinal study. *Neurology* 2013; 80: 1728–1733.

30. Sumowski JF, Rocca MA, Leavitt VM, et al. Brain reserve and cognitive reserve in multiple sclerosis: what you've got and how you use it. *Neurology* 2013; 80: 2186–2193.

31. Schwartz CE, Quaranto BR, Healy BC, Benedict RH and Vollmer TL. Cognitive reserve and symptom experience in multiple sclerosis: a buffer to disability progression over time? *Arch Phys Med Rehabil* 2013; 94: 1971–1981.

32. Fisher E, Rudick RA, Cutter G, et al. Relationship between brain atrophy and disability: an 8-year follow-up study of multiple sclerosis patients. *Mult Scler* 2000; 6: 373–377.