Serial decline of no evidence of disease activity-4 in Latin-American patients with multiple sclerosis after 1 and 2 years of follow-up

Short title: NEDA-4 in Latin-American patients

Carlos Guevara[1], Violeta Diaz [1], Cristian Garrido[1], Melissa Martinez[1], Patricia Orellana[1], Pablo Alarcón[1], Carlos Silva-Rosas[1], Gareth J Barker[2], Matthew J Kempton[2], José de Grazia[1]

[1]Faculty of Medicine, Hospital Clínico José Joaquín Aguirre, Santos Dumont 999, Universidad de Chile, Santiago, Chile

[2]Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom

Corresponding author

Carlos Guevara

Santos Dummont 999, Universidad de Chile, Santiago, Chile

Tel.: +56994423761

E-mail: neurocrs@gmail.com

Keywords

Relapsing/remitting; Quantitative MRI; Outcome measurement (NEDA-4); Multiple sclerosis;Brain Atrophy.

Abstract

***Background*:** In relapsing-remitting multiple sclerosis (RRMS), no evidence of disease activity-3 (NEDA-3) is defined as: no relapses, no disability progression, and no active lesions on MRI scans. NEDA-4 status is defined as meeting all NEDA-3 criteria plus having an annualized brain volume loss (a-BVL) of ≤ 0.4%. There are currently no prospective real-world studies presenting data on NEDA-4. We aim to determine the proportion of patients failing to meet one or more NEDA-4 criteria and the contribution of each component to this. The study recruited Latin-American Chilean RRMS patients in the early disease stages, at a single center and during routine clinical assessment. We hypothesize that NEDA-4 has the potential to capture the impact of DMT therapies in RRMS and that considering NEDA-4 status, clinicians may select a second-line DMT earlier in the disease course to control subclinical disease activity of RRMS.

Methods: Forty-five patients were prospectively followed 1 and 2 years after their baseline assessment at the University of Chile Hospital. SIENA software was used to assess a-BVL.

Results: At baseline, the patients had a mean age of 33.0 years (range 18-57), disease duration of 1.9 years (0.4-4), Expanded Disability Status Scale score of 1.3 (0-4), and 67% were female. The majority had RRMS (91% while 9% had clinically isolated syndrome (CIS)). Seventy-three percent were on first line DMT (interferons (53%), glatiramer acetate (11%), teriflunomide (9.0%)), CIS patients were without DMT, and 18% (8/45) were on fingolimod. There was a serial decline in the proportion of no evidence of disease activity: after 1 and 2 years of follow-up 60% and 47% met NEDA-3 status, and 38% and 27% met NEDA-4, respectively. At the last follow-up 21% remained on interferons, 47% were now on fingolimod, 4% on alemtuzumab and 2% on natalizumab.

Conclusion: NEDA-4 has the potential to capture the impact of DMT therapies in RRMS. a-BVL may be a key domain to consider when assessing the effectiveness of the first line DMT. Considering NEDA-4 status, clinicians may select a second-line DMT earlier in the disease course to control subclinical disease activity of RRMS.

Introduction

Combined disease status assessments are increasingly explored to evaluate the overall impact of disease modifying treatments (DMTs) in multiple sclerosis.[1] During 2018-2019 a number of papers with real world clinical data on No Evidence of Disease Activity-3 (NEDA-3) have been published (Table 1). NEDA-3 is defined as the absence of all of the following: relapses, disability progression and MRI activity (new/enlarged T2 lesions and/or gadolinium-enhanced T1 lesions). However, individual components of NEDA-3 are somewhat impractical and their exact definitions vary in routine clinical practice, and even in randomized controlled trials (RCT) for assessing ongoing disease activity. [2]

Disease activity follow-up may be improved by considering quantitative magnetic resonance imaging (qMRI). In the present study, structural image evaluation, using normalization, of atrophy (SIENA)[3,4] was used to explore brain volume loss (BVL) in relapsing-remitting multiple sclerosis (RRMS) and to investigate the addition to NEDA-3 of a fourth criterion—no pathological BVL—proposed by Kappos et al.[5] During normal aging, the annual rate of cerebral BVL has been estimated as ≤0.4%.[6,7] NEDA-4 status is therefore defined as meeting all NEDA-3 criteria plus having an annualized BVL (a-BVL) of ≤0.4%.[5] NEDA-4 outcomes have exclusively been reported from *post-hoc* analyses from RCTs in carefully selected groups of patients, from multicentric and/or retrospective studies;[5] *to our knowledge there are no prospective real-world studies* presenting data on NEDA-4, let alone in Latin-American populations (Table 1). Khan et al[8] reported that a review of PubMed conducted in 2014 revealed that there were nearly 60,000 articles in total published on MS, with approximately 52,000 written in English; however, only 113 focused on African American (or black) patients and only 23 focused on Latin American patient populations with MS. Thus less than 0.04% of the literature on MS has focused on Latin American MS patients.

Table 1 shows that only two studies have reported NEDA-4 outcomes, both retrospective and in patients using fingolimod. It is noticeable that much discussion on NEDA-4 has taken place in the absence of prospective real word data. On one side this reflects the scepticism about using brain volume loss as a clinical biomarker, but also the challenge for neurologists and even neuroradiologists, in accessing quantitative MRI processing techniques.

To address these issues our prospective and observational study investigates for first time a real word experience of calculating serial NEDA-4 outcomes. We present prospectively collected clinical data from Latin-American Chilean RRMS patients in early disease stages, at a single center and during the routine clinical assessment. We aim to determine the proportion of patients failing to meet one or more NEDA-4 criteria after 1 and 2 years of follow-up and the contribution of each NEDA-4 component to this.

We hypothesize that NEDA-4 has the potential to capture the impact of DMT therapies in RRMS and that considering NEDA-4 status, clinicians may select a second-line DMT earlier in the disease course to control subclinical disease activity of RRMS.

Materials and methods

Patient assessment

Forty-five adult patients with RRMS according to the revised McDonald criteria[9] were followed at the University of Chile Hospital as part of routine clinical practice. This observational study was integrated into routine clinical practice and therefore the treating neurologists were free to make any therapeutic change. Between January 2016 and March 2019. Patients were scheduled to undergo clinical and MRI assessments at baseline screening, and at months 12 and 24 resulting in a total of three clinical assessments and three MRI brain scans. For both the baseline and follow-up assessments, the clinical data and MRI scans were acquired within 1 week of each other. Patients were assessed for the first time

when relapse free and at least 90 days after their last relapse. All RRMS patients were on DMT (Table 2) and patients were excluded if they had a clinical disease duration of ≥4 years and an EDSS score of ≥4. Due to the observational nature of the study, no further inclusion criteria were defined. The following definitions were used for the individual components of NEDA-3[1]. Relapse: the appearance of a new or worsening of a previously stable neurological abnormality, present for at least 24 hours and occurring in the absence of fever or infection, confirmed within 7 days of symptom onset. Focal MRI activity: new or enlarged T2 lesions and/or gadolinium-enhanced T1 lesions. Disability progression: an increase in the EDSS score of at least 0.5 points from the baseline score. Neuropsychological status was assessed using the Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS) and scores were assessed relative to normative data in a Hispanic population.[10,11] At 12 and 24 months worsening of at least 10% of the composite BICAMS measures and of the three instruments that compose BICAMS (the Symbol Digit Modalities Test (SDMT), California Verbal Learning Test 2 (CVLT2) and the Brief Visual Spatial Memory Test-Revised (BVMTR) were considered to be clinically meaningful. [10,11]

MRI acquisition

Between January 2016 and March 2019, patients underwent three MRI brain scans. Brain MRI was performed at baseline and follow-up on the same MRI system using the same imaging protocol (same pulse sequences and spatial resolution). MRI images were acquired on a 1.5 T Siemens MRI scanner. Axial T1-weighted images of the whole brain were obtained using a 3D inversion-recovery prepared spoiled gradient-echo (IR-SPGR) sequence. The following parameters were used in this volumetric sequence: field of view of 250x234 mm; matrix of 256x240 mm; repetition time of 12 ms; echo time of 5.68 ms; excitation flip angle of 15º; isotropic voxel size of 0.98 × 0.98 × 0.98 mm. Patients underwent a second and third MRI brain scan within two years of follow-up. Two neuroradiologists (JdG and PO, one of them with 20 years' experience and the other with 5 years' experience) assessed the MRI scans of every patient to rule out gross anatomical abnormalities. No MRI images included in this study showed

any structural abnormalities other than atrophy-related changes and demyelinating lesions. New lesions at follow up (T2 lesions and gadolinium-enhancing lesions) were evaluated by eye and were defined as hyperintensities > approximately 3 mm on T2-weighted/FLAIR scans or showing gadolinium enhancement on T1-weighted scans. The enlargement of T2 lesions from baseline to follow-up was also evaluated by eye. No quantitative measure was made regarding T2 lesion load. The number of T2 lesions per patient scan was divided into three categories: less than 10; 11 to 49; and 50 or more than.[12]

MRI analysis

All images were converted to NIFTI format using MRIcron software (http://people.cas.sc.edu/rorden/mricron/dcm2nii.html). Cross-sectional whole brain volumes and brain tissue volumes were estimated using SIENAX.[3] Before further processing, all data were anonymized by removing any reference to the patients' names and other demographic data from the image headers. Briefly, SIENAX extracts brain and skull images from the acquired MRI data. The brain image is then affine registered to Montreal Neurological Institute 152 space, using the skull image to determine the registration scaling. The registration scaling is then used to obtain a volumetric scaling factor, which is employed to normalize the tissue volume estimates. Segmentation with partial volume estimation is subsequently performed to calculate the total volume of brain tissue, including separate estimates of the volumes of gray matter (GM) and white matter (WM). Gray matter is divided into cortical gray matter (cGM) and deep gray matter (dGM). [13] The longitudinal SIENA processing algorithm has been validated and described in detail elsewhere.[3] The processing steps are as follows: (1) Brain extraction: segmentation of brain from nonbrain tissue for each scan, followed by skull extraction. (2) Registration: registration of the segmented brain from the second (follow-up) scan to that of the first (baseline) scan using a linear transformation. The two skull images are used as normalizing factors to constrain the scale and skew. (3) Tissue type segmentation: white matter and gray matter tissues are treated as one tissue and the cerebrospinal fluid as another tissue. (4) Change analysis: detection of the edge of the brain on the

registered baseline and follow-up image. At each edge point the displacement between the baseline brain edge and follow-up brain edge is determined. Finally, the mean displacement of brain surface at each edge point is converted to a global percentage change in brain volume by taking into account the baseline brain volume. Subjects were included in the study if they had three MRI scans of adequate quality and the brain extraction step in SIENA functioned correctly. One patient was excluded because of the presence of a frontal arachnoid cyst. SIENA is freely available (http://www.fmrib.ox.ac.uk/analysis/research/siena).

Statistical analyses

Statistical analyses of the clinical data were performed using Statistical Package for Social Sciences, version 22 (IBM). The results are presented as the means ± SDs. In all cases, a two-sided p of < 0.05 was considered significant. Comparisons between groups were assessed using a *t*-test. Changes in clinical scores were assessed using the Wilcoxon's signed rank test.

a-BVL was calculated by dividing the BVL values by the interscan interval in years. Clinical scores were annualized by dividing the unit change between the assessments by the assessment interval in years. To assess the contribution of the four different components of the NEDA-4 measure, hierarchical analysis of patients was performed according to individual disease activity criteria using the following hierarchy: relapses, disability progression, MRI activity and accelerated a-BVL.[5] In this analysis, patients who had an event for one outcome were removed from evaluation for any other outcomes from that point on. For example, if a patient had a relapse, the patient was removed from subsequent evaluation for disability progression, MRI activity and a-BVL (Figure 2 and 3).

Standard protocol approval and patient consent

Prior to inclusion, patients provided informed written consent for participation in the study. The study was conducted in accordance with international standards of good clinical practice (ICH guidelines and the Declaration of Helsinki). The project was approved by the local research ethics committees of the University of Chile Hospital, Santiago, Chile.

Results

Clinical findings (Table 2)

At baseline. Forty-five patients were followed up with three clinical and MRI assessments. 91% (41/45) had relapsing-remitting MS (RRMS) and 9% (4/45) clinically isolated syndrome (CIS). At baseline, the patients had a mean age of 33.0 years (18-57), disease duration of 1.9 years (0.4-4), EDSS score of 1.3 (0-4), and female 67%.

Seventy-three percent of patients were on the so called first-line DMT (interferons (53%), glatiramer acetate (11%), teriflunomide (9%)), 18% were on fingolimod because of intolerance to interferons or used as first DMTs. CIS (9%) patients were without DMT.

First follow-up (at year 1, Figure 1): 20% (9/45) of the patients had at least one relapse, 16.0% (7/45) had disability progression, 11.0% (5/45) had new T2 lesions, and 9.0% (4/45) had gadolinium-enhanced T1 lesions. Sixty percent (27/45) achieved NEDA-3. An a-BVL of >0.4% was observed in 52.0% (23/45) (mean: 0.45% ±0.68). Including a-BVL (threshold of 0.4%), the proportion of patients achieving NEDA-4 was 38% (17/45). Thirty-three percent (15/45) were now on second-line DMT: 29% (13/45) on fingolimod and 2% (1/45) on natalizumab.

Second follow-up (at year 2, Figure 2): Thirty-three percent (15/45) had at least one relapse, 11% (5/45) had disability progression, 7.0% (3/45) had new/enlarged T2 lesions, and 13.0% (6/45) had gadolinium-

enhanced T1 lesions. 47% (21/45) of patients achieved NEDA-3. An a-BVL of >0.4% was observed in 49% (22/45) (mean: 0.48% ±0.57) and 27% (12/45) achieved NEDA-4. Fifty-one percent (23/45) were now on second-line DMTs: fingolimod 47% (21/45), alemtuzumab 4% (2/45) and natalizumab 2% (1/45).

Cognitive measures (Table 2)

At baseline, fifty-three percent of patients performed badly on the Symbol Digit Modalities Test (mean ±SD: 43 ±13), 42% on the California Verbal Learning Test 2 (54 ±10), and 20% performed badly on the Brief Visual spatial Memory Test-Revised (23 ±7). The average BICAMS was 121±25. At **1 year**, BICAMs deterioration by at least 10% was seen in 13% of patients (6/45). Forty-seven percent of patients performed badly on the Symbol Digit Modalities Test (mean± SD: 45±13), 40% on the California Verbal Learning Test 2 (54 ±12), and 18% performed badly on the Brief Visual spatial Memory Test-Revised (25 ±7). In the hierarchical analysis, adding BICAMS would result in 50% achieving a putative NEDA-4. **At 2 years,** BICAMS deterioration was seen in 7% of patients (3/45) relative to baseline measurement. 39% of patients performed badly on the Symbol Digit Modalities Test (46 ±15), 39% on the California Verbal Learning Test 2 (54 ±12), and 16% performed badly on the Brief Visual spatial Memory Test-Revised (26 ±10). However, there was a significant increase between the baseline BICAMS and the last follow-up BICAMS (p=0.02) and between the baseline SDMT and the last SDMT follow up (p=0.07). At this time point, inclusion of BICAMS did not contributed to change the proportion of a putative NEDA-4 status including this measure.

Tissue partition MRI-derived measures (Table 3) Baseline qMRI volumes were between within normal limits. From baseline to 1.8 years after baseline, white matter volume mean decreased from 816 ml ±34 to 808 ml ± 31 (p= 0.015), peripheral gray matter decreased from 584 ml ± 30 to 578 ml ±29 (p=0.049), and whole brain volume decreased from 1565 ml ±51ml to 1553 ml± 48 (p=0.011). MRI-derived measures during the first period of follow-up (baseline- year 1) did not show significant changes.

Discussion

This longitudinal study showed a serial decline in the proportion of patients with no evidence of disease activity as defined by the concepts of NEDA-3 and NEDA-4; after 1 and 2 years of follow-up 60% and 47% met NEDA-3 status and 38% and 27% met NEDA-4, respectively. At baseline, 53% of the patients were on interferons and at the 2 year follow-up 21% remained on interferons. Based on NEDA-3 criteria, the treating neurologist replaced the use of interferons by fingolimod, natalizumab and alemtuzumab as second line therapies.

The three components of NEDA-3 may not be suitable for the determination of timely treatment failure in routine clinical practice. In a large retrospective cohort (1594 patients with RRMS), 810 patients showed evidence of disease activity (≥ 1 relapse or an increase in the EDSS score by ≥0.5 points and/or MRI activity) after at least 2 years of follow-up.[14] Of these 810 patients, 32% were assessed as having progressive disease and 65% as stable; despite the clinical and MRI changes, the treating neurologist did not recommend treatment optimization in the putatively stable patients. The components of NEDA-3 reflect the ongoing disease status imperfectly, and their variability in terms of definitions and acquisition limits their effectiveness as outcome measures. Relapses and gadolinium-enhanced lesions on T1-weighted scans reflect only focal inflammatory disease activity, underestimating the presence of early diffuse and clinically silent neurodegeneration in RRMS.[15] The use of relapses as outcome measures also has major limitations. Relapses are relative rare events[2]; in a pivotal clinical trial untreated MS patients showed an annual relapse rates of 0.4 (0.34-0.47).[16] Particularly in early stages of the disease the number of relapses may therefore underestimate the ongoing pathological changes and overestimate the efficacy of DMTs.

New or enlarging lesions on T2-weighted scans are one of the main parameters used for following disease activity in RRMS. However, in clinical practice, the detection of new or enlarged T2 lesions is limited by technical and methodological factors. Manual counting of T2 lesions is imprecise, and the number of new

T2 lesions is typically specified only approximately, or as >10 when many are present.[12] Moreover, the cortical lesion burden is poorly visualized by most routine MRI protocols.[17] In RCTs, T2 lesion volumes are often automatically counted, which may be more accurate to assess T2 lesion burden.[16] In the current study, the longitudinal measurement of T2 lesions was a rather insensitive parameter for assessing disease progression, as the T2 lesion burden accounted for only 11% (5/45, Figure 3) of the unique events.

Identifying subclinical signs of disease activity is imperative to prevent the neurodegenerative aspects of RRMS and reduce progression to irreversible disability. The frequency of conversion from RRMS to a secondary progressive multiple sclerosis (SPMS) increases with duration of disease (12% at 5 years; 41% at 10 years).[18] Assessing BVL early during the course of the disease could help identify groups of people with RRMS who may benefit from particular types of therapies before the progression to SPMS, when patients seem to receive no benefit from DMTs. Accelerated a-BVL is predictive of disability progression and cognitive decline in the long term.[19,20]

In this clinical series two questions arise. Firstly, should this data encourage early treatment with fingolimod or monoclonal antiboidies instead of interferons? Secondly should the treating neurologist consider switching therapy to potentially more effective drugs in patients who have not achieved NEDA-4 status.? This topic is controversial, particularly regarding patients who have accelerated a-BVL only (22 % in this study). Although SIENA has been shown to have a low estimation error for atrophy rate over the whole brain (0.5%)[3,4,13] confounding factors in determining the rate of BVL require further discussion. The two-year follow-up period in the present study seems to be clinically meaningful for switching the initial DMT owing to disease activity. A one-year follow up period may overestimate BVL because of the resolution of the early anti-inflammatory effect of DMTs and steroids (pseudoatrophy) and BVL would need to be assessed considering this caveat. Thus, a two-year period has been suggested as a more robust approach when measuring BVL.[21] BVL assessment at 12 and 24 months may be an

appropriate approach for assessing the pattern of disease activity and overcoming the confounding factor of pseudoatrophy.

Depicting precise cognitive profiles in patients with RRMS would potentially assists therapeutic decisions, especially at earlier stages. Thus, cognitive impairment has been proposed as another component to assess evidence of disease activity and for being integrated into therapeutic algorithms for RRMS. This clinical study has assessed BICAMS as a fast neurocognitive tool capable of being applied in the routine clinical practice by the treating neurologist. In order to improve the sensitivity for detection of chronic clinical deterioration, Sacca *et al* [22] have integrated BICAMS and the results of two orientation tests (Mini Mental State Examination and the Montreal Cognitive Assessment) into the EDSS scoring system in place of the Cerebral Functional System. They have shown that the sensitivity to detect cognitive impairment in a cross-sectional fashion increased by 25% in the group of patients with EDSS score <4. In the current cohort, cognitive impairment was already present at baseline in at least 50% of these rather young patients.[23] At the group level, this percentage did not increase in the following years and thus BICAMS may have not been a sensitive marker of disease progression when there is a floor effect due to a high proportion of cognitive impairment at baseline.

The pathological processes responsible for atrophy are likely to involve the death of different types of brain cells. In the cohort reported here, cortical gray matter (cGM) and deep gray matter (dGM) losses were the contributors to accelerated a-BVL, but supratentorial white matter volume was also reduced significantly after 2 years of follow-up. These *in vivo* data provide further evidence that tissue loss in the cGM and dGM structures occurs early in the course of RRMS. However, longitudinal studies have shown that gray matter atrophy is a better predictor of disease progression than white matter atrophy.[17,24] A recent large multicenter longitudinal study in 3604 patients augmented this finding by showing that that dGM loss drives disability accumulation in RRMS.[25] However, the accurate segmentation of GM is difficult to achieve. The cortex (cGM) is a thin layer of GM surrounded by WM on one side and CSF on

the other, both of which produce partial volume effects that confound its delineation. Moreover, the automated segmentation of dGM is much less accurate than that of cGM. For instance, the automated techniques tended to misclassify large portions of dGM as WM [26]. These shortcomings in the segmentation of GM structures should be taken into account when interpreting these findings.

This study has a number of limitations. Concerns regarding the biological validity of these BVL changes in RRMS remain. Factor such as alcohol,[27] mild traumatic brain injury, [28] smoking, genetics, diabetes mellitus [29], and hydration/dehydration can cause changes in brain size.[30] Moreover brain volumes seem to fluctuate throughout the day, decreasing from morning to evening.[31] Although these clinical factors may add to the variance in the measurements at the individual level, making it more difficult to detect real changes, clinicians and patients may be able to allow for these potential confounders. Various other sources of error related to image acquisition can affect MRI atrophy quantification: image artifacts due to head motion, poor signal-to-noise ratio, partial head coverage, imperfect patient repositioning in a longitudinal study and image acquisition with nonidentical scan parameters. Even small changes could be argued to be an artifact caused by, for example, cardiac pulsations. However, due to the duration of the MRI acquisition, all images are effectively an average over several minutes, so the effects of cardiac pulsation should be averaged out. A strength of this study is the prospectively collected clinical data, with a high-quality control standard. Throughout the duration of the study, the patients underwent the same MRI protocol on the same MRI scanner at a single site. When the patients are prospectively recruited in this way, through a single center, the risks of data variability may be substantially reduced.

We have not included the spinal cord assessment related to the lesions and atrophy, which may be of clinical significance as spinal cord pathology is a major contributor to RRMS disability. Indeed, the rate of spinal cord atrophy is greater than that of brain atrophy (1.78% versus 0.5% per year)[32,33] suggesting that spinal cord atrophy is a sensitive and meaningful marker of neurodegeneration. [32] Spinal cord atrophy-related measures are typically calculated using semi-automated segmentation-based methods, which are subject to inter-rater variability. Future directions of research to fully automated analysis

methods, including segmentation of gray matter and intramedullary lesions will facilitate the use of spinal cord atrophy in the clinical and research arenas. [33]

Conclusion

Substantial evidence indicates that uncontrolled clinical and subclinical disease activity in early stages of RRMS may be critical for the evolution of long-term disability. The sequential addition of the individual components of NEDA-4 results in fewer patients achieving NEDA status at one and two years of follow-up. Brain atrophy is a good marker of disease progression in RRMS, and a-BVL is a parameter to continue investigating for guiding clinical practice.

The current diagnostic criteria and the follow-up tools of disease progression in RRMS lack relevance to the neurodegenerative aspects and concentrate mainly on the inflammatory process. BVL may be a cornerstone of measurement of neurodegenerative components of disease progression RRMS, which should lead to improvement in treatment strategies and patient outcomes. However, while current methods provide sufficient precision for cohort studies, they are not adequate for confidently assessing changes in individual patients. Advances in imaging and processing techniques are needed in order to enable neurologists to probe BVL along with the clinical endpoints in RRMS, and ultimately to improve treatment.[34]

References

1. Banwell B, Giovannoni G, Hawkes C, Lublin F. Editors' welcome and a working definition for a multiple sclerosis cure. *Mult Scler Relat Disord* 2013; **2**(2): 65-7.
2. Hegen H, Bsteh G, Berger T. 'No evidence of disease activity' - is it an appropriate surrogate in multiple sclerosis? *European journal of neurology* 2018; **25**(9): 1107-e101.
3. Smith SM, Zhang Y, Jenkinson M, et al. Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. *NeuroImage* 2002; **17**(1): 479-89.
4. Smith SM, De Stefano N, Jenkinson M, Matthews PM. Normalized accurate measurement of longitudinal brain change. *J Comput Assist Tomogr* 2001; **25**(3): 466-75.
5. 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. *Multiple sclerosis (Houndmills, Basingstoke, England)* 2016; **22**(10): 1297-305.
6. Anderson VM, Bartlett JW, Fox NC, Fisniku L, Miller DH. Detecting treatment effects on brain atrophy in relapsing remitting multiple sclerosis: sample size estimates. *J Neurol* 2007; **254**(11): 1588-94.
7. De Stefano N, Stromillo ML, Giorgio A, et al. Establishing pathological cut-offs of brain atrophy rates in multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2016; **87**(1): 93-9.
8. Khan O, Williams MJ, Amezcua L, Javed A, Larsen KE, Smrtka JM. Multiple sclerosis in US minority populations: Clinical practice insights. *Neurology Clinical practice* 2015; **5**(2): 132-42.
9. Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann Neurol* 2011; **69**(2): 292-302.
10. Vanotti S, Smerbeck A, Benedict RH, Caceres F. A new assessment tool for patients with multiple sclerosis from Spanish-speaking countries: validation of the Brief International Cognitive Assessment for MS (BICAMS) in Argentina. *Clin Neuropsychol* 2016; **30**(7): 1023-31.
11. Benedict RH, Amato MP, Boringa J, et al. Brief International Cognitive Assessment for MS (BICAMS): international standards for validation. *BMC Neurol* 2012; **12**: 55.
12. Arnold DL, Li D, Hohol M, et al. Evolving role of MRI in optimizing the treatment of multiple sclerosis: Canadian Consensus recommendations. *Mult Scler J Exp Transl Clin* 2015; **1**: 2055217315589775.
13. Smith SM, Jenkinson M, Woolrich MW, et al. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* 2004; **23 Suppl 1**: S208–S19.
14. Schmidt S, Koehler J, Winterstein C, Schicklmaier P, Kallmann B. An epidemiological study on the course of disease and therapeutic considerations in relapsing-remitting multiple sclerosis patients receiving injectable first-line disease-modifying therapies in Germany (EPIDEM). *Therapeutic advances in neurological disorders* 2018; **11**: 1756285617749802.
15. Lu G, Beadnall HN, Barton J, Hardy TA, Wang C, Barnett MH. The evolution of "No Evidence of Disease Activity" in multiple sclerosis. *Mult Scler Relat Disord* 2018; **20**: 231-8.
16. Kappos L, Radue EW, O'Connor P, et al. A placebo-controlled trial of oral fingolimod in relapsing multiple sclerosis. *N Engl J Med* 2010; **362**(5): 387-401.
17. Geurts JJ, Calabrese M, Fisher E, Rudick RA. Measurement and clinical effect of grey matter pathology in multiple sclerosis. *Lancet Neurol* 2012; **11**(12): 1082-92.
18. Kremenchutzky M, Cottrell D, Rice G, et al. The natural history of multiple sclerosis: a geographically based study. 7. Progressive-relapsing and relapsing-progressive multiple sclerosis: a re-evaluation. *Brain : a journal of neurology* 1999; **122 (Pt 10)**: 1941-50.
19. De Stefano N, Airas L, Grigoriadis N, et al. Clinical relevance of brain volume measures in multiple sclerosis. *CNS Drugs* 2014; **28**(2): 147-56.
20. Horakova D, Dwyer MG, Havrdova E, et al. Gray matter atrophy and disability progression in patients with early relapsing-remitting multiple sclerosis: a 5-year longitudinal study. *J Neurol Sci* 2009; **282**(1-2): 112-9.

21. Rotstein DL, Healy BC, Malik MT, Chitnis T, Weiner HL. Evaluation of no evidence of disease activity in a 7-year longitudinal multiple sclerosis cohort. *JAMA Neurol* 2015; **72**(2): 152-8.
22. Sacca F, Costabile T, Carotenuto A, et al. The EDSS integration with the Brief International Cognitive Assessment for Multiple Sclerosis and orientation tests. *Multiple sclerosis (Houndmills, Basingstoke, England)* 2017; **23**(9): 1289-96.
23. Oliva CG, Martinez M, de Grazia J, et al. Early disease stage of multiple sclerosis in Chilean patients: cognitive impairment is a key domain and does not relate to MRI measures (P1.412). *Neurology* 2018; **90**(15 Supplement): DOI: 10.1016/j.clinph.2006.11.176
24. Bergsland N, Horakova D, Dwyer MG, et al. Gray matter atrophy patterns in multiple sclerosis: A 10-year source-based morphometry study. *Neuroimage Clin* 2018; **17**: 444-51.
25. Eshaghi A, Prados F, Brownlee WJ, et al. Deep gray matter volume loss drives disability worsening in multiple sclerosis. *Ann Neurol* 2018; **83**(2): 210-22.
26. Derakhshan M, Caramanos Z, Giacomini PS, et al. Evaluation of automated techniques for the quantification of grey matter atrophy in patients with multiple sclerosis. *NeuroImage* 2010; **52**(4): 1261-7.
27. Bartsch AJ, Homola G, Biller A, et al. Manifestations of early brain recovery associated with abstinence from alcoholism. *Brain : a journal of neurology* 2007; **130**(Pt 1): 36-47.
28. MacKenzie JD, Siddiqi F, Babb JS, et al. Brain atrophy in mild or moderate traumatic brain injury: a longitudinal quantitative analysis. *AJNR American journal of neuroradiology* 2002; **23**(9): 1509-15.
29. Enzinger C, Fazekas F, Matthews PM, et al. Risk factors for progression of brain atrophy in aging: six-year follow-up of normal subjects. *Neurology* 2005; **64**(10): 1704-11.
30. Kempton MJ, Ettinger U, Schmechtig A, et al. Effects of acute dehydration on brain morphology in healthy humans. *Hum Brain Mapp* 2009; **30**(1): 291-8.
31. Nakamura K, Brown RA, Narayanan S, Collins DL, Arnold DL. Diurnal fluctuations in brain volume: Statistical analyses of MRI from large populations. *NeuroImage* 2015; **118**: 126-32.
32. Tur C, Moccia M, Barkhof F, et al. Assessing treatment outcomes in multiple sclerosis trials and in the clinical setting. *Nature reviews Neurology* 2018; **14**(2): 75-93.
33. Moccia M, Ruggieri S, Ianniello A, Toosy A, Pozzilli C, Ciccarelli O. Advances in spinal cord imaging in multiple sclerosis. *Therapeutic advances in neurological disorders* 2019; **12**: 1756286419840593.
34. Alroughani R, Deleu D, El Salem K, et al. A regional consensus recommendation on brain atrophy as an outcome measure in multiple sclerosis. *BMC neurology* 2016; **16**(1): 240.
35. Nygaard GO, Celius EG, de Rodez Benavent SA, et al. A Longitudinal Study of Disability, Cognition and Gray Matter Atrophy in Early Multiple Sclerosis Patients According to Evidence of Disease Activity. *PloS one* 2015; **10**(8): e0135974.
36. Huhn K, Senger D, Utz KS, et al. No evidence of disease activity status over 3 years in a real-world cohort of relapsing remitting MS patients in Germany. *Multiple sclerosis and related disorders* 2019; **27**: 133-8.
37. Rojas JI, Sanchez F, Caro F, et al. Brain volume loss and no evidence of disease activity over 3years in multiple sclerosis patients under interferon beta 1a subcutaneous treatment. *Journal of clinical neuroscience : official journal of the Neurosurgical Society of Australasia* 2019; **59**: 175-8.
38. Perumal J, Fox RJ, Balabanov R, et al. Outcomes of natalizumab treatment within 3 years of relapsing-remitting multiple sclerosis diagnosis: a prespecified 2-year interim analysis of STRIVE. *BMC neurology* 2019; **19**(1): 116.
39. Weinstock-Guttman B, Medin J, Khan N, et al. Assessing 'No Evidence of Disease Activity' Status in Patients with Relapsing-Remitting Multiple Sclerosis Receiving Fingolimod in Routine Clinical Practice: A Retrospective Analysis of the Multiple Sclerosis Clinical and Magnetic Resonance Imaging Outcomes in the USA (MS-MRIUS) Study. *CNS drugs* 2018; **32**(1): 75-84.

40. Zivadinov R, Khan N, Korn JR, et al. No evidence of disease activity in patients receiving fingolimod at private or academic centers in clinical practice: a retrospective analysis of the multiple sclerosis, clinical, and magnetic resonance imaging outcomes in the USA (MS-MRIUS) study. *Current medical research and opinion* 2018; **34**(8): 1431-40.
41. Yokote H, Kamata T, Toru S, Sanjo N, Yokota T. Brain volume loss is present in Japanese multiple sclerosis patients with no evidence of disease activity. *Neurological sciences : official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology* 2018; **39**(10): 1713-6.
42. Kim H, Lee EJ, Kim SK, Kim KK, Lim YM. Efficacy and safety of alemtuzumab in Korean multiple sclerosis patients. *Multiple sclerosis and related disorders* 2019; **30**: 247-51.
43. Frau J, Coghe G, Lorefice L, Fenu G, Musu L, Cocco E. Efficacy and safety of alemtuzumab in a real-life cohort of patients with multiple sclerosis. *Journal of neurology* 2019; **266**(6): 1405-11.
44. Prosperini L, Annovazzi P, Boffa L, et al. No evidence of disease activity (NEDA-3) and disability improvement after alemtuzumab treatment for multiple sclerosis: a 36-month real-world study. *Journal of neurology* 2018; **265**(12): 2851-60.