

OPEN



## MAGNIMS consensus recommendations on the use of brain and spinal cord atrophy measures in clinical practice

Jaume Sastre-Garriga<sup>1</sup>✉, Deborah Pareto<sup>2</sup>, Marco Battaglini<sup>3</sup>, Maria A. Rocca<sup>4</sup>, Olga Ciccarelli<sup>5,6</sup>, Christian Enzinger<sup>7</sup>, Jens Wuerfel<sup>8</sup>, Maria P. Sormani<sup>9,10</sup>, Frederik Barkhof<sup>6,11,12</sup>, Tarek A. Yousry<sup>5,13</sup>, Nicola De Stefano<sup>3</sup>, Mar Tintoré<sup>1</sup>, Massimo Filippi<sup>4,14</sup>, Claudio Gasperini<sup>15</sup>, Ludwig Kappos<sup>16</sup>, Jordi Río<sup>1</sup>, Jette Frederiksen<sup>17</sup>, Jackie Palace<sup>18</sup>, Hugo Vrenken<sup>11</sup>, Xavier Montalban<sup>1,19</sup>, Àlex Rovira<sup>2</sup>✉ and on behalf of the MAGNIMS study group

**Abstract** | Early evaluation of treatment response and prediction of disease evolution are key issues in the management of people with multiple sclerosis (MS). In the past 20 years, MRI has become the most useful paraclinical tool in both situations and is used clinically to assess the inflammatory component of the disease, particularly the presence and evolution of focal lesions — the pathological hallmark of MS. However, diffuse neurodegenerative processes that are at least partly independent of inflammatory mechanisms can develop early in people with MS and are closely related to disability. The effects of these neurodegenerative processes at a macroscopic level can be quantified by estimation of brain and spinal cord atrophy with MRI. MRI measurements of atrophy in MS have also been proposed as a complementary approach to lesion assessment to facilitate the prediction of clinical outcomes and to assess treatment responses. In this Consensus statement, the Magnetic Resonance Imaging in MS (MAGNIMS) study group critically review the application of brain and spinal cord atrophy in clinical practice in the management of MS, considering the role of atrophy measures in prognosis and treatment monitoring and the barriers to clinical use of these measures. On the basis of this review, the group makes consensus statements and recommendations for future research.

The inflammatory component of multiple sclerosis (MS) pathology can be focal or diffuse and is associated with neurodegenerative processes that ultimately lead to irreversible tissue damage and neuronal loss<sup>1</sup>. Neurodegeneration was originally thought to be a late-stage phenomenon with limited clinical relevance, but it is now recognized as being associated with acute inflammation from the early stages of MS and as the main driver of irreversible disability<sup>2–5</sup>. In parallel with improvements in our understanding of the mechanisms of neurodegeneration, advances in imaging techniques have enabled in vivo assessment of brain and spinal cord area and volumes using MRI. Although brain and spinal cord volume loss observed with MRI cannot be equated with atrophy<sup>6</sup>, because the latter implies pathologically proven and irreversible tissue loss, changes in these MRI measures are associated with atrophy<sup>7</sup> and the level of disability in MS<sup>8,9</sup>.

MRI-based quantification of inflammatory activity in MS — on the basis of lesion counts and lesion volumes — is established as the main efficacy outcome in phase II clinical trials<sup>10</sup>. Currently, brain and spinal cord volume measures have no role in the MS diagnostic criteria<sup>11,12</sup> or disease course classification<sup>13</sup>, but a body of evidence that these measures are valuable for early evaluation of treatment responses and prediction of disease evolution has been steadily growing alongside improvements in methodology that could facilitate widespread implementation of these measures in clinical practice<sup>14,15</sup>. A key difficulty arises in this implementation because translation of group-based results into actionable, patient-level information must be made with extreme caution.

In this Consensus Statement, we, on behalf of the **Magnetic Resonance in Imaging in MS (MAGNIMS) study group**, provide specific recommendations for the

✉e-mail: [jsastre-garriga@cem-cat.org](mailto:jsastre-garriga@cem-cat.org); [alex.rovira.idi@gencat.cat](mailto:alex.rovira.idi@gencat.cat)  
<https://doi.org/10.1038/s41582-020-0314-x>

implementation of brain and spinal cord atrophy measures in the clinical management of patients with MS and on the directions of future research to improve our knowledge in this field. The recommendations are based on a critical review of the literature and the personal experience of MAGNIMS study group members. We discuss the difficulties of translating group-based data into clinical application and highlight where particular caution is appropriate. We first discuss the role of atrophy measures on prognosis, then treatment monitoring and, finally, the barriers to implementation in clinical practice. Each of these three sections comprises a review of the available evidence and a set of consensus guidelines.

## Methods

A multicentre international panel on the implementation of brain and spinal cord atrophy measures in clinical practice convened in Barcelona, Spain, under the auspices of MAGNIMS, an independent European network of clinical research groups with a common interest in the study of MS with MRI. The panel was made up of experts in the diagnosis and management of MS, including neuroradiologists, neurologists, physicists, imaging methodologists and statisticians, who were selected by the workshop organizers (with approval from all members of the Steering Committee) on the basis of their personal expertise, from MAGNIMS centres from seven different countries. The purpose of this face-to-face

meeting was to review and discuss all published data on brain and spinal cord atrophy in MS and to consider whether the previously published recommendations<sup>16,17</sup> on its use for diagnosis, prognosis and monitoring of patients with MS needed to be revised and updated in view of technical advances and numerous clinical studies of atrophy in MS. The panel agreed that updated recommendations were necessary. After this meeting, the panel members formulated specific recommendations in relation to the implementation of brain and spinal cord atrophy measures in clinical practice.

The authors of the Consensus statement are members of the MAGNIMS Study Group. The network is independent of any other organization and, at the time of the workshop mentioned above, was run by a Steering Committee whose members were Å. Rovira (Barcelona, co-chair), C. Enzinger (Graz, co-chair), F. Barkhof (Amsterdam), O. Ciccarelli (London), N. de Stefano (Siena), M. Filippi (Milan), J. Frederiksen (Copenhagen), C. Gasperini (Rome), L. Kappos (Basel), J. Palace (Oxford), M.A. Rocca (Milan), J. Sastre-Garriga (Barcelona), H. Vrenken (Amsterdam) and T. Yousry (London). The first draft of the recommendations was written by the principal authors (J.S.-G., D.P. and M.B.) on the basis of the panellists' presentations and contributions to discussions on specific topics, which were assigned to individuals according to each member's area of expertise. The initial draft was then circulated among all authors (who were all presenters and/or discussants at the meeting). Modifications were made iteratively until consensus was reached on all recommendations; all panel members agreed on the full contents of the final recommendations.

## Author addresses

<sup>1</sup>Multiple Sclerosis Centre of Catalonia (Cemcat), Department of Neurology/Neuroimmunology, Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona, Barcelona, Spain.

<sup>2</sup>Section of Neuroradiology and Magnetic Resonance Unit, Department of Radiology, Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona, Barcelona, Spain.

<sup>3</sup>Department of Medicine, Surgery and Neuroscience, University of Siena, Siena, Italy.

<sup>4</sup>Neuroimaging Research Unit, Institute of Experimental Neurology, Division of Neuroscience, IRCCS San Raffaele Scientific Institute, Milan, Italy.

<sup>5</sup>NMR Research Unit, University College London Queen Square Institute of Neurology, London, UK.

<sup>6</sup>National Institute for Health Research Biomedical Research Centre, University College London Hospitals, London, UK.

<sup>7</sup>Department of Neurology and Division of Neuroradiology, Vascular and Interventional Radiology, Department of Radiology, Medical University of Graz, Graz, Austria.

<sup>8</sup>Medical Image Analysis Center (MIAC AG) and Department of Biomedical Engineering, University of Basel, Basel, Switzerland.

<sup>9</sup>Biostatistics Unit, Department of Health Sciences, University of Genoa, Genoa, Italy.

<sup>10</sup>IRCCS, Ospedale Policlinico San Martino, Genoa, Italy.

<sup>11</sup>Amsterdam Neuroscience, MS Center Amsterdam, Department of Radiology and Nuclear Medicine, Amsterdam UMC, Amsterdam, Netherlands.

<sup>12</sup>Institutes of Neurology and Healthcare Engineering, University College London, London, UK.

<sup>13</sup>Lysholm Department of Neuroradiology, University College London Hospitals National Hospital for Neurology and Neurosurgery, University College London Institute of Neurology, London, UK.

<sup>14</sup>Vita-Salute San Raffaele University, Milan, Italy.

<sup>15</sup>Multiple Sclerosis Center, Department of Neurosciences, San Camillo-Forlanini Hospital, Rome, Italy.

<sup>16</sup>Neurologic Clinic and Policlinic, Departments of Medicine, Clinical Research and Biomedical Engineering, University Hospital, University of Basel, Basel, Switzerland.

<sup>17</sup>Department of Neurology, Rigshospitalet-Glostrup and University of Copenhagen, Glostrup, Denmark.

<sup>18</sup>Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK.

<sup>19</sup>Division of Neurology, St Michael's Hospital, University of Toronto, Toronto, Canada.

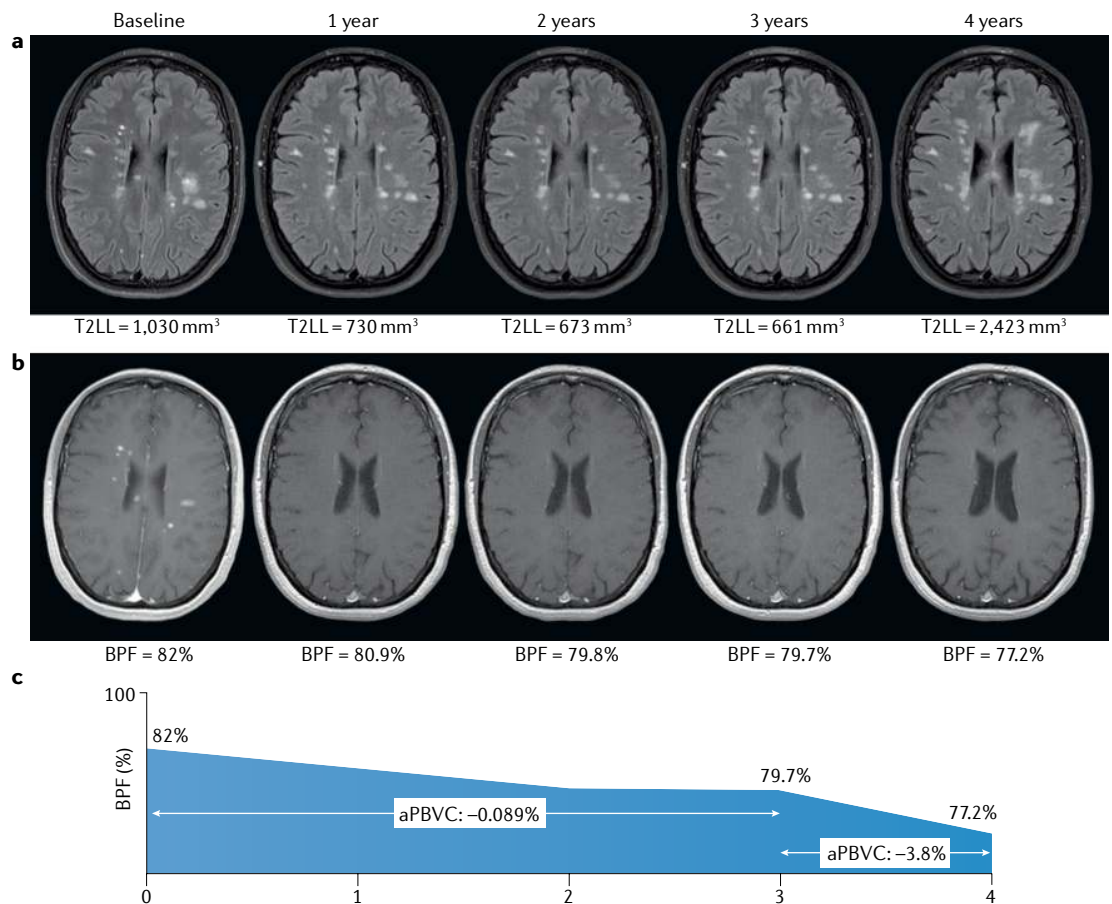
## Defining and predicting MS severity

### Evidence review

**Global brain volume measures to define and predict MS severity.** The initial studies to investigate clinical correlates of brain atrophy in MS focused on patients with well-established disease and severe clinical manifestations, particularly in the cognitive sphere<sup>18–20</sup>, but later studies included disability, as measured with the Expanded Disability Status Scale<sup>8</sup>. Evidence from these studies made it clear that neurodegenerative processes occur in the earliest phases of MS<sup>21</sup>, even before the disease becomes symptomatic<sup>22</sup>.

Yearly global brain volume loss in healthy ageing individuals ranges from  $-0.05\%$  at 20–30 years of age to  $-0.3\%$  at 60–70 years of age<sup>23</sup>. A change of  $-0.4\%$  per year has been proposed as the cut-off for pathological brain atrophy in MS<sup>24</sup> (FIG. 1), although care must be taken before applying this threshold as a marker of therapeutic efficacy owing to the phenomenon of pseudo-atrophy (see Brain volume as an outcome measure in randomized clinical trials)<sup>25,26</sup>. Multiple studies have shown that short-term changes (over as little as 1 year) in brain volume are predictive of clinical status (diagnosis of MS or disability status) at various follow-up times in clinically isolated syndromes<sup>27,28</sup>, relapsing–remitting MS (RRMS)<sup>29</sup> and primary progressive MS<sup>30–32</sup>, either in isolation or together with lesion-related parameters<sup>33,34</sup>.

The findings above are group-based results, and translation of these findings to the individual level is not



**Fig. 1 | Lesion load and brain atrophy in relapsing–remitting multiple sclerosis. a** | Transverse T2-weighted fluid attenuation inversion recovery images from a patient with highly active relapsing–remitting multiple sclerosis (MS) who started a disease-modifying therapy at baseline. The T2 lesion load (T2LL) is stable during the first 3 years of treatment while the patient remained clinically stable (no relapses and no disability worsening), but markedly increases at the fourth year after treatment discontinuation associated with clinical activity (the rebound effect). **b** | Contrast-enhanced T1-weighted images from the same patient showing the change in brain parenchymal fraction (BPF) over time. The decrease in global brain volume in the first 3 years is mild (annualized percentage of brain volume change (aPBVC)  $-0.089\%$ ), but the volume loss at the fourth year is severe (aPBVC  $-3.8\%$ ), matching the change in T2LL and clinical evolution. The severe loss observed in year 4 is well beyond the  $-0.4\%$  suggested as a pathological cut-off for brain volume loss in MS<sup>24</sup>. **c** | Graphical representation of the changes in BPF over time, emphasizing the dramatic loss of volume in year 4.

straightforward. In a study published in 2017, Sormani et al.<sup>35</sup> made the first attempt to define individual cut-off values for brain volume changes according to patients' baseline characteristics. Pooled baseline data from the placebo arms of two large international clinical trials that involved a total of 2,342 patients with RRMS showed that expected normalized brain volumes can be calculated from demographic (age and sex), clinical (Expanded Disability Status Scale score and disease duration) and neuroradiological (T2-weighted lesion volume) parameters for individuals. Deviation of the true brain volume from this expected value enabled classification of individuals with MS as having low, medium or high brain volume. Patients with low brain volume had a 2.4-fold higher risk of disability progression over the next 2 years than patients with high brain volume.

**Spinal cord atrophy measures to define and predict MS severity.** Early, seminal studies of cervical cord atrophy in MS already suggested that cervical spinal cord area is

an important marker of disability status in MS<sup>9</sup>. Further studies demonstrated that spinal cord area and volume are affected differently in different MS subtypes, with the most profound atrophy in cross-sectional studies being seen in patients with progressive MS<sup>36–39</sup>. Since 2015, an association between reduced cervical cord area and increased disability and motor dysfunction, independent of brain atrophy, has been confirmed<sup>40–43</sup>. An association between cord atrophy and reduced peripapillary retinal nerve fibre layer thickness has been identified, indicating that cervical cord atrophy reflects, at least in part, global pathological processes and not only specific damage of long tracts<sup>41</sup>. Most studies of spinal cord area have focused on global cervical cord area measurements, but some work has highlighted that damage in particular locations in the spinal cord, such as cervical grey matter<sup>44</sup>, the thoracolumbar segment<sup>45</sup> and the posterior and lateral cord segments<sup>46</sup>, are also relevant to disability.

Longitudinal studies indicate that atrophy rates in the spinal cord are higher than those in the brain and higher

in progressive MS than in established RRMS<sup>47,48</sup>. Higher rates of cervical cord area loss have been associated with disability progression, independent of other clinical and MRI parameters<sup>30,47</sup> including spinal cord lesions<sup>49</sup>. However, as for brain atrophy, use of such group-level evidence to inform clinical decisions at the individual level is not easy. Results that can be used at the individual level are slowly emerging; for example, Tsagkas et al.<sup>43</sup> have shown that a 1% increase in the annual rate of spinal cord atrophy increases the risk of disability progression by 28%, reinforcing the notion that spinal cord atrophy is a reliable and independent tool for monitoring disease progression.

**Regional and tissue-specific brain volumetry measures to define and predict MS severity.** Early cross-sectional studies of brain white matter and grey matter changes in patients with MS indicated that both white matter and grey matter loss occurred early in the disease course, regardless of disease phenotype<sup>50–53</sup>. Evidence also indicates that grey matter damage can occur before white matter atrophy and can occur independently of white matter lesions<sup>54–56</sup>. Results of further longitudinal studies have identified larger decreases in grey matter volumes than in white matter volumes<sup>57–59</sup> and that grey matter damage is more relevant than white matter injury to clinical outcomes, both concurrent and forthcoming<sup>56,60–62</sup>. Two studies — one in which cortical thickness was estimated<sup>63</sup> and one meta-analysis of voxel-based morphometry studies<sup>64</sup> — have revealed statistically significant associations between disability end points and grey matter atrophy<sup>65</sup>, which occurs bilaterally, predominantly in the cingulate, pre-central and/or post-central gyri and the thalami and basal ganglia. Despite these results, global brain volume changes seem to be more strongly associated with clinical outcomes than are regional changes. This observation is unexpected because grey matter loss is thought to underlie disability accumulation. Associations between grey matter volume change and disability accumulation might be masked by the effects of high variability of regional segmentations, which makes clinical application of these regional measures inadvisable at present<sup>62,66</sup>.

### Statements and recommendations

1. We recommend measurement of global brain volume to better gauge global disease burden in patients with MS because brain volume loss is associated with and predicts disability in all clinical MS phenotypes, including the earliest stages of the condition.
2. We recommend measurement of cervical cord area loss because this measure is associated with and predicts disability in all clinical MS phenotypes, including the earliest stages of the condition.
3. Grey matter volume changes in the brain are more pronounced and clinically relevant than white matter volume changes, even in the earliest stages of MS, but their exact relevance in clinical practice is unclear. We recommend further research to clarify this relevance.
4. Some cerebral grey matter regions (including the thalami, basal ganglia and specific cortical areas) are affected particularly strongly by atrophy in MS, but

whether the pathological involvement of these areas is relevant in clinical practice remains unclear. We recommend further research to determine the clinical relevance of atrophy in these regions.

### Monitoring therapeutic effect

#### Evidence review

**Brain volume as an outcome measure in randomized clinical trials.** Many trials of disease-modifying therapies for MS have included brain atrophy as an outcome measure (TABLE 1). Most early studies of interferon- $\beta$  (IFN $\beta$ ) and glatiramer acetate did not include preplanned brain volume measures as secondary MRI outcomes. Those that did include a sound comparison of brain volume changes between intervention arms or between intervention and placebo arms produced mixed results<sup>67</sup>.

The only study of IFN $\beta$  that provided evidence for a positive effect of treatment of brain atrophy was the ETOMS trial<sup>68</sup>. In this study, accrual of atrophy was reduced by 30% in patients with clinically isolated syndromes who received low-dose subcutaneous IFN $\beta$ 1a compared with patients who received placebo<sup>68</sup>. In several trials — particularly the trial of intramuscular IFN $\beta$ 1a in RRMS<sup>69,70</sup> — negative results were at least partly attributed to a pseudoatrophy effect, caused by brain volume loss linked to the presumed treatment-associated resolution of inflammatory activity and oedema. In the RRMS intramuscular IFN $\beta$ 1a trial, significant differences that favoured treatment with IFN $\beta$ 1a were only observed in the second year<sup>69,70</sup>. A post-hoc analysis of grey matter and white matter atrophy during the 2 years of the trial confirmed this finding and indicated that pseudoatrophy of white matter contributed most to the observed effect<sup>71</sup>. The same effect has been described in observational studies of patients taking natalizumab<sup>72</sup> or IFN $\beta$ <sup>73</sup>, although more research is needed to confirm these findings. Results with glatiramer acetate were also mixed, though some nonprimary analyses have suggested a positive effect of the treatment in patients who received glatiramer acetate from the beginning of the trial when compared with those who received the treatment later<sup>74</sup>. Trials of IFN $\beta$  and of glatiramer acetate in progressive MS have been negative<sup>75</sup> or have also suggested a pseudoatrophy effect<sup>76</sup>.

Trials of natalizumab provided a clear demonstration of pseudoatrophy. In the AFFIRM trial<sup>77</sup>, brain volume decreases among patients who received natalizumab were larger in the first year than among patients who received placebo, but the observation was reversed in the second year. Subsequent clinical trials of newer drugs (including fingolimod, dimethyl fumarate, teriflunomide, ocrelizumab and alemtuzumab) have all incorporated brain volume measures as secondary or tertiary outcomes, and results have been positive overall<sup>78</sup>, although studies are not readily comparable. Of note, in studies of powerful anti-inflammatory drugs against active comparators, the trial drugs have been superior at decreasing accrual of atrophy<sup>79–81</sup>, indicating that the pseudoatrophy effect can be overcome by the beneficial effects of anti-inflammatory drugs on neurodegeneration in MS. Strategies to minimize the effect of

Table 1 | Brain atrophy outcomes in pivotal trials of approved disease-modifying drugs

Drug	Clinical trial	Phenotype	Comparator	Time frame	Software	Treatment favoured
IFNβ1a (SC)	ETOMS <sup>68</sup>	CIS	Placebo	0–24 months	SIENA <sup>113</sup>	IFNβ1a (SC)
	REFLEX <sup>155</sup>	CIS	Placebo	0–24 months	SIENA <sup>113</sup>	None
	PRISMS <sup>156</sup>	RRMS	Placebo <sup>a</sup>	0–6 years	Kappos et al. <sup>156</sup>	None
IFNβ1b (SC)	BENEFIT <sup>157</sup>	CIS	Placebo <sup>a</sup>	0–36 months	SIENA <sup>113</sup>	None
	EUSPMS <sup>158</sup>	SPMS	Placebo	0–36 months	Losseff et al. <sup>8</sup>	None
	Montalban et al. <sup>76</sup>	PPMS	Placebo	0–24 months	SPM <sup>51</sup>	None
IFNβ1a (IM)	Rudick et al. <sup>69</sup>	RRMS	Placebo	0–24 months	Rudick et al. <sup>70</sup>	None
	Leary et al. <sup>90</sup>	PPMS	Placebo	0–24 months	Fox et al. <sup>140</sup>	None
Glatiramer acetate	PRECISE <sup>159</sup>	CIS	Placebo	0–36 months	SIENA <sup>113</sup>	None
	Sormani et al. <sup>160</sup>	RRMS	Placebo <sup>a</sup>	0–18 months	SIENA <sup>113</sup>	Glatiramer acetate
	PROMISE <sup>75</sup>	PPMS	Placebo	0–36 months	Wolinsky et al. <sup>75</sup>	None
Teriflunomide	TOPIC <sup>161</sup>	CIS	Placebo	0–24 months	Miller et al. <sup>161</sup>	None
	TEMSO <sup>162</sup>	RRMS	Placebo	0–24 months	Wolinsky et al. <sup>163</sup>	None
	TEMSO <sup>164,b</sup>	RRMS	Placebo	0–24 months	SIENA <sup>113</sup>	Teriflunomide
Dimethyl fumarate	DEFINE <sup>82</sup>	RRMS	Placebo	6–24 months	SIENA <sup>113</sup>	Dimethyl fumarate
	CONFIRM <sup>165</sup>	RRMS	Placebo	0–24 months	SIENA <sup>113</sup>	None
Natalizumab	AFFIRM <sup>77</sup>	RRMS	Placebo	0–24 months	Rudick et al. <sup>70</sup>	None <sup>d</sup>
	SENTINEL <sup>166</sup>	RRMS	Placebo <sup>c</sup>	0–24 months	Rudick et al. <sup>70</sup>	None <sup>d</sup>
	ASCEND <sup>167</sup>	SPMS	Placebo	24–96 weeks	SIENAX <sup>113</sup>	None
Fingolimod	FREEDOMS 1 (REF. <sup>168</sup> )	RRMS	Placebo	0–24 months	SIENA <sup>113</sup>	Fingolimod
	FREEDOMS 2 (REF. <sup>169</sup> )	RRMS	Placebo	0–24 months	SIENA <sup>113</sup>	Fingolimod
	TRANSFORMS <sup>81</sup>	RRMS	IFNβ1a (IM)	0–12 months	SIENA <sup>113</sup>	Fingolimod
	INFORMS <sup>91</sup>	PPMS	Placebo	0–36/60 months	SIENA <sup>113</sup>	None
Alemtuzumab	CARE-MS 1 (REF. <sup>79</sup> )	RRMS	IFNβ1a (SC)	0–24 months	Rudick et al. <sup>70</sup>	Alemtuzumab
	CARE-MS 2 (REF. <sup>80</sup> )	RRMS	IFNβ1a (SC)	0–24 months	Rudick et al. <sup>70</sup>	Alemtuzumab
Ocrelizumab	OPERA 1 (REF. <sup>170</sup> )	RRMS	IFNβ1a (SC)	24–96 weeks	SIENA <sup>113</sup>	Ocrelizumab
	OPERA 2 (REF. <sup>170</sup> )	RRMS	IFNβ1a (SC)	24–96 weeks	SIENA <sup>113</sup>	Ocrelizumab
	ORATORIO <sup>171</sup>	PPMS	Placebo	24–120 weeks	SIENA <sup>113</sup>	Ocrelizumab
Cladribine	ORACLE <sup>172</sup>	CIS	Placebo	0–24 months	SIENA <sup>113</sup>	None
	CLARITY <sup>83</sup>	RRMS	Placebo	6–24 months	SIENA <sup>113</sup>	Cladribine

CIS, clinically isolated syndrome; IFN, interferon; IM, intramuscular; PPMS, primary progressive multiple sclerosis; RRMS, relapsing–remitting multiple sclerosis; SC, subcutaneous; SIENA, Structural Image Evaluation, using Normalization, of Atrophy; SPMS, secondary progressive multiple sclerosis. <sup>a</sup>Includes a period receiving the active drug. <sup>b</sup>Reanalysis of TEMSO trial data using SIENA. <sup>c</sup>As an add-on to IFNβ1a (IM). <sup>d</sup>Results favoured natalizumab in the 12–24-month period.

pseudoatrophy on clinical measures include, but are not restricted to, obtaining baseline measurements once the anti-inflammatory effect is well established (for example, re-baseline with MRI at 6 or 12 months after treatment initiation)<sup>82,83</sup>.

Further support for the clinical relevance of brain volume outcomes in trials of treatment for RRMS comes from a meta-analysis that included >13,500 patients from 13 different clinical trials<sup>84</sup>. The conclusion of the analysis was that the effect of a given therapy on changes in brain volume over 2 years is associated with the effect of the drug on disability outcomes and that this association is, at least in part, independent of its anti-inflammatory effect on active MRI lesions<sup>84</sup>. This close association between brain atrophy and disability outcomes in clinical

trials has driven the adoption of brain volume change as a primary outcome in phase II trials in cohorts of patients with progressive MS<sup>85,86</sup>.

**Spinal cord atrophy as an outcome in randomized clinical trials.** Despite the relevance of spinal cord atrophy to long-term disability, this measure has scarcely been used as an outcome in clinical trials<sup>87</sup>; when it has been used, the results have been negative. For example, spinal cord atrophy was an outcome measure in an investigator-initiated study of lamotrigine for neuroprotection in secondary progressive MS, but no differences were seen between the treatment and placebo arms<sup>88</sup>. Spinal cord atrophy measures have been used in several other studies in progressive MS<sup>89</sup> but the atrophy and clinical results

have either been negative<sup>76,90</sup> or were not published with the rest of the trial<sup>91</sup>.

**Brain volume and spinal cord atrophy to monitor clinical treatment response.** The relevance of brain volume measures to the evolution of disability in MS clinical trials is beyond any doubt<sup>84</sup>. The evidence from trials is complemented by that from studies of individual-level data from clinical trials<sup>92,93</sup> and from observational studies of real-world cohorts<sup>25,94</sup>, which confirm a close association between brain volume changes with therapy and concurrent<sup>95</sup> or subsequent<sup>96</sup> disability progression. These studies also indicate that the association between brain volume loss and disability progression is independent of clinical and MRI inflammatory markers.

Most models for the prediction of disability progression have included brain volume change combined with either the appearance of new T2 lesions or the presence of clinical relapses<sup>25,92–94</sup>. Brain volume changes have also been proposed as an addition to the ‘no evidence of disease activity’<sup>97,98</sup> outcome measure so as to enable assessment of neurodegenerative processes as well as inflammatory processes, with the aim of achieving full remission that includes an absence of disease-specific neurodegeneration; the proposed cut-off for this measure is  $-0.4\%$  change in volume per year<sup>24</sup>. In a potentially more realistic ‘minimal evidence of disease activity’ approach<sup>99</sup>, a less stringent cut-off has been suggested that would allow for pseudoatrophy-driven brain volume loss<sup>25</sup>. However, all these data need confirmation, and different cut-offs might be needed for different calculation methods and for different drugs or groups of drugs according to different temporal patterns of brain volume effects of each drug<sup>6,78</sup>.

### Statements and recommendations

1. We recommend the use of whole brain atrophy over a minimum period of 12 months as a secondary end point in clinical trials in MS and even as a primary outcome measure in trials in the progressive forms of MS to show the effects of the drug on the neurodegenerative component of the disease.
2. Ongoing and forthcoming trials are expected to include grey matter volume loss as an outcome measure, as atrophy in the grey matter compartment is more substantial and more clinically relevant than atrophy in the white matter and is likely to be affected less by pseudoatrophy; however, data on pseudoatrophy remain discordant and we recommend further research to clarify the contribution of grey matter atrophy.
3. Pseudoatrophy effects mostly occur within the first 6–12 months from treatment initiation with any anti-inflammatory therapy, so we recommend re-baseline MRI at 6–12 months after initiation of any therapy to mitigate the impact of pseudoatrophy on outcome measures.
4. Associations between treatment effects on brain volume and disability have been demonstrated in clinical trials and indicated by evidence at the individual level, but we recommend further research to

confirm these associations before brain volume can be considered for use as a treatment-monitoring tool.

5. Use of spinal cord atrophy as a treatment-monitoring tool in clinical trials and in clinical practice has been scarce, but the rate of spinal cord atrophy is faster than that of brain atrophy and methodological advances could improve reproducibility and reliability, so we recommend further research to establish the role of spinal cord atrophy for treatment monitoring.

### Barriers to clinical implementation

#### Evidence review: technical barriers

Several technical aspects of image acquisition and quantification can affect the measurement of brain and spinal cord volumes and thereby affect the accuracy of estimated values. These technical barriers are discussed below.

**Acquisition protocols.** The choice of the acquisition parameters (usually repetition time, echo time, inversion time or flip angle) is usually based on the image contrast, as assessed visually by an expert neuroradiologist. Changes in scan parameters, which tend to happen in a clinical environment, affect quantification and hamper reliable cross-sectional and longitudinal comparisons. Image contrast also depends greatly on the age of the population that undergoes MRI. The Alzheimer’s Disease Neuroimaging Initiative<sup>100</sup> has made a large effort to homogenize acquisition protocols across vendors.

**Gradient distortion.** By design, the gradients applied to the magnetic field in MRI are generally not uniform, which affects the geometry of the image. Small displacements of the patient’s head in the  $z$  axis have a notable effect on the estimated brain volume change<sup>101</sup>. Positioning of the patient identically across scanning sessions can minimize this effect, but this is time-consuming and difficult; a better solution is to apply approaches developed by MRI scanner manufacturers for 3D correction for the gradient nonlinearity effect<sup>102</sup>.

**Intrascanner variability.** Any MRI-derived measure is inherently variable, even when technical and physiological conditions are controlled<sup>103–108</sup>. Global estimates, such as that of the whole brain volume, are the least variable ( $<1\%$ )<sup>106</sup>, whereas measures of smaller structures, such as the amygdala, are much more variable ( $\sim 5\%$ )<sup>104,105</sup>. Such variability must be taken into account because changes that are smaller than the estimated variability cannot be reliably detected. This limitation is highly relevant to small grey matter structures and when follow-up periods are short because the expected change is small<sup>23</sup>.

**Movement.** Movement of the patient during image acquisition generates characteristic artefacts that affect image quality; as a result, estimated volumes are substantially decreased<sup>109</sup>. Visual verification of image quality is important because the problem is resolved when the only images included in an analysis are those that an expert considers artefact free<sup>109</sup>. Various approaches have been developed to correct for movement, but an accurate method is still not available<sup>110</sup>.

**Scanner system upgrades and interscanner variability.** Scanner upgrades are unavoidable, particularly during the course of longitudinal studies, and can affect the image contrast even if the same acquisition parameters are used. Previous studies have shown that the system upgrade should be included as a variable in the statistical analysis<sup>103,111,112</sup>. Quantification methods based on the subtraction of images, rather than on differences in brain parenchymal fraction between two time points, seem to be more sensitive to system upgrades<sup>113</sup>, although no studies have been performed to confirm this observation. Reliable quantification of longitudinal changes in MS requires scans to be acquired with the same magnet and exactly the same sequence protocol. Variability between different scanners is higher than all the factors above together<sup>108</sup>. If data acquired in different scanners need to be merged, a variable that accounts for the scanner should be taken into consideration.

**Evidence review: confounding factors**

Numerous factors can have confounding effects on the quantification of brain volume (and its changes) and thereby cause overestimation or underestimation<sup>114</sup>. These factors are discussed below.

**Age, sex and brain size.** Several physiological factors influence brain volume estimations in healthy individuals. Studies of healthy elderly individuals have demonstrated ongoing brain volume loss, which tends to accelerate with age<sup>115</sup>. This age-related effect is particularly pronounced for specific CNS structures, such as the hippocampus<sup>116</sup>.

Sex is another key factor in brain volume changes. Sex differences in global brain size in humans are well established; on average, the total volume of men's brains is ~10% larger than that of women's brains<sup>117</sup>. Differential patterns of age-related brain volume loss<sup>118</sup> and sex-specific differences in brain morphology have also been demonstrated<sup>119,120</sup>. Global and regional volumetric studies have suggested that hormonal status can contribute to these sex-related differences<sup>121</sup>.

**Diurnal fluctuations and hydration state.** Studies of healthy individuals have shown that estimations of brain volume fluctuate with the time of scanning and the hydration state of the individuals. Analysis of MRI data from patients with MS ( $n = 755$ , 3,269 scans) and from participants in the Alzheimer's Disease Neuroimaging Initiative ( $n = 834$ , 6,114 scans) revealed that time of day had a notable effect on estimates of the brain parenchymal fraction in both groups. Brain volumes were substantially larger in the morning<sup>122</sup>, and the effect size was comparable to the yearly rate of brain atrophy in MS and in healthy elderly people<sup>122</sup>. Similarly, in studies in which hydration status was manipulated by overnight thirsting and subsequent drinking of water, hydration-related changes in brain volume were as large as -0.55% on dehydration and +0.72% on rehydration<sup>123</sup>.

**Lifestyle and risk factors.** Many lifestyle factors, including physical activity<sup>124</sup>, influence estimates of brain volume. A higher level of alcohol intake has been associated

with a higher rate of brain atrophy over a 6-year period<sup>115</sup> and with a specific pattern of regional involvement of the white matter and grey matter<sup>125</sup>. A similar effect has been described for cigarette smoking and substance abuse (for example, marijuana use)<sup>115,126</sup>. Many systemic conditions, such as diabetes, chronic kidney disease, hypertension, obesity and vascular conditions can also accelerate brain atrophy<sup>115,127,128</sup>.

**The MS brain.** All confounding factors previously discussed can interact with features of MS and affect estimates of brain atrophy in patients with the disease; these interactions can also affect comparisons between groups. For instance, more severe brain atrophy has been observed in patients with MS who have one or more cardiovascular risk factors<sup>129</sup>, although their impact on longitudinal assessments might be limited, as vascular risk factors were not associated with greater brain volume loss during 3.5 years of follow-up in the same study<sup>129</sup>. In addition, white matter lesions in MS influence the accuracy of most available software for estimation of atrophy because they alter the image intensity histogram and influence the detection of intensity borders between grey matter, white matter and cerebrospinal fluid (CSF). This effect can be minimized by use of lesion filling techniques<sup>130,131</sup>, which enable replacement of lesions in the image with voxels that have intensities that closely resemble normal-appearing white matter.

**Pseudoatrophy.** As discussed above, studies of the correlation between inflammatory disease activity (new T2 and/or gadolinium enhancing lesions) and brain volume have shown that inflammation can cause a transient increase in brain volume. This increase can dramatically resolve following treatment with steroids<sup>132</sup> or other disease-modifying drugs, and the resultant reduction in brain volume can be erroneously interpreted as atrophy<sup>133</sup>.

**Evidence review: volumetry tools**

Several free-to-use online libraries of software for neuroimaging analyses include fully automated pipelines for quantification of brain volume (TABLE 2). On the basis of the current literature that relates to this software, these software tools can be classified into two broad categories. The first are 'segmentation-based' tools, which use a priori localization-related and intensity information to classify the brain voxels of each MRI without using information from brain MRI images taken at different time points. These tools do not enable direct evaluation of volumetric changes over time. This type of software is mostly used in cross-sectional analyses. The second are the 'registration-based' tools, which enable comparison of brain MRI images from the same individual acquired over time and are based on an initial registration step; this type of software is used in longitudinal analyses<sup>134</sup>.

Most segmentation-based software packages provide measures of total brain volume, grey matter volume and white matter volume based on the partial volume estimation (PVE) of each tissue in each voxel. The initial step is assignment of the PVE to a given brain voxel on the basis of its intensity and the intensities of the surrounding

**Brain parenchymal fraction**  
The percentage of the intracranial volume that is occupied by brain parenchyma, calculated as the total intracranial volume minus the volume of cerebrospinal fluid divided by the total intracranial volume.

**Partial volume estimation (PVE).** The proportions of white matter, grey matter and cerebrospinal fluid present in each voxel of an MRI image.

Table 2 | Available brain and spinal cord volumetry tools

Tool	Freely available?	Measures	Major limitations <sup>a</sup>
SIENAX	Yes	Global and regional brain volumes for cross-sectional comparisons	Segmentations are affected by the presence of brain lesions
SPM/VBM	Yes <sup>b</sup>	Global and regional brain volumes, pixel-to-pixel statistical comparisons between two groups or time points	Segmentations are affected by the presence of brain lesions
GIF	Yes	Regional brain volumes for cross-sectional comparisons	Time consuming; data analysed remotely
Atropos	Yes	Regional brain volumes for cross-sectional and longitudinal comparisons	Limited information about the method as it has not been used extensively
FreeSurfer	Yes	Cortical thickness, global and regional grey matter and white matter volumes	Time-consuming; requires manual correction of the segmented surfaces; segmentations are affected by the presence of lesions
CIVET	No	Cortical thickness	Software not freely available
SIENA	Yes	Percentage brain volume change between two time points	Only provides global measures that include grey matter and white matter
SIENA-XL	No	Grey matter and white matter volumes for longitudinal comparisons	Software not freely available
SIENAX-MTP	No	Grey matter and white matter volumes for longitudinal comparisons	Software not freely available
BBSI	Yes	Percentage brain volume change between two time points	Only provides global measures that include grey matter and white matter
CLADA	No	Cortical thickness	Software not freely available
NeuroQuant (FDA clearance and CE mark received)	No	Global and regional grey matter volumes	Validation of results is only external; segmentations affected by the presence of brain lesions; only images from the scanner can be analysed (that is, filled images cannot be used)
Icometrix (FDA clearance and CE mark received)	No	Global and regional grey matter volumes	Whole verification of the results is not direct; data analysed remotely
Biometrica (CE mark received)	No	Global and regional grey matter volumes	Whole verification of the results is not direct
Quantib (FDA clearance and CE mark received)	No	Global and regional grey matter volumes	Whole verification of the results is not direct
Cordial	Yes	Spinal cord volume	Limited information about the method as it has not been used extensively
Spinal Cord Toolbox	Yes	Spinal cord area, volume and length	Regions of interest should be edited and manually corrected
JIM	No	Spinal cord area, volume and length	Needs several reference marks for accurate estimates

BBSI, Brain boundary shift integral; CE, Conformité Européenne; CLADA, cortical longitudinal atrophy detection; GIF, Geodesical information flows; JIM, Jacobian integration method; SIENA, Structural Image Evaluation, using Normalization, of Atrophy; SPM, statistical parametric mapping; VBM, voxel-based morphometry. <sup>a</sup>Not exhaustive, only major limitations are included. <sup>b</sup>SPM itself is free but a MATLAB licence is needed.

voxels<sup>113</sup>. To improve the segmentation, the a priori spatial information for each voxel can be included, thereby increasing the probability that each voxel belongs to specific tissue type on the basis of its location<sup>135</sup>, although the accuracy of this step strongly depends on the anatomical similarity between the MRI image and the a priori tissue maps used. To avoid problems due to an anatomical mismatch with the atlas, only MRI images with high anatomical similarity should be used to provide the voxel location information<sup>136</sup>. Use of different anatomical maps, such as probability maps of tissues or structure labelling maps, can also offer improvements<sup>137</sup>. Other approaches that do not depend on the PVE can provide a measure of cortical thickness by calculating the distances between pairs of voxels at the grey matter–white matter and grey matter–CSF interfaces perpendicular to the grey matter–white matter surface interface.

These methods tend to be more susceptible than some of the previously mentioned methods to the low-intensity contrast between tissues because they heavily rely on the gradient intensities between tissue interfaces<sup>138,139</sup>.

Registration-based software packages provide measures of total brain volume, grey matter volume and white matter volume changes by comparison of serially acquired MRI images from the same individual. A common preliminary step in most of these procedures is registration of all MRI images from the same subject on the same virtual space. The first such software packages that were used in longitudinal analyses<sup>113,140</sup> involved registration of two MRI images of the same individual and measurement of whole brain volume change by analysing the shift of the parenchyma–CSF border over time. Newer approaches apply different methods to enable assessment of grey matter and white matter volume changes. In one, for each



voxel, the intensity information from neighbouring voxels at each time point is used<sup>141</sup>. In another, a new intensity harmonization scheme is applied to all MRI images from one individual, with the aim of assigning similar intensity to voxels with similar content of PVE<sup>142</sup>. Another approach, known as the Jacobian integration method<sup>143,144</sup>, is based on local assessment of relative volumetric differences between two MRI images of the same individual, one of which is usually the baseline image; the net sum of all local volumetric changes provides an estimate of total volume changes over time. Finally, cortical thickness changes can be detected by the use of a within-subject template (an MRI image created by merging all MRI images from one individual) to improve cortical thickness estimation at each time point, or by fitting a subject-specific cortical deformable model at each time point<sup>145,146</sup>.

Assessment of spinal cord atrophy is more difficult than brain segmentation owing to particular anatomical (higher mobility and smaller dimensions than the brain) and imaging (lower tissue contrast) features of the spinal cord. Semiautomated (Cordial)<sup>147</sup> and automated (Spinal Cord Toolbox)<sup>148</sup> tools have now been developed, based on deformable models. These promising new software tools still need to be extensively validated on independent datasets before they can even be considered for use in clinical practice.

Academic software packages have important advantages over commercial software packages, such as the fact that they have been validated in many studies under a plethora of different MRI conditions over the past decade. However, they have the severe limitation of being highly technically demanding and their use is therefore limited to centres that are specialized in MRI processing. In addition, clinical application of software to support diagnosis or care is only permitted with products that have received the “Conformité Européenne” (CE) mark in Europe or FDA clearance in the USA. For this reason, translation of imaging analysis software tools to clinical practice is challenging and almost unfeasible for academic neuroimaging laboratories.

In the past 10 years, several companies have proposed centralized MRI reading services, often using their in-house software for quantification of atrophy (TABLE 2). Four software packages have been approved for use in Europe and three of these have also received FDA clearance in the USA. The IcoBrain MS (Icometrix, previously MSmetrix)<sup>149</sup> quantifies cross-sectional volumes with software based on Nifty Seg and quantifies longitudinal changes in grey matter and white matter with software that implements Jacobian integration. NeuroQuant (CorTechs Labs)<sup>150</sup> provides both cross-sectional and longitudinal quantification of atrophy<sup>151</sup>, building on approaches already developed by previous methods<sup>138</sup>. Biometrica MS (Jung Diagnostics) builds on developments of Statistical Parametric Mapping, a software library for neuroimaging analysis, for atrophy measurement and of Lesion Segmentation Tool software for automatic lesion segmentation<sup>152,153</sup>. Quantib Brain (Quantib) is a platform that is integrated into the General Electric MRI scanner and can assess cross-sectional brain volumes and longitudinal changes in volume. IcoBrain MS and Biometrica MS are offered as remote analysis

services, Quantib Brain can be run locally or on a vendor console (General Electric), and NeuroQuant can be a remote analysis service or local installation. All packages have the CE mark and, with the exception of Biometrica, FDA clearance. These certifications guarantee standardization of procedures and results, meaning the software can be used as medical devices.

Importantly, the companies must provide the magnitude of the error in their results, and health care professionals should use this information to validate or discard findings of analyses. All four commercial software packages have been evaluated scientifically to some extent but not completely. To our knowledge, only MSmetrix has been validated by an independent group in the context of MS<sup>154</sup>. Furthermore, the real-world clinical value of these software packages has not yet been assessed, and the procedures are not widely reimbursed (with a few exceptions, such as in the USA). Although promising, these analytical approaches should therefore be more extensively validated by expert groups in the field of MRI preprocessing, especially in the context of MS<sup>134</sup>, before they can be considered for use in the routine clinical setting.

#### Statements and recommendations

1. We recommend appropriate management of several scanner-related factors (including, but not limited to, variation in acquisition protocols, different scanner systems and upgrades, movement artefacts and gradient distortions) to ensure reliability of brain volume estimates, particularly at an individual patient level.
2. We recommend appropriate management of physiological and MS-related factors (including, but not limited to, age, sex, hydration status, time of day, steroid use and MS-related parenchymal alterations).
3. Brain volume measures are software-dependent so the use of software that has been approved as a medical device and independently evaluated in MS is a prerequisite; we recommend further research to validate existing software tools in MS and assess their clinical value.

#### Conclusions and future directions

Based on the evidence reviewed, the idea that brain volume changes and, to a lesser extent, spinal cord atrophy are helpful predictors of the evolution of MS before initiation of therapy is undisputed, so these measures could be valid treatment-decision tools. The evidence reviewed also supports the idea that brain volume measures have value in monitoring the effects of MS drugs as part of the no evidence of disease activity outcome measure or minimal evidence of disease activity outcome measure. However, several potential sources of substantial error remain, including, but not limited to, differential effects of drugs on brain volume measures, confounding physiological and technical factors and the performance and value of volumetric tools. To make implementation of volume measurements in clinical practice feasible, these potential sources of error need to be accounted for and appropriately managed, and further research is needed to ensure the accuracy and reliability of the measurements.

Published online 24 February 2020

#### Deformable model

A mathematical construct capable of representing a broad range of shapes that can be used in MRI image registration or segmentation.

#### Nifty Seg

An open-source image analysis tool developed at University College London to perform segmentations of MRI images.

# CONSENSUS STATEMENT

- Kawachi, I. & Lassmann, H. Neurodegeneration in multiple sclerosis and neuromyelitis optica. *J. Neurol. Neurosurg. Psychiatry* **88**, 137–145 (2017).
- Ferguson, B., Matyszak, M. K., Esiri, M. M. & Perry, V. H. Axonal damage in acute multiple sclerosis lesions. *Brain* **120**, 393–399 (1997).
- Trapp, B. D. et al. Axonal transection in the lesions of multiple sclerosis. *N. Engl. J. Med.* **338**, 278–285 (1998).
- Kuhlmann, T., Lingfeld, G., Bitsch, A., Schuchardt, J. & Bruck, W. Acute axonal damage in multiple sclerosis is most extensive in early disease stages and decreases over time. *Brain* **125**, 2202–2212 (2002).
- Bjartmar, C., Kidd, G., Mork, S., Rudick, R. & Trapp, B. D. Neurological disability correlates with spinal cord axonal loss and reduced N-acetyl aspartate in chronic multiple sclerosis patients. *Ann. Neurol.* **48**, 895–901 (2000).
- Sastre-Garriga, J. et al. Brain atrophy in natalizumab-treated patients: a 3-year follow-up. *Mult. Scler.* **21**, 749–756 (2015).
- Popescu, V. et al. Postmortem validation of MRI cortical volume measurements in MS. *Hum. Brain Mapp.* **37**, 2223–2233 (2016).
- Losseff, N. A. et al. Progressive cerebral atrophy in multiple sclerosis. A serial MRI study. *Brain* **119**, 2009–2019 (1996).
- Losseff, N. A. et al. Spinal cord atrophy and disability in multiple sclerosis. A new reproducible and sensitive MRI method with potential to monitor disease progression. *Brain* **119**, 701–708 (1996).
- Miller, D. H. et al. Guidelines for the use of magnetic resonance techniques in monitoring the treatment of multiple sclerosis. US National MS Society task force. *Ann. Neurol.* **39**, 6–16 (1996).
- Thompson, A. J. et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol.* **17**, 162–173 (2017).
- McDonald, W. I. et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the international panel on the diagnosis of multiple sclerosis. *Ann. Neurol.* **50**, 121–127 (2001).
- Lublin, F. D. et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. *Neurology* **83**, 278–286 (2014).
- Miller, D. H., Barkhof, F., Frank, J. A., Parker, G. J. & Thompson, A. J. Measurement of atrophy in multiple sclerosis: pathological basis, methodological aspects and clinical relevance. *Brain* **125**, 1676–1695 (2002).
- Rocca, M. A. et al. Brain MRI atrophy quantification in MS: from methods to clinical application. *Neurology* **88**, 403–413 (2017).
- Wattjes, M. P. et al. Evidence-based guidelines: MAGNIMS consensus guidelines on the use of MRI in multiple sclerosis — establishing disease prognosis and monitoring patients. *Nat. Rev. Neurol.* **11**, 597–606 (2015).
- Rovira, A. et al. Evidence-based guidelines: MAGNIMS consensus guidelines on the use of MRI in multiple sclerosis — clinical implementation in the diagnostic process. *Nat. Rev. Neurol.* **11**, 471–482 (2015).
- Rao, S. M. et al. Chronic progressive multiple sclerosis. Relationship between cerebral ventricular size and neuropsychological impairment. *Arch. Neurol.* **42**, 678–682 (1985).
- Comi, G. et al. Brain magnetic resonance imaging correlates of cognitive impairment in multiple sclerosis. *J. Neurol. Sci.* **115**, S66–S73 (1993).
- Huber, S. J. et al. Magnetic resonance imaging correlates of dementia in multiple sclerosis. *Arch. Neurol.* **44**, 732–736 (1987).
- De Stefano, N. et al. Assessing brain atrophy rates in a large population of untreated multiple sclerosis subtypes. *Neurology* **74**, 1868–1876 (2010).
- Amato, M. P. et al. Association of MRI metrics and cognitive impairment in radiologically isolated syndromes. *Neurology* **78**, 309–314 (2012).
- Battaglini, M. et al. Lifespan normative data on rates of brain volume changes. *Neurobiol. Aging* **81**, 30–37 (2019).
- De Stefano, N. et al. Establishing pathological cut-offs of brain atrophy rates in multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry* **87**, 93–99 (2016).
- Perez-Miralles, F. C. et al. Predictive value of early brain atrophy on response in patients treated with interferon beta. *Neurol. Neuroimmunol. Neuroinflamm.* **2**, e132 (2015).
- Andorra, M. et al. Assessing biological and methodological aspects of brain volume loss in multiple sclerosis. *JAMA Neurol.* **75**, 1246–1255 (2018).
- Perez-Miralles, F. et al. Clinical impact of early brain atrophy in clinically isolated syndromes. *Mult. Scler.* **19**, 1878–1886 (2013).
- Di Filippo, M. et al. Brain atrophy and lesion load measures over 1 year relate to clinical status after 6 years in patients with clinically isolated syndromes. *J. Neurol. Neurosurg. Psychiatry* **81**, 204–208 (2010).
- Minneboo, A. et al. Predicting short-term disability progression in early multiple sclerosis: added value of MRI parameters. *J. Neurol. Neurosurg. Psychiatry* **79**, 917–923 (2008).
- Sastre-Garriga, J. et al. Long-term clinical outcome of primary progressive MS: predictive value of clinical and MRI data. *Neurology* **65**, 633–635 (2005).
- Khaleeli, Z. et al. Predicting progression in primary progressive multiple sclerosis: a 10-year multicenter study. *Ann. Neurol.* **63**, 790–793 (2008).
- Rocca, M. A. et al. Long-term disability progression in primary progressive multiple sclerosis: a 15-year study. *Brain* **140**, 2814–2819 (2017).
- Popescu, V. et al. Brain atrophy and lesion load predict long term disability in multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry* **84**, 1082–1091 (2013).
- Miller, D. H. et al. Brain atrophy and disability worsening in primary progressive multiple sclerosis: insights from the INFORMS study. *Ann. Clin. Transl. Neurol.* **5**, 346–356 (2018).
- Sormani, M. P. et al. Defining brain volume cutoffs to identify clinically relevant atrophy in RRMS. *Mult. Scler.* **23**, 656–664 (2017).
- Daams, M. et al. Mean upper cervical cord area (MUCCA) measurement in long-standing multiple sclerosis: relation to brain findings and clinical disability. *Mult. Scler.* **20**, 1860–1865 (2014).
- Bieniek, M. et al. Cord atrophy separates early primary progressive and relapsing remitting multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry* **77**, 1036–1039 (2006).
- Lukas, C. et al. Relevance of spinal cord abnormalities to clinical disability in multiple sclerosis: MR imaging findings in a large cohort of patients. *Radiology* **269**, 542–552 (2013).
- Rocca, M. A. et al. A multicenter assessment of cervical cord atrophy among MS clinical phenotypes. *Neurology* **76**, 2096–2102 (2011).
- Daams, M. et al. Unraveling the neuroimaging predictors for motor dysfunction in long-standing multiple sclerosis. *Neurology* **85**, 248–255 (2015).
- Oh, J. et al. Relationships between quantitative spinal cord MRI and retinal layers in multiple sclerosis. *Neurology* **84**, 720–728 (2015).
- Kearney, H. et al. Magnetic resonance imaging correlates of physical disability in relapse onset multiple sclerosis of long disease duration. *Mult. Scler.* **20**, 72–80 (2014).
- Tsagkas, C. et al. Spinal cord volume loss: a marker of disease progression in multiple sclerosis. *Neurology* **91**, e349–e358 (2018).
- Schlaeger, R. et al. Spinal cord gray matter atrophy correlates with multiple sclerosis disability. *Ann. Neurol.* **76**, 568–580 (2014).
- Yiannakas, M. C., Kakar, P., Hoy, L. R., Miller, D. H. & Wheeler-Kingshott, C. A. The use of the lumbosacral enlargement as an intrinsic imaging biomarker: feasibility of grey matter and white matter cross-sectional area measurements using MRI at 3 T. *PLoS One* **9**, e105544 (2014).
- Rocca, M. A. et al. Voxel-wise mapping of cervical cord damage in multiple sclerosis patients with different clinical phenotypes. *J. Neurol. Neurosurg. Psychiatry* **84**, 35–41 (2013).
- Lukas, C. et al. Cervical spinal cord volume loss is related to clinical disability progression in multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry* **86**, 410–418 (2015).
- Tsagkas, C. et al. Preferential spinal cord volume loss in primary progressive multiple sclerosis. *Mult. Scler.* **25**, 947–957 (2019).
- Kearney, H. et al. Cervical cord lesion load is associated with disability independently from atrophy in MS. *Neurology* **84**, 367–373 (2015).
- Chard, D. T. et al. Brain atrophy in clinically early relapsing-remitting multiple sclerosis. *Brain* **125**, 327–337 (2002).
- Sastre-Garriga, J. et al. Grey and white matter atrophy in early clinical stages of primary progressive multiple sclerosis. *Neuroimage* **22**, 353–359 (2004).
- De Stefano, N. et al. Evidence of early cortical atrophy in MS: relevance to white matter changes and disability. *Neurology* **60**, 1157–1162 (2003).
- Quarantelli, M. et al. Brain tissue volume changes in relapsing-remitting multiple sclerosis: correlation with lesion load. *Neuroimage* **18**, 360–366 (2003).
- Fisher, E., Lee, J. C., Nakamura, K. & Rudick, R. A. Gray matter atrophy in multiple sclerosis: a longitudinal study. *Ann. Neurol.* **64**, 255–265 (2008).
- Roosendaal, S. D. et al. Grey matter volume in a large cohort of MS patients: relation to MRI parameters and disability. *Mult. Scler.* **17**, 1098–1106 (2011).
- Dalton, C. M. et al. Early development of multiple sclerosis is associated with progressive grey matter atrophy in patients presenting with clinically isolated syndromes. *Brain* **127**, 1101–1107 (2004).
- Valsasina, P. et al. Evidence for progressive gray matter loss in patients with relapsing-remitting MS. *Neurology* **65**, 1126–1128 (2005).
- Tiberio, M. et al. Gray and white matter volume changes in early RRMS: a 2-year longitudinal study. *Neurology* **64**, 1001–1007 (2005).
- Sastre-Garriga, J. et al. Grey and white matter volume changes in early primary progressive multiple sclerosis: a longitudinal study. *Brain* **128**, 1454–1460 (2005).
- Fisniku, L. K. et al. Gray matter atrophy is related to long-term disability in multiple sclerosis. *Ann. Neurol.* **64**, 247–254 (2008).
- Filippi, M. et al. Gray matter damage predicts the accumulation of disability 13 years later in MS. *Neurology* **81**, 1759–1767 (2013).
- Jacobsen, C. et al. Brain atrophy and disability progression in multiple sclerosis patients: a 10-year follow-up study. *J. Neurol. Neurosurg. Psychiatry* **85**, 1109–1115 (2014).
- Narayana, P. A. et al. Regional cortical thickness in relapsing remitting multiple sclerosis: a multi-center study. *Neuroimage Clin.* **2**, 120–131 (2012).
- Lansley, J., Mataix-Cols, D., Grau, M., Radua, J. & Sastre-Garriga, J. Localized grey matter atrophy in multiple sclerosis: a meta-analysis of voxel-based morphometry studies and associations with functional disability. *Neurosci. Biobehav. Rev.* **37**, 819–830 (2013).
- Eshaghi, A. et al. Deep gray matter volume loss drives disability worsening in multiple sclerosis. *Ann. Neurol.* **83**, 210–222 (2018).
- Pichler, A. et al. Combined analysis of global and compartmental brain volume changes in early multiple sclerosis in clinical practice. *Mult. Scler.* **22**, 340–346 (2016).
- Vidal-Jordana, A., Sastre-Garriga, J., Rovira, A. & Montalban, X. Treating relapsing-remitting multiple sclerosis: therapy effects on brain atrophy. *J. Neurol.* **262**, 2617–2626 (2015).
- Filippi, M. et al. Interferon beta-1a for brain tissue loss in patients at presentation with syndromes suggestive of multiple sclerosis: a randomised, double-blind, placebo-controlled trial. *Lancet* **364**, 1489–1496 (2004).
- Rudick, R. A., Fisher, E., Lee, J. C., Duda, J. T. & Simon, J. Brain atrophy in relapsing multiple sclerosis: relationship to relapses, EDSS, and treatment with interferon beta-1a. *Mult. Scler.* **6**, 365–372 (2000).
- Rudick, R. A., Fisher, E., Lee, J. C., Simon, J. & Jacobs, L. Use of the brain parenchymal fraction to measure whole brain atrophy in relapsing-remitting MS. Multiple Sclerosis Collaborative Research Group. *Neurology* **53**, 1698–1704 (1999).
- Fisher, E. et al. Effect of intramuscular interferon beta-1a on gray matter atrophy in relapsing-remitting multiple sclerosis: a retrospective analysis. *Mult. Scler.* **22**, 668–676 (2016).
- Vidal-Jordana, A. et al. Early brain pseudoatrophy while on natalizumab therapy is due to white matter volume changes. *Mult. Scler.* **19**, 1175–1181 (2013).
- Vidal-Jordana, A. et al. Brain volume loss during the first year of interferon-beta treatment in multiple sclerosis: baseline inflammation and regional brain volume dynamics. *J. Neuroimaging* **26**, 532–538 (2016).
- Comi, G. et al. Effects of early treatment with glatiramer acetate in patients with clinically isolated syndrome. *Mult. Scler.* **19**, 1074–1083 (2013).
- Wolinsky, J. S. et al. Glatiramer acetate in primary progressive multiple sclerosis: results of a multinational, multicenter, double-blind, placebo-controlled trial. *Ann. Neurol.* **61**, 14–24 (2007).
- Montalban, X. et al. A single-center, randomized, double-blind, placebo-controlled study of interferon beta-1b on primary progressive and transitional multiple sclerosis. *Mult. Scler.* **15**, 1195–1205 (2009).

77. Miller, D. H. et al. MRI outcomes in a placebo-controlled trial of natalizumab in relapsing MS. *Neurology* **68**, 1390–1401 (2007).
78. Branger, P., Parienti, J. J., Sormani, M. P. & Defer, C. The effect of disease-modifying drugs on brain atrophy in relapsing-remitting multiple sclerosis: a meta-analysis. *PLoS One* **11**, e0149685 (2016).
79. Cohen, J. A. et al. Alemtuzumab versus interferon beta 1a as first-line treatment for patients with relapsing-remitting multiple sclerosis: a randomised controlled phase 3 trial. *Lancet* **380**, 1819–1828 (2012).
80. Coles, A. J. et al. Alemtuzumab for patients with relapsing multiple sclerosis after disease-modifying therapy: a randomised controlled phase 3 trial. *Lancet* **380**, 1829–1839 (2012).
81. Cohen, J. A. et al. Oral fingolimod or intramuscular interferon for relapsing multiple sclerosis. *N. Engl. J. Med.* **362**, 402–415 (2010).
82. Arnold, D. L. et al. Effects of delayed-release dimethyl fumarate on MRI measures in the phase 3 DEFINE study. *J. Neurol.* **261**, 1794–1802 (2014).
83. De Stefano, N. et al. Reduced brain atrophy rates are associated with lower risk of disability progression in patients with relapsing multiple sclerosis treated with cladribine tablets. *Mult. Scler.* **24**, 222–226 (2018).
84. Sormani, M. P., Arnold, D. L. & De Stefano, N. Treatment effect on brain atrophy correlates with treatment effect on disability in multiple sclerosis. *Ann. Neurol.* **75**, 43–49 (2014).
85. Chataway, J. et al. Effect of high-dose simvastatin on brain atrophy and disability in secondary progressive multiple sclerosis (MS-STAT): a randomised, placebo-controlled, phase 2 trial. *Lancet* **383**, 2213–2221 (2014).
86. Fox, R. J. et al. Phase 2 trial of ibudilast in progressive multiple sclerosis. *N. Engl. J. Med.* **379**, 846–855 (2018).
87. Kearney, H., Miller, D. H. & Ciccarelli, O. Spinal cord MRI in multiple sclerosis — diagnostic, prognostic and clinical value. *Nat. Rev. Neurol.* **11**, 327–338 (2015).
88. Kapoor, R. et al. Lamotrigine for neuroprotection in secondary progressive multiple sclerosis: a randomised, double-blind, placebo-controlled, parallel-group trial. *Lancet Neurol.* **9**, 681–688 (2010).
89. Cawley, N. et al. Spinal cord atrophy as a primary outcome measure in phase II trials of progressive multiple sclerosis. *Mult. Scler.* **24**, 932–941 (2018).
90. Leary, S. M. et al. Interferon beta-1a in primary progressive MS: an exploratory, randomized, controlled trial. *Neurology* **60**, 44–51 (2003).
91. Lublin, F. et al. Oral fingolimod in primary progressive multiple sclerosis (INFORMS): a phase 3, randomised, double-blind, placebo-controlled trial. *Lancet* **387**, 1075–1084 (2016).
92. Sormani, M. P. et al. Fingolimod effect on brain volume loss independently contributes to its effect on disability. *Mult. Scler.* **21**, 916–924 (2015).
93. Fisher, E. et al. Eight-year follow-up study of brain atrophy in patients with MS. *Neurology* **59**, 1412–1420 (2002).
94. Rojas, J. I., Patrucco, L., Miguez, J., Besada, C. & Cristiano, E. Brain atrophy as a non-response predictor to interferon-beta in relapsing-remitting multiple sclerosis. *Neurol. Res.* **36**, 615–618 (2014).
95. Radue, E. W. et al. Correlation between brain volume loss and clinical and MRI outcomes in multiple sclerosis. *Neurology* **84**, 784–793 (2015).
96. Jeffery, D. R. et al. The relationship between the rate of brain volume loss during first 24 months and disability progression over 24 and 48 months in relapsing MS. *J. Neurol.* **263**, 299–305 (2016).
97. Kappos, L. 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.* **22**, 1297–1305 (2016).
98. Gasperini, C. et al. Unraveling treatment response in multiple sclerosis: a clinical and MRI challenge. *Neurology* **92**, 180–192 (2019).
99. Rio, J. et al. Disability progression markers over 6–12 years in interferon-beta-treated multiple sclerosis patients. *Mult. Scler.* **24**, 322–330 (2018).
100. Mueller, S. G. et al. Ways toward an early diagnosis in Alzheimer's disease: the Alzheimer's Disease Neuroimaging Initiative (ADNI). *Alzheimers Dement.* **1**, 55–66 (2005).
101. Caramanos, Z. et al. Gradient distortions in MRI: characterizing and correcting for their effects on SIENA-generated measures of brain volume change. *Neuroimage* **49**, 1601–1611 (2010).
102. Jovicich, J. et al. Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. *Neuroimage* **30**, 436–443 (2006).
103. Han, X. et al. Reliability of MRI-derived measurements of human cerebral cortical thickness: the effects of field strength, scanner upgrade and manufacturer. *Neuroimage* **32**, 180–194 (2006).
104. MacLaren, J., Han, Z., Vos, S. B., Fischbein, N. & Bammer, R. Reliability of brain volume measurements: a test–retest dataset. *Sci. Data* **1**, 140037 (2014).
105. Morey, R. A. et al. Scan–rescan reliability of subcortical brain volumes derived from automated segmentation. *Hum. Brain Mapp.* **31**, 1751–1762 (2010).
106. Sampat, M. P. et al. Disease modeling in multiple sclerosis: assessment and quantification of sources of variability in brain parenchymal fraction measurements. *Neuroimage* **52**, 1367–1373 (2010).
107. Yang, C. Y. et al. Reproducibility of brain morphology from short-term repeat clinical MRI examinations: a retrospective study. *PLoS One* **11**, e0146913 (2016).
108. Biberacher, V. et al. Intra- and interscanner variability of magnetic resonance imaging based volumetry in multiple sclerosis. *Neuroimage* **142**, 188–197 (2016).
109. Reuter, M. et al. Head motion during MRI acquisition reduces gray matter volume and thickness estimates. *Neuroimage* **107**, 107–115 (2015).
110. Godenschweger, F. et al. Motion correction in MRI of the brain. *Phys. Med. Biol.* **61**, R32–R56 (2016).
111. Jovicich, J. et al. MRI-derived measurements of human subcortical, ventricular and intracranial brain volumes: reliability effects of scan sessions, acquisition sequences, data analyses, scanner upgrade, scanner vendors and field strengths. *Neuroimage* **46**, 177–192 (2009).
112. Barnes, J. et al. Head size, age and gender adjustment in MRI studies: a necessary nuisance? *Neuroimage* **53**, 1244–1255 (2010).
113. Smith, S. M. et al. Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. *Neuroimage* **17**, 479–489 (2002).
114. Amiri, H. et al. Urgent challenges in quantification and interpretation of brain grey matter atrophy in individual MS patients using MRI. *Neuroimage Clin.* **19**, 466–475 (2018).
115. Enzinger, C. et al. Risk factors for progression of brain atrophy in aging: six-year follow-up of normal subjects. *Neurology* **64**, 1704–1711 (2005).
116. Fraser, M. A., Shaw, M. E. & Cherbuin, N. A systematic review and meta-analysis of longitudinal hippocampal atrophy in healthy human ageing. *Neuroimage* **112**, 364–374 (2015).
117. Ruigrok, A. N. et al. A meta-analysis of sex differences in human brain structure. *Neurosci. Biobehav. Rev.* **39**, 34–50 (2014).
118. Kiraly, A. et al. Male brain ages faster: the age and gender dependence of subcortical volumes. *Brain Imaging Behav.* **10**, 901–910 (2016).
119. Good, C. D. et al. Cerebral asymmetry and the effects of sex and handedness on brain structure: a voxel-based morphometric analysis of 465 normal adult human brains. *Neuroimage* **14**, 685–700 (2001).
120. Luders, E. et al. Gender effects on cortical thickness and the influence of scaling. *Hum. Brain Mapp.* **27**, 314–324 (2006).
121. De Bondt, T., Pullens, P., Van Hecke, W., Jacquemyn, Y. & Parizel, P. M. Reproducibility of hormone-driven regional grey matter volume changes in women using SPM8 and SPM12. *Brain Struct. Funct.* **221**, 4631–4641 (2016).
122. Nakamura, K. et al. Diurnal fluctuations in brain volume: statistical analyses of MRI from large populations. *Neuroimage* **118**, 126–132 (2015).
123. Duning, T. et al. Dehydration confounds the assessment of brain atrophy. *Neurology* **64**, 548–550 (2005).
124. Klaren, R. E. et al. Objectively measured physical activity is associated with brain volumetric measurements in multiple sclerosis. *Behav. Neurol.* **2015**, 482536 (2015).
125. Mechtcheriakov, S. et al. A widespread distinct pattern of cerebral atrophy in patients with alcohol addiction revealed by voxel-based morphometry. *J. Neurol. Neurosurg. Psychiatry* **78**, 610–614 (2007).
126. Mashhoon, Y., Sava, S., Sneider, J. T., Nickerson, L. D. & Silveri, M. M. Cortical thickness and volume differences associated with marijuana abuse in emerging adults. *Drug Alcohol Depend.* **155**, 275–283 (2015).
127. Beauchet, O. et al. Blood pressure levels and brain volume reduction: a systematic review and meta-analysis. *J. Hypertens.* **31**, 1502–1516 (2013).
128. Hooshmand, B. et al. Association of vitamin B12, folate, and sulfur amino acids with brain magnetic resonance imaging measures in older adults: a longitudinal population-based study. *JAMA Psychiatry* **73**, 606–613 (2016).
129. Pichler, A. et al. The impact of vascular risk factors on brain volume and lesion load in patients with early multiple sclerosis. *Mult. Scler.* **25**, 48–54 (2019).
130. Chard, D. T., Jackson, J. S., Miller, D. H. & Wheeler-Kingshott, C. A. Reducing the impact of white matter lesions on automated measures of brain gray and white matter volumes. *J. Magn. Reson. Imaging* **32**, 223–228 (2010).
131. Battaglini, M., Jenkinson, M. & De Stefano, N. Evaluating and reducing the impact of white matter lesions on brain volume measurements. *Hum. Brain Mapp.* **33**, 2062–2071 (2012).
132. Cheriyan, J., Kim, S., Wolansky, L. J., Cook, S. D. & Cadavid, D. Impact of inflammation on brain volume in multiple sclerosis. *Arch. Neurol.* **69**, 82–88 (2012).
133. Zivadinov, R. et al. Mechanisms of action of disease-modifying agents and brain volume changes in multiple sclerosis. *Neurology* **71**, 136–144 (2008).
134. Storelli, L. et al. Measurement of whole-brain and gray matter atrophy in multiple sclerosis: assessment with MR imaging. *Radiology* **288**, 554–564 (2018).
135. Ashburner, J. & Friston, K. J. Unified segmentation. *Neuroimage* **26**, 839–851 (2005).
136. Cardoso, M. J. et al. Geodesic information flows: spatially-variant graphs and their application to segmentation and fusion. *IEEE Trans. Med. Imaging* **34**, 1976–1988 (2015).
137. Avants, B. B., Tustison, N. J., Wu, J., Cook, P. A. & Gee, J. C. An open source multivariate framework for n-tissue segmentation with evaluation on public data. *Neuroinformatics* **9**, 381–400 (2011).
138. Fischl, B., Sereno, M. I. & Dale, A. M. Cortical surface-based analysis. II: inflation, flattening, and a surface-based coordinate system. *Neuroimage* **9**, 195–207 (1999).
139. MacDonald, D., Kabani, N., Avis, D. & Evans, A. C. Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI. *Neuroimage* **12**, 340–356 (2000).
140. Fox, N. C. & Freeborough, P. A. Brain atrophy progression measured from registered serial MRI: validation and application to Alzheimer's disease. *J. Magn. Reson. Imaging* **7**, 1069–1075 (1997).
141. Dwyer, M. G., Bergsland, N. & Zivadinov, R. Improved longitudinal gray and white matter atrophy assessment via application of a 4-dimensional hidden Markov random field model. *Neuroimage* **90**, 207–217 (2014).
142. Battaglini, M., Jenkinson, M. & De Stefano, N. SIENA-XL for improving the assessment of gray and white matter volume changes on brain MRI. *Hum. Brain Mapp.* **39**, 1063–1077 (2018).
143. Ashburner, J. & Ridgway, G. R. Symmetric diffeomorphic modeling of longitudinal structural MRI. *Front. Neurosci.* **6**, 197 (2012).
144. Nakamura, K. et al. Jacobian integration method increases the statistical power to measure gray matter atrophy in multiple sclerosis. *Neuroimage Clin.* **4**, 10–17 (2014).
145. Reuter, M. & Fischl, B. Avoiding asymmetry-induced bias in longitudinal image processing. *Neuroimage* **57**, 19–21 (2011).
146. Nakamura, K., Fox, R. & Fisher, E. CLADA: cortical longitudinal atrophy detection algorithm. *Neuroimage* **54**, 278–289 (2011).
147. Amann, M. et al. Reliable volumetry of the cervical spinal cord in MS patient follow-up data with cord image analyzer (Cordial). *J. Neurol.* **263**, 1364–1374 (2016).
148. De Leener, B. et al. SCT: spinal cord toolbox, an open-source software for processing spinal cord MRI data. *Neuroimage* **145**, 24–43 (2017).
149. Jain, S. et al. Automatic segmentation and volumetry of multiple sclerosis brain lesions from MR images. *Neuroimage Clin.* **8**, 367–375 (2015).
150. Ross, D. E., Ochs, A. L., Seabaugh, J. M. & Shrader, C. R. Man versus machine: comparison of radiologists' interpretations and NeuroQuant® volumetric analyses of brain MRIs in patients with traumatic brain injury. *J. Neuropsychiatry Clin. Neurosci.* **25**, 32–39 (2013).
151. Brewer, J. B. Fully-automated volumetric MRI with normative ranges: translation to clinical practice. *Behav. Neurol.* **21**, 21–28 (2009).
152. Opfer, R. et al. Atlas based brain volumetry: how to distinguish regional volume changes due to biological or physiological effects from inherent noise of the

methodology. *Magn. Reson. Imaging* **34**, 455–461 (2016).

153. Spies, L. et al. Fully automatic detection of deep white matter T1 hypointense lesions in multiple sclerosis. *Phys. Med. Biol.* **58**, 8323–8337 (2013).

154. Steenwijk, M. D. et al. Agreement of MSmetrix with established methods for measuring cross-sectional and longitudinal brain atrophy. *Neuroimage Clin.* **15**, 843–853 (2017).

155. De Stefano, N. et al. Efficacy of subcutaneous interferon beta-1a on MRI outcomes in a randomised controlled trial of patients with clinically isolated syndromes. *J. Neurol. Neurosurg. Psychiatry* **85**, 647–653 (2014).

156. Kappos, L. et al. Long-term subcutaneous interferon beta-1a therapy in patients with relapsing-remitting MS. *Neurology* **67**, 944–953 (2006).

157. Kappos, L. et al. Effect of early versus delayed interferon beta-1b treatment on disability after a first clinical event suggestive of multiple sclerosis: a 3-year follow-up analysis of the BENEFIT study. *Lancet* **370**, 389–397 (2007).

158. Molyneux, P. D. et al. The effect of interferon beta-1b treatment on MRI measures of cerebral atrophy in secondary progressive multiple sclerosis. European Study Group on Interferon beta-1b in secondary progressive multiple sclerosis. *Brain* **123**, 2256–2263 (2000).

159. Comi, G. et al. Effect of glatiramer acetate on conversion to clinically definite multiple sclerosis in patients with clinically isolated syndrome (PreCISE study): a randomised, double-blind, placebo-controlled trial. *Lancet* **374**, 1503–1511 (2009).

160. Sormani, M. P. et al. Measurement error of two different techniques for brain atrophy assessment in multiple sclerosis. *Neurology* **62**, 1432–1434 (2004).

161. Miller, A. E. et al. Oral teriflunomide for patients with a first clinical episode suggestive of multiple sclerosis (TOPIC): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Neurol.* **13**, 977–986 (2014).

162. O'Connor, P. et al. Randomized trial of oral teriflunomide for relapsing multiple sclerosis. *N. Engl. J. Med.* **365**, 1295–1303 (2011).

163. Wolinsky, J. S. et al. Magnetic resonance imaging outcomes from a phase III trial of teriflunomide. *Mult. Scler.* **19**, 1310–1319 (2013).

164. Radue, E. W. et al. Teriflunomide slows BVL in relapsing MS: a reanalysis of the TEMSO MRI data set using SIENA. *Neurol. Neuroimmunol. Neuroinflamm.* **4**, e390 (2017).

165. Miller, D. H. et al. Effects of delayed-release dimethyl fumarate on MRI measures in the phase 3 CONFIRM study. *Neurology* **84**, 1145–1152 (2015).

166. Radue, E. W. et al. Natalizumab plus interferon beta-1a reduces lesion formation in relapsing multiple sclerosis. *J. Neurol. Sci.* **292**, 28–35 (2010).

167. Kapoor, R. et al. Effect of natalizumab on disease progression in secondary progressive multiple sclerosis (ASCEND): a phase 3, randomised, double-blind, placebo-controlled trial with an open-label extension. *Lancet Neurol.* **17**, 405–415 (2018).

168. Radue, E. W. et al. Impact of fingolimod therapy on magnetic resonance imaging outcomes in patients with multiple sclerosis. *Arch. Neurol.* **69**, 1259–1269 (2012).

169. Calabresi, P. A. 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.* **13**, 545–556 (2014).

170. Hauser, S. L. et al. Ocrelizumab versus interferon beta-1a in relapsing multiple sclerosis. *N. Engl. J. Med.* **376**, 221–234 (2017).

171. Montalban, X. et al. Ocrelizumab versus placebo in primary progressive multiple sclerosis. *N. Engl. J. Med.* **376**, 209–220 (2017).

172. Leist, T. P. et al. Effect of oral cladribine on time to conversion to clinically definite multiple sclerosis in patients with a first demyelinating event (ORACLE MS): a phase 3 randomised trial. *Lancet Neurol.* **13**, 257–267 (2014).

### Acknowledgements

O.C., F.B. and T.A.Y. receive support from the National Institute for Health Research University College London Hospitals Biomedical Research Centre.

### Author contributions

J.S.-G. and A.R. developed the idea for the manuscript and created the structure. J.S.-G., D.P. and M.B. each wrote specific sections of the manuscript. All authors participated in the workshop, commented on and revised the manuscript and approved the final version.

### Competing interests

J.S.-G. declares grants and personal fees from Genzyme, personal fees from Almirall, Biogen, Celgene, Merck, Novartis, Roche and Teva and is a member of the Editorial Committee of *Multiple Sclerosis Journal* and Director of the Scientific Committee of *Revista de Neurologia*, all unrelated to this Consensus Statement. D.P. declares personal fees from Novartis and Sanofi-Genzyme, unrelated to this Consensus Statement. M.A.R. declares personal fees from Biogen Idec, Genzyme, Merck Serono, Novartis, Roche, Sanofi-Aventis and Teva, all unrelated to this Consensus Statement. O.C. declares grants from the Spinal Cord Research Foundation, the Rosetree Trust, the Progressive MS Alliance, Bioclinica, GE Neuro, EU-H2020, the UK MS Society, the National MS Society, and the National Institute for Health Research Biomedical Research Centre, University College London Hospitals, and declares personal fees from Merck, *Neurology*, Novartis, Roche and Teva, all unrelated to this Consensus Statement. O.C. also declares grants from the UK MS Society, the National Institute for Health Research Biomedical Research Centre, University College London Hospitals, and the National MS Society during the course of this Consensus Statement. C.E. declares personal fees from Bayer, Biogen, Celgene, Genzyme, Merck, Novartis, Roche, Sanofi-Aventis, Shire and Teva, all unrelated to this Consensus Statement. J.W. declares employment with Medical Image Analysis Center, scientific advisory board membership for Biogen, Genzyme, Novartis, TEVA and Roche, personal fees from Bayer, Novartis and Roche, and grants from the European Union (Horizon2020), German Federal Ministries of Education and Research (BMBF), and Economic Affairs and Energy (BMWi), all unrelated to this Consensus Statement. M.P.S. declares personal fees from Actelion, Biogen, Genzyme, Merck, Novartis, Roche, Serono, Synthon and Teva, all unrelated to this Consensus Statement. F.B. acts as a consultant for Apitope, Bayer-Schering, Biogen-Idec, GeNeuro, Sanofi-Genzyme, Ixico, Janssen Research, Merck-Serono, Novartis, Roche and TEVA. He has received grants, or grants are pending, from the Amyloid Imaging to Prevent Alzheimer's Disease (AMYPAD) initiative, the Biomedical Research Centre at University College London Hospitals, the Dutch MS Society, ECTRIMS–MAGNIMS, EU-H2020, the Dutch Research Council (NWO), the UK MS Society, and the National Institute for Health Research, University College London. He has received payments for the development of educational presentations from Ixico and to his institution from Biogen-Idec. He is on the editorial board of *Radiology*, *Brain*, *European Radiology*, *Multiple Sclerosis Journal* and *Neurology*. T.A.Y. declares research support from Biogen, GlaxoSmithKline, Novartis and Schering, and honoraria from Bayer Schering, Biogen, Hikma and Novartis. N.D.S. declares personal fees from Biogen-Idec, Celgene, Sanofi-Genzyme, Merck Serono, Novartis, Roche and Teva, grants from Merck Serono, Novartis and the Italian MS Society, and nonfinancial support from Biogen-Idec, Sanofi-Genzyme, Merck Serono, Novartis,

Roche and Teva, all unrelated to this Consensus Statement. M.T. declares personal fees from Almirall, Bayer Healthcare, Biogen Idec, Merck Serono, Novartis, Roche, Sanofi-Genzyme and Teva, grants from Sanofi-Genzyme, and other competing interests related to Biogen Idec, all unrelated to this Consensus Statement. M.F. declares personal fees from Biogen Idec, Merck-Serono, Novartis and Teva, and grants from Biogen Idec, Merck-Serono, Novartis, Roche and Teva, all unrelated to this Consensus Statement. He is the Editor-in-Chief of the *Journal of Neurology*. C.G. declares advisory board membership for Biogen, Genzyme, Merck-Serono, Roche and Teva, and honoraria for speaking or consultation fees from Almirall, Bayer Schering, Biogen Idec, Genzyme, Merck Serono, Novartis, Roche and Teva. L.K. declares competing interests related to Actelion, Almirall, Alkermes, Bayer, Biogen, df-mp, The European Union, Excemed, GeNeuro, Genzyme, Japan Tobacco, Merck, Minoryx, Mitsubishi Pharma, Neurostat, Novartis, Receptos–Celgene, Roche, Roche Research Foundations, Sanofi-Aventis, Santhera, the Swiss Multiple Sclerosis Society, Swiss National Research Foundation, Teva and Vianex, all unrelated to this Consensus Statement. J.R. declares personal fees from Biogen, Merck, Novartis and Roche, and other competing interests related to Merck and Mylan, all unrelated to this Consensus Statement. J.F. declares personal fees from Biogen Idec, Merck Serono and Sanofi-Aventis, participation in scientific advisory boards for Almirall, Genzyme and Novartis, personal fees (speaker honoraria) from Biogen Idec, Merck Serono, Santhera and Teva, and participation as an advisor on preclinical development for Takeda, all unrelated to this Consensus Statement. J.P. declares grants and personal fees from Alexion, Biogen Idec, Chugai, and Merck Serono, personal fees from ABIDE, ARGEX, Med Day, Novartis, Roche and Teva, and grants from ABIDE and MedImmune, all unrelated to this Consensus Statement. She also has a patent (patent number 13704627.2-1408) with Isis: Diagnosing Multiple Sclerosis. H.V. declares grants from Merck Serono, Novartis and Teva, and other competing interests related to Merck and Novartis, all unrelated to this Consensus Statement. He also has a patent related to brain atrophy measurement pending. X.M. declares personal fees from Actelion, Biogen, Celgene, Merck, Novartis, Roche, Sanofi-Genzyme and Teva, all unrelated to this Consensus Statement. A.R. declares personal fees from Bayer, Biogen, Icometrix, Merck Serono, Novartis, Sanofi-Genzyme, Roche and Teva, and other competing interests related to Bayer, Novartis, OLEA Medical, Sanofi-Genzyme and Synthetic MRI, all unrelated to this Consensus Statement. M.B. declares no competing interests.

### Peer review information

*Nature Reviews Neurology* thanks N. Bergsland, D. Ontaneda and A. Trabulsee for their contribution to the peer review of this work.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

### RELATED LINKS

Magnetic Resonance in Imaging in MS (MAGNIMS) study group: [www.magnims.eu](http://www.magnims.eu)

© The Author(s) 2020