

REVIEW ARTICLE

Measurement of atrophy in multiple sclerosis: pathological basis, methodological aspects and clinical relevance

David H. Miller,¹ Frederik Barkhof,³ Joseph A. Frank,⁴ Geoffrey J. M. Parker² and Alan J. Thompson¹

¹Department of Neuroinflammation, Institute of Neurology, London, ²Department of Imaging Science and Biomedical Engineering, University of Manchester, Manchester, UK, ³Department of Radiology, Free University Hospital, Amsterdam, The Netherlands and ⁴Experimental Neuroimaging Section, Laboratory of Diagnostic Radiology Research, National Institutes of Health, Washington, DC, USA

Correspondence to: D. H. Miller, NMR Research Unit, Institute of Neurology, Queen Square, London WC1N 3BG, UK
E-mail: d.miller@ion.ucl.ac.uk

Summary

MRI methods are widely used to follow the pathological evolution of multiple sclerosis in life and its modification by treatment. To date, measures of the number and volume of macroscopically visible lesions have been studied most often. These MRI outcomes have demonstrated clear treatment effects but without a commensurate clinical benefit, suggesting that there are other aspects of multiple sclerosis pathology that warrant investigation. In this context, there has been considerable interest in measuring tissue loss (atrophy) as a more global marker of the adverse outcome of multiple sclerosis pathology, whether it arises in macroscopic lesions or in the normal appearing tissues. An International Workshop recently considered the measurement of atrophy in multiple sclerosis and provided the basis for this review. Brain white matter bulk consists predominantly of axons (46%) followed by myelin (24%), and progressive atrophy implies loss of these structures, especially axons, although variable effects on tissue volumes may also arise from glial cell proliferation or loss, gliosis, inflammation and oedema. Significant correlations found between brain volume and other putative MR neuronal markers also indicate that atrophy reflects axonal loss.

Numerous methods are available for the measurement of global and regional brain volumes and upper cervical cord cross-sectional area that are highly reproducible and sensitive to changes within 6–12 months. In general, 3D-T₁-weighted acquisitions and largely automated segmentation approaches are optimal. Whereas normalized volumes are desirable for cross-sectional studies, absolute volume measures are adequate for serial investigation. Atrophy is seen at all clinical stages of multiple sclerosis, developing gradually following the appearance of inflammatory lesions. This probably reflects both inflammation-induced axonal loss followed by Wallerian degeneration and post-inflammatory neurodegeneration that may be partly due to failure of remyelination. One component of atrophy appears to be independent of focal lesions. Existing immunomodulatory therapies have had limited effects on progressive atrophy, concordant with their modest effects on progressive disability. Atrophy provides a sensitive measure of the neurodegenerative component of multiple sclerosis and should be measured in trials evaluating potential anti-inflammatory, remyelinating or neuroprotective therapies.

Keywords: multiple sclerosis; MRI; atrophy

Abbreviations: BFV = brain fraction volume; BICCR = brain intracranial capacity ratio; BPF = brain parenchymal fraction; CIS = clinically isolated syndromes; COV = coefficient of variance; EDSS = expanded disability status score; FLAIR = fluid-attenuated inversion recovery; IVMP = intravenous methylprednisolone; MR = magnetic resonance; MSFC scale = Multiple Sclerosis Functional Composite scale; NAA = *N*-acetyl aspartate; NAWM = normal appearing white matter; SPM = statistical parametric mapping; WBR = whole brain ratio

Introduction

In the last 15 years, MRI has assumed an important role as a tool to assist in the diagnosis of multiple sclerosis and to monitor the evolving disease *in vivo*. Because of the high conspicuity of lesions seen on MRI, it provides a sensitive measure of the evolving pathology, which is not clinically apparent.

While the value of a sensitive measure has been appreciated in providing new insights into the natural history of the disease and its modification by treatment, it has become apparent that there is more to the pathology of multiple sclerosis than the mere depiction of lesions and quantification of their number and volume. Abnormalities are reported in normal appearing white matter (NAWM) and cortex using quantitative magnetic resonance (MR) methods, and pathological studies confirm the presence of disease in these regions. It has also emerged that destructive or degenerative changes with axonal loss in lesions are relatively common. The extent of such changes varies markedly between lesions, but appears to be greatest in the progressive forms of multiple sclerosis, indicating that this is a more relevant mechanism of disability than the extent of lesions *per se*.

With this background, there has been increasing interest in measuring tissue loss in the CNS in multiple sclerosis, since this should represent the ultimate consequences of destructive pathological changes, whether they occur within lesions or in the normal appearing tissues. This interest has been stimulated by the development of a range of methods that are sensitive and reproducible in measuring even small changes in tissue volume, and by the observations of more robust correlations of the patient's level of disability with measures of atrophy than had been previously obtained by examining correlations with lesion volume or activity measures.

Stimulated by these emerging data, an International Workshop on the Measurement of Atrophy in Multiple Sclerosis was held in London, UK, in December 2000. The Workshop covered: (i) the pathological basis of atrophy in multiple sclerosis; (ii) methods for measuring tissue loss and anticipated developments; (iii) clinical studies in multiple sclerosis; (iv) the present role for atrophy measurement in multiple sclerosis research and clinical trials; and (v) future research and clinical applications. This review reflects the content of the Workshop and the authors' perspectives.

Pathological basis of atrophy in multiple sclerosis

It is logical to assume that loss of myelin, the classical feature of multiple sclerosis lesions, would result in a loss of tissue volume, and hence contribute to the atrophy observed in multiple sclerosis. However, other tissue elements within the brain also contribute to tissue bulk. Even in the white matter, where myelin is most concentrated, there are axons, glial cells, blood vessels, blood and tissue fluid, all of which contribute to the white matter volume. In bovine white matter, myelin constitutes 50–60% of the dry weight but only

25% of the wet weight—the latter figure is a more realistic assessment of the contribution of myelin to white matter volume *in vivo*. This is supported by experimental evidence: an estimation of the proportions in normal human white matter due to different tissue elements has been made by point counting from a stereological grid applied to an electron microscopic image published in an atlas (Adams, 1989; M. Esiri, personal communication); the proportion due to axons was 46%, with smaller contributions from myelin (24%), glial cells (17%) and other elements including blood vessels, blood and tissue fluid (13%).

The same assessment of four multiple sclerosis plaques showed no myelin in three demyelinated plaques and 23% myelin in the fourth plaque, which exhibited remyelination. The proportion of the plaques containing axons varied from 6 to 50%, whereas glial and other cellular elements contributed 30–80% of the plaque volume (M. Esiri, personal communication). This evidence that a marked degree of axonal loss can occur in some chronic multiple sclerosis plaques is supported by previously published studies in both new and old literature (Charcot, 1868; Lassmann *et al.*, 1994). Signs of axonal damage and loss have also been observed in active inflammatory lesions using amyloid precursor protein staining (Ferguson *et al.*, 1997) or by confocal microscopy, which has revealed axonal spheroids and transections (Trapp *et al.*, 1998).

A consequence of axonal loss in a lesion is Wallerian degeneration along the fibre pathways that traverse it. Axonal loss in lesions may therefore cause atrophy by two mechanisms: tissue loss within the lesion *per se*, and Wallerian degeneration in related fibre pathways. Given the large proportion that axons contribute to white matter volume, and evidence for considerable axonal damage in multiple sclerosis, axonal loss seems likely to be an important contributor to the atrophy observed in multiple sclerosis.

Detailed studies in the spinal cord and lateral geniculate body in multiple sclerosis have shown a proportionately greater loss of small axons with relative preservation of larger ones (Ganter *et al.*, 1999)—smaller axons may be more susceptible to inflammatory mediators such as nitric oxide. In multiple sclerosis corpus callosum, both decreases in area and in axonal density have been reported (Evangelou *et al.*, 2000a). This emphasizes that the full extent of axonal loss in multiple sclerosis may be underestimated by simply measuring tissue volume loss. Investigators have found a relationship between axonal density in the normal appearing corpus callosum and cerebral hemisphere white matter lesion load, suggesting that the abnormalities seen in the NAWM using a number of MR techniques may in part reflect Wallerian degeneration (Evangelou *et al.*, 2000b). It also has been suggested that an important part of the axonal damage found in the spinal cord is secondary to cerebral disease resulting in Wallerian degeneration (Lovas *et al.*, 2000).

Although there has been less pathological investigation of multiple sclerosis grey matter, plaques in grey matter are commonly recognized with careful post mortem examination

(Kidd *et al.*, 1999; Bø *et al.*, 2000). Such plaques are rarely seen on conventional MR sequences, but loss of myelin or axons within them may also contribute to the global brain atrophy seen in multiple sclerosis. It is not known whether there are decreases in the number or size of neuronal cell bodies in multiple sclerosis.

The contribution of glial cells to tissue bulk in normal white matter is less than that of myelin and axons. In plaques, an increase in astrocytes and activated microglia appears to outweigh the decrease in oligodendrocytes, and glial cells appear to constitute an increased proportion of the lesion volume compared with normal white matter. The net effect of glial pathology on tissue volume is uncertain—potentially, reactive astrocytosis has a positive mass effect and isomorphic gliosis a negative or neutral effect.

Changes in tissue water content can have an important effect on global measures of brain volume. Changes as much as 30–40 ml have been observed after dialysis in patients with renal failure (Walters *et al.*, 2001). In multiple sclerosis, acute lesions are associated with breakdown of the blood–brain barrier, inflammation and vasogenic oedema. Visible swelling is sometimes seen on MRI in these lesions in the brain, spinal cord or optic nerves. The extent and volume of gadolinium-enhancing lesions in the brain may therefore affect overall brain volume, and is a variable to consider when assessing and interpreting tissue volume changes. This is especially relevant in therapeutic trials or natural history studies, where therapies are known to have an anti-oedema or anti-inflammatory effect. There is evidence that chronic plaques also have a degree of blood–brain barrier impairment (Broman, 1964; Silver *et al.*, 2002) and that the water content of multiple sclerosis NAWM is increased (Tourtellotte and Parker, 1968). The potential for more subtle variations in the blood–brain barrier and brain water content to affect volume measures is therefore not excluded.

Methodological requirements: general issues

The optimal technique for measuring tissue volume should be reproducible, sensitive to change, accurate and practical to implement. Of these requirements, accuracy is least easy to verify, and small errors of accuracy are probably insignificant in the study of atrophy, as long as they are constant between subjects and over time, and do not impact on the sensitivity and precision of the technique. The two distinct components involved in measuring tissue volumes are data acquisition and data analysis.

Data acquisition

The principal aspects of data acquisition that influence the efficacy of volume measurements (and therefore any atrophy measurement derived from them) are image resolution in three dimensions and image contrast. The desirability of high resolution scans to reduce partial volume errors means that 3D (volumetric) acquisitions are attractive, although both 2D (Molyneux *et al.*, 2000) and 3D (Liu *et al.*, 1999) sequences

have been used successfully to derive volume measures in the CNS. 3D sequences also allow more accurate re-slicing of datasets when registering volumes between serial acquisitions.

The choice of image contrast is also important in reliable volume measurements. The contrast weighting chosen depends on the study's aims and, to an extent, on the choice of analysis technique. For whole-brain atrophy measurements, segmentation of the brain is necessary, meaning that suppression of CSF, for example using fluid-attenuated inversion recovery (FLAIR) sequences, is desirable to generate a sharp distinction in signal between cerebral and extra-cerebral matter. For this reason, the most widely used 3D sequence is a T₁-weighted gradient echo, with or without added CSF suppression, the latter provided by an inversion recovery pre-pulse. Such an acquisition allows voxels with dimensions of the order of 1 × 1 × 1 mm, and can be completed in ~10 min.

Studies focusing on the measurement of atrophy of individual structures within the CNS may require alternative imaging strategies. For instance, a study of white matter atrophy requires good contrast between white matter, grey matter, CSF and possibly lesions. Segmentation may be aided by multiple-contrast (multi-spectral) acquisitions, typically T₁-weighted, T₂-weighted and proton-density-weighted data acquisitions (all of which should be of the same resolution and in register).

Data analysis methods

Manual and semi-automated methods

Manual outlining or linear measurements of tissues provide simple approaches to measuring changes in volume or dimensions over time. For these methods to be reproducible, an experienced observer is required, who is familiar with normal neuroanatomy and the appearance of CNS tissues and pathology on the sequences employed. Manual segmentation of the whole brain from multislice or 3D image sets is laborious and rarely used; it is more practical in outlining or defining small structures or selected regions, for example outlining the cross section of the spinal cord or measuring third ventricle dimensions. Moderately reproducible measures of area or width have been obtained and significant atrophy (or enlargement of the third ventricle) has been reported in multiple sclerosis (Simon *et al.*, 1999).

Techniques requiring a high degree of manual input typically have the following advantages: they are methodologically simple; little specialist software is required; and the segmentation results relate well to the operator's perceptions of the limits of a particular structure. Typical disadvantages include: they are prone to operator bias; the precision is low when compared with many automated techniques; and analysis times can be long.

Semi-automated methods can supplement or replace manual outlining methods, resulting in improved speed and

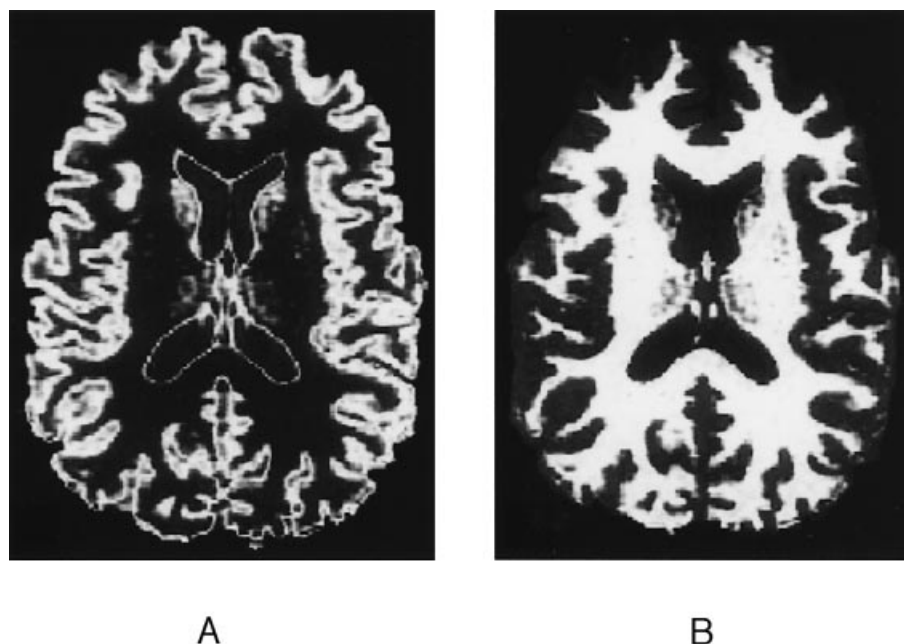


Fig. 1 Results of segmentation into grey matter (**A**) and white matter (**B**), using a fuzzy C-means method [reprinted from Pham and Prince (1999b), copyright 1999, with permission from Elsevier Science].

reproducibility. The use of regional segmentation algorithms, such as seed growing (Rovaris *et al.*, 2000), and local edge detection algorithms and contouring (Plummer, 1992) has assisted in outlining lesions and also small structures such as the spinal cord, optic nerves and ventricles (Losseff *et al.*, 1996a; Lycklama *et al.*, 1998). Measurement reproducibility improves from ~3–5% for manual outlining to 1–3% for semi-automated approaches for measuring spinal cord and ventricles. Liu *et al.* (1999) and Edwards *et al.* (1999) have applied semi-automated stereological techniques to assess atrophy in the corpus callosum, cerebellum and brainstem of multiple sclerosis patients.

Automated segmentation

Analysis methods should be automated as much as possible to provide good reproducibility and reduced reliance on time-consuming operator input. Many automated methods exist for segmentation (and thus volume measurement) of the brain parenchyma, and for brain white and grey matter. Fewer proven approaches exist for automated segmentation of individual brain structures or for the spinal cord or optic nerve.

Both single contrast acquisitions (Chard *et al.*, 2001) and multi-spectral data (Udupa *et al.*, 1997; Ge *et al.*, 2000a) have been utilized for whole-brain segmentation and volumetry. Usually the difference in signal intensity between brain parenchyma and CSF on a single contrast acquisition is enough to drive the segmentation process. Methods for removing any residual signal from scalp in the resulting

segmentation include automated detection of the intracranial contour (Zijbendos *et al.*, 1994) and segmentation based on probabilistic spatial tissue distributions (Ashburner and Friston, 2000).

The segmentation of grey and white matter volumes may again be accomplished utilizing either single-contrast or multi-spectral data, although additional sophistication is required to separate the two tissue types. Methods include statistical parametric mapping (SPM)-based segmentation (Wellcome Department of Functional Neurology, Institute of Neurology, London, UK; Ashburner and Friston, 2000) and the fuzzy C-means algorithm, as described by Pham and Prince (1999) (Fig. 1). A method for determining the regions of tissue occupied by multiple sclerosis lesions is necessary if mis-classification of these regions is to be avoided and accurate tissue volumes are to be obtained. Automated segmentation of multiple sclerosis lesions has been investigated by a number of researchers (Zijbendos *et al.*, 1994, 1996; Udupa *et al.*, 1997; Guttmann *et al.*, 1999; Kikinis *et al.*, 1999); if total lesion load is small, manual segmentation may be a reasonable alternative. In the case of negligible lesion volume, lesion segmentation may be unnecessary.

Automated measurement of atrophy in individual CNS regions has received relatively little attention to date. De Carli *et al.* (1992) used a trace-then-threshold approach to delineate specific structures. This method requires significant operator input, and is therefore subject to many of the drawbacks of other semi-automated methods. Recent work has seen progress towards

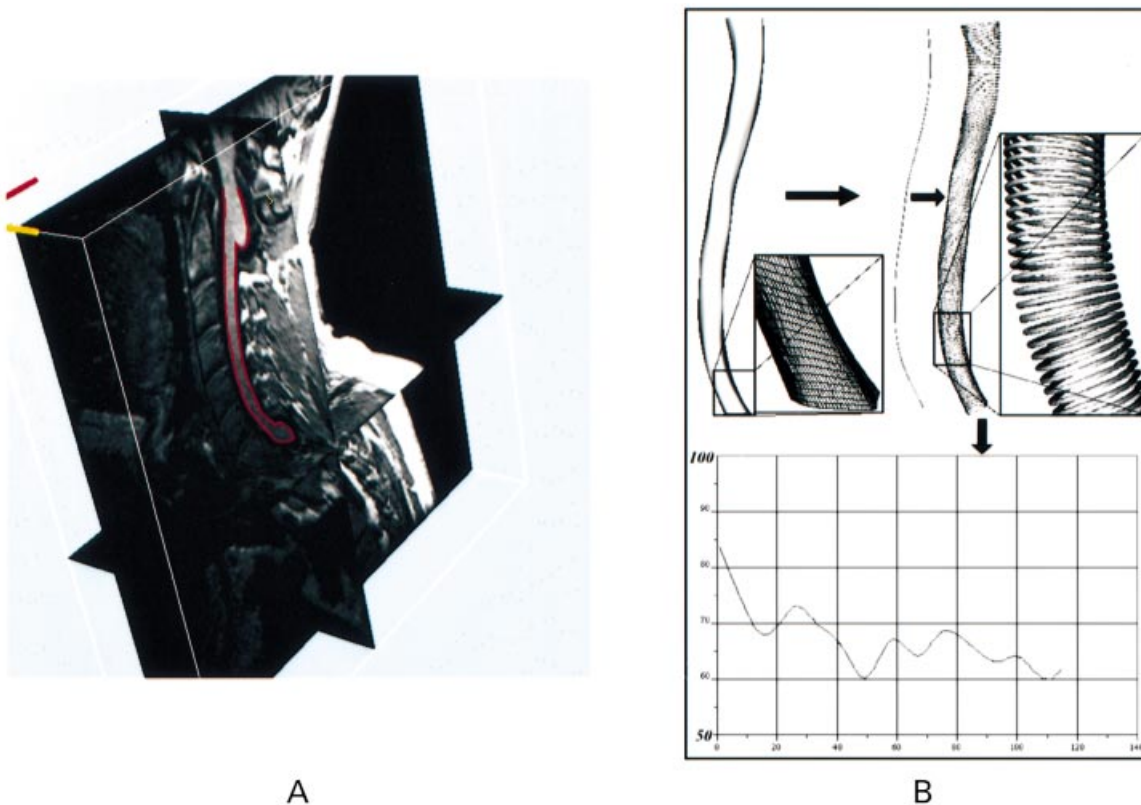


Fig. 2 Segmentation of the spinal cord using a B-spline active surface model. (A) 3D visualization of cord plus segmentation (red). (B) Steps in the generation of cross-sectional areas of cord as a function of position: identification of cord surface; identification of medial axis of cord; calculation of cross-sections of cord; plot of area as a function of position (courtesy of O. Coulon, Computer Sciences Department, UCL, London, UK).

highly automated segmentation of the optic nerve and spinal cord (Coulon *et al.*, 2000, 2001), allowing the cross-sectional area of the structure to be defined as a function of position (Fig. 2).

Registration- and boundary-based methods

As atrophy is the measurement of change in volume, measurements of absolute volumes at separate time points are not necessarily needed; information may be obtained by looking for *differences* between serial scans. Non-linear registration of such scans produces deformation fields that yield information concerning regional and global atrophy (Freeborough and Fox, 1998). Statistical analysis of the differences between groups of patients and normal controls may identify regions of the brain subject to atrophy across a patient population (Ashburner and Friston, 2000). Rigid body registration, on the other hand, can be used to track the displacement of the surface of the brain during atrophy. Using the inner table of the skull as standard of reference, Freeborough and Fox (1997) have developed a method to quantify atrophy with high sensitivity in many neurological diseases.

The use of relative volumes

Comparisons between groups of patients are confounded by the presence of substantial inter-subject variations in head size that can mask differences attributable to atrophy. Normalizing the volumes of brain, specific tissues and structures to head size may be successful in reducing these confounding effects (Mathalon *et al.*, 1993; Whitwell *et al.*, 2001). Relative volumes also remove variability in volume data due to scanner instability, as all structures in the image will experience the same amount of artefactual scaling due to such effects.

A number of normalization methods have emerged: the scalp (Freeborough and Fox, 1997), total intracranial capacity determined by the sum of the volumes of grey matter, white matter and CSF (Guttmann *et al.*, 2000a; Chard *et al.*, 2002), and the sum of the brain, and ventricular and sulcal CSF (Rudick *et al.*, 1999), have all been used to create volumes against which normalization can be performed.

Non-uniformity correction

Reliable segmentation is, in many cases, aided by the application of image non-uniformity correction (De Carli

Table 1 Methods used for measuring atrophy

Location/generic methodology	Method	Advantages	Disadvantages
Brain/linear and regional measures	Third ventricle width	Easy to implement; rapid analysis; standard acquisition methods; enables targeting of eloquent regions	Limited anatomical scope; may miss subtle effects; may exhibit user bias; high user input
	Third ventricle volume Brain width Corpus callosum width Volume on central brain slices Stereology CSF volumes		
Brain/whole brain segmentation approaches	BPF WBR BICCR Fuzzy connectedness Probabilistic segmentation (SPM) TDS SIENAX MIDAS	Increased automation reduces user bias and user input; generally higher measurement precision	Complex analysis methods; possibly more complex acquisition schemes
	Brain/registration-based methods		
Spinal cord	Voxel-based morphometry SIENA Manual outlining Semi-automated outline of 3D axial images	Straightforward Precise	Possible user bias; high user input
	Automated whole cord volume measurement		
Optic nerve	Manual outlining Semi-automated outline of 3D axial images	Straightforward Precise	Possible user bias; high user input
	Automated whole nerve volume measurement		

See Appendices I and II for references and detailed description. TDS = template-driven segmentation; MIDAS = Medical Image Display and Analysis Software.

et al., 1992; Pham and Prince, 1999; Ashburner and Friston 2000). Such non-uniformity is caused by the spatially varying radiofrequency electromagnetic fields associated, to varying degrees, with all imaging coils. This typically causes signal drop-off in the cerebellum and brainstem when using head coils. Similarly, studies of the spinal cord using phased array coils may require non-uniformity correction (Losseff *et al.*, 1996a). Methods include fitting a smoothly varying bias field to the image data (Sled *et al.*, 1998). Alternatively, data acquired from phantoms within the coil in question, and using the sequence used for the atrophy measurement, may be used to correct the images (Losseff *et al.*, 1996a).

Specific methods

A considerable number of specific methods for measuring tissue volumes in the brain, spinal cord and optic nerves have been developed and applied in multiple sclerosis. They are

discussed further in Appendices I and II. A summary of the benefits and limitations of current methods for brain volume measurement is provided in Table 1.

Clinical studies

Factors causing short-term changes in brain volume

A strong appeal of measuring atrophy is the potential to quantify, in the longer term, the loss of CNS tissue, especially axons, which are the likely substrate of irreversible and progressive functional deterioration. However, other factors, most notably changes in brain water content, may contaminate the measurement of true tissue loss, at least in the short term. In one study, up to 20 serial scans were performed over 1 year, and short-term changes in brain parenchymal fraction (BPF) in individuals were larger than the average change seen over the whole 12-month period (Guttmann *et al.*, 2000a).

Greater short-term fluctuations were seen in relapsing–remitting than in progressive patient cohorts. White matter lesion volume was also correlated with BPF in relapsing–remitting but not progressive patients. These observations are consistent with shorter term fluctuations in brain volume being related to variation in the extent of inflammatory oedematous lesions, which are more commonly found in the relapsing–remitting stage of multiple sclerosis (Tubridy *et al.*, 1998); they also caution against using atrophy to evaluate short-term disease progression in individual patients.

The effect of steroids on brain fraction volume (BFV) has been studied in a group of 26 patients with relapsing–remitting multiple sclerosis who were followed with monthly brain MRI (Rao *et al.*, 2002). Patients were given courses of 3–5 days of high-dose intravenous methylprednisolone (IVMP) as required to treat acute relapses. BFV was determined from T₁-weighted sequences using an adaptive Bayesian segmentation using K-means clustering (Goldszal *et al.*, 1998). A total of 56 courses of steroids were given during the prolonged follow up, 40 while taking β -interferon and 16 while on no other disease-modifying treatment. BFV was stable for the 3 months prior to IVMP, but when all 56 events were combined, there was a significant reduction in BFV at month 1 and 2 following IVMP ($P = 0.009$, non-parametric Wilcoxon test). When considered according to prior treatment status a significant reduction in BFV was seen only at month 1 ($P < 0.04$ for both untreated and β -interferon-treated subjects). This short-term decrease in BFV after IVMP was accompanied by a significant reduction in the number of contrast-enhancing lesions at months 1 and 2. The study revealed less variability in BFV measures during periods on β -interferon compared with periods untreated, probably due to fewer inflammatory/oedematous lesions. In long-term treatment trials, measurement of brain volume should probably be avoided in the 2 months after IVMP. Other therapies with anti-inflammatory effects, including β -interferons, are also likely to reduce brain volumes in the short term.

Relationship of brain atrophy to disability and other MR measures of axonal loss

Several studies have investigated the relationship between brain atrophy and either disability or putative MR measures of axonal loss. In a study of 137 patients, disability was measured using the Multiple Sclerosis Functional Composite (MSFC) scale, and correlated with a normalized measure of atrophy, defined as the brain volume fraction (BVF), and with T₁ and T₂ lesion load (Kalkers *et al.*, 2001a). Stronger correlations were observed with the BVF ($r = 0.36$) than with either lesion load measure ($r = -0.21$ for T₁ and $r = -0.22$ for T₂ load). These correlations are still low, which may partly relate to variations in image acquisition sequences and to the heterogeneous patient features, which exhibited a wide variation in disease duration and included a group with a

primary progressive course. When only relapse onset patients were considered, the correlation between MSFC and BVF improved while the relationship of MSFC with T₂ and T₁ lesion load decreased.

A second study investigated a normalized measure of brain volume [the brain intracranial capacity ratio (BICCR); Collins *et al.*, 2001] in conjunction with MR spectroscopy examination of the axonal marker, *N*-acetyl aspartate (NAA), which was measured from a large voxel acquired from the cerebral hemispheres above the level of the lateral ventricles (Collins *et al.*, 2000). Both increasing atrophy and decreasing NAA were correlated with increasing disability. The most marked gradient of reduction in NAA was observed in the lower expanded disability status score (EDSS) patients with relapsing–remitting disease. A larger gradient of reduction in BICCR was seen in more disabled secondary progressive patients. The investigators proposed that in the early multiple sclerosis, there is a potentially reversible dysfunction of axons, manifested as a substantial decrease in NAA with a lesser decrease in brain volume, but with continued progression in disability, irreversible axonal loss occurs, manifested as a greater decrease in BICCR.

A third study in secondary progressive multiple sclerosis reported a correlation between cerebral volume and NAA (Coles *et al.*, 1999), which suggests a direct association between atrophy and axonal loss. Two further studies investigated the relationship between T₁ hypointense lesion load and cerebral atrophy in relapsing–remitting or secondary progressive multiple sclerosis (Paolillo *et al.*, 2000; Sailer *et al.*, 2001). Chronic T₁ hypointense lesions are associated with more marked axonal loss than lesions that are T₁ isointense (van Walderveen *et al.*, 1998). In both studies, decreased cerebral volume was correlated significantly with T₁ hypointense lesion load but not with total T₂ load.

Taken together, these studies provide consistent evidence of a relationship between atrophy, disability and axonal loss.

Clinical subgroup studies: clinically isolated syndromes

Many patients with clinically isolated syndromes (CIS) affecting the optic nerve, brainstem or spinal cord already have disseminated cerebral white matter lesions typical of multiple sclerosis, and develop clinically definite multiple sclerosis on follow up (O’Riordan *et al.*, 1998). A cohort of 55 CIS patients (Dalton *et al.*, 2001) has been followed over a 1-year period to investigate ventricular enlargement. Significant enlargement was seen in the 18 patients who developed clinical multiple sclerosis (i.e. experienced a second clinical relapse), and in the 40 patients with abnormal brain MRI at presentation. There was no change in ventricular volume in the 15 patients with normal imaging. Thus, brain atrophy occurs at the earliest clinical stage of multiple sclerosis, and is associated with short-term clinical disease activity.

A 2-year placebo-controlled trial of β -interferon, which demonstrated a delay in the primary outcome (time to development of clinically definite multiple sclerosis; Comi *et al.*, 2001), also demonstrated significant brain atrophy in the placebo arm over the course of the trial (Filippi *et al.*, 2001). Mean brain volume at entry in the placebo group (135 patients) was 1210 ml and mean loss over the 2 years was 15 ml, which corresponds to an annual reduction of $\sim 0.6\%$. There were significant, but modest, negative correlations of change in brain volume with the number of enhancing lesions ($r = -0.25$, $P = 0.01$) and the number of new T₂ lesions ($r = -0.36$, $P = 0.001$) in the placebo group. Even weaker correlations were evident in the interferon-treated cohort, who nevertheless also exhibited significant atrophy over the study period. In addition to CIS, subjects were included who had polysymptomatic or multifocal presentations, and all patients were required to have at least four lesions on T₂-weighted MRI at entry. The weak relationship between atrophy and lesion evolution in this early stage of disease suggests that atrophy is evolving by mechanisms that are at least partly independent of lesions.

A cross-sectional study performed a mean of 19 months following a single attack of optic neuritis revealed a mean reduction in optic nerve area of $\sim 15\%$ (11.2 mm² in affected versus 12.9 mm² in unaffected nerves) (Hickman *et al.*, 2001). Serial follow up in a subgroup of patients over 1 year revealed a further, significant increase in atrophy (from 11.1 to 10.2 mm², $P = 0.01$) in the affected nerve (Hickman *et al.*, 2000). These observations indicate that inflammatory demyelinating lesions result in atrophy that evolves for months to years. Progressive atrophy may be due to continued axonal loss in the lesion, and may reflect Wallerian degeneration in the white matter tract in which the lesions arises, in this case the optic nerve (Stoll *et al.*, 1989).

Clinical subgroup studies: relapsing–remitting multiple sclerosis

Using the whole brain ratio (WBR) method, brain atrophy was assessed over 2 years in a placebo-controlled trial of β -interferon in relapsing–remitting multiple sclerosis (Jones *et al.*, 2001). Atrophy measures were available in 519 patients, 172 of whom were on placebo. Significant brain atrophy was seen in the total cohort over 2 years: the mean WBR decreased by 1.4%. The baseline WBR was weakly correlated with baseline age, EDSS and T₂ load, with average r -values of 0.3. Two-year WBV correlated weakly with relapse rate, EDSS change, T₂ lesion load change and T₂ lesion activity, with r -values ranging from 0.2 to 0.4. No difference in the rate of atrophy was seen between treatment arms.

Cerebral atrophy has been evaluated from 52 relapsing–remitting patients for 6 months prior to and 24 months following β -interferon treatment and correlated with other MRI lesion and clinical parameters (Gasparini *et al.*, 2002). During the 2 years of treatment there was a significant

reduction of brain volume (mean -2.2%) that was correlated with the mean number of enhancing lesions on monthly scans during the 6 months pre-treatment ($r = -0.32$, $P = 0.02$). Thus, the extent of inflammation over a relatively short period of time contributes to atrophy that develops later and over a longer period of time. During the 2 years of treatment, 26 patients exhibited significant atrophy and 26 did not; in the former group, 13 experienced an increase in disability, whereas in the latter group only three became more disabled. This confirms other studies in showing a link between increasing atrophy and disability (Losseff *et al.*, 1996b; Paolillo *et al.*, 1999).

In a 2-year placebo controlled trial of β -interferon in relapsing–remitting multiple sclerosis, atrophy was measured from yearly scans using BPF (Fisher, 1997). The mean BPF decrease was similar in both arms in year 1 (-0.763% in β -interferon and -0.699% in placebo arm), but was smaller in the β -interferon arm in year 2 (-0.233% versus -0.521%) (Rudick *et al.*, 1999). The changes in BPF during this 2-year period showed little or no correlation with lesion measures. Prolonged follow up of 134 out of 172 of the placebo cohort from this trial assessed the longer term relationship between BPF and clinical change (Fisher *et al.*, 2000a, b). Many received disease-modifying therapies after the original trial finished. Whereas the mean BPF at entry was 0.832, it was 0.807 at follow up, a mean of 8 years later. The average yearly reduction in BPF over 8 years was 0.38%, which is about a half the rate seen during the 2 years of the original trial. This observation could be related to cohort selection bias, disease natural history, treatment effect (most patients went on to disease-modifying treatments after the original trial ended) or changes in scanner/image data. Comparison of patient quartiles based on change in BPF over the first 2 years revealed that 57% of those in the fourth quartile (greatest atrophy) had progressed to EDSS of 6 or more at follow up, whereas only 13% of those in the first quartile (least atrophy) had reached this level of disability.

A 9 month, placebo-controlled trial of glatiramer acetate in 239 relapsing–remitting multiple sclerosis revealed a mean 0.7–0.8% reduction in central cerebral volume (Rovaris *et al.*, 2001). Whereas an earlier, small study (27 patients) over 2 years suggested that the rate of atrophy is reduced in glatiramer acetate-treated patients versus controls (Ge *et al.*, 2000b), no significant group differences during the 9 month study were revealed in the larger study (Rovaris *et al.*, 2001). Follow up to 18 months with all patients converted to active treatment from month 9 revealed a trend to less atrophy in those treated with glatiramer acetate from onset. The study showed a weak association between enhancing lesion numbers and atrophy ($r = -0.21$, $P = 0.002$).

Clinical subgroup studies: progressive forms of multiple sclerosis

Atrophy is seen in both the brain and spinal cord in secondary and primary progressive multiple sclerosis. The most marked

Table 2 Clinical studies of brain and spinal cord atrophy in multiple sclerosis

Condition	Reference [first author (year)]	No. patients studied	Atrophy detected	Method	Comments
Clinically isolated syndromes	Brex (2000)	17	+	Ventricular volume	Small, significant increase in 1 year
	Dalton (2001)	55	+	Ventricular volume	Significant increase in those developing multiple sclerosis or with T ₂ lesions
	Filippi (2000)		+	Whole brain volume	Significant atrophy over 2 years; no effect of BIFN
	Brex (2001) Hickman (2001)	43 17	+	C2 area Optic nerve area	Significant atrophy at presentation Atrophy increases over months to years after an attack of optic neuritis
Relapsing–remitting multiple sclerosis	Losseff (1996b)	29	+	Central cerebral volume	Strong correlation with disability but not with lesions over 18 months
	Gasperini (2002)	52	+	Regional brain volumes	Enhancing lesions over 6 months correlated with atrophy in next 2 years
	Rudick (1999)		+	BPF	Slowing of atrophy during year 2 of BIFN therapy
	Collins (2000)		+	BICCR	Mild atrophy; marked decrease NAA
	Rovaris (2000)	50	+	Central cerebral volume	Significant atrophy over 1 year
	Sandaine (2000)	24	+	Fuzzy connectedness	No correlation with enhancing lesions
	Wolinsky (2000)	30	+	Normalized CSF volume	% increase CSF volume/year
	Rovaris (2001)	239	+	Central cerebral volume	No effect of glatiramer acetate over 9 months
	Chard (2001)	28	+	BPF (SPM99)	Grey and white matter atrophy in early RR multiple sclerosis
	Kalkers (2001)		+	BVF	Atrophy correlates better with disability than lesion load
Secondary progressive multiple sclerosis	Paty (2001)	519	+	WBR	No effect of BIFN on atrophy
	Losseff (1996a)	15	+	C2 area	Marked atrophy correlated with disability
	Molyneux (2000) Collins (2000)	95	+	Central cerebral volume BICCR	No effect of BIFN on atrophy More atrophy than RR multiple sclerosis and increases with increasing disability
	Ge (2000)	36	+	Fuzzy connectedness	Correlated with disability in SP but not RR multiple sclerosis
	Filippi (2000) Wolinsky (2000)	Many	+	Whole brain volume Normalized CSF volume	No effect of cladribine on atrophy Higher CSF volume in SP than PP multiple sclerosis
	Kalkers (2001)		+	BVF	Atrophy correlates better with disability than lesion load
Primary progressive multiple sclerosis	Losseff (1996a)	15	+	C2 area	Correlation between C2 area and disability
	Stevenson (2000)	167	+	Central cerebral slices	Progressive atrophy over 1 year
	Barkhof (2000)	16	+	C2 area	Riluzole may slow cord atrophy
	Filippi (2000) Kalkers (2001)		+	Whole brain volume BVF	No effect of cladribine on disability Atrophy present

+ = atrophy detected; BPF = brain parenchymal fraction; BIFN = β -interferon; BICCR = brain intracranial capacity ratio; BVF = brain volume fraction; WBR = whole brain ratio; SPM = statistical parametric mapping; RR = relapsing–remitting; SP = secondary progressive; PP = primary progressive; C2 area = cross-sectional area of cervical cord at C2 level.

atrophy occurs in secondary progressive disease and correlates with disability (Losseff *et al.*, 1996a; Kalkers *et al.*, 2001). In primary progressive multiple sclerosis, significant atrophy of brain and cord over 1 year was evident in a large cohort of primary progressive patients drawn from six European centres (Stevenson *et al.*, 2000). Change in cerebral volume over 1 year correlated only weakly with change in T₁ and T₂ brain load, suggesting that the measurement of atrophy

provides complementary information. In cross-sectional analysis, there was a correlation between neuropsychological impairment and cerebral volume (Camp *et al.*, 1999).

A small study of 16 patients with primary progressive multiple sclerosis evaluated riluzole, a neuroprotective agent (which acts as a glutamate antagonist), using change in cord area as a putative measure of progressive axonal loss (Kalkers *et al.*, 2001b). Patients were followed untreated for 1 year

followed by 1 year on riluzole. During the pre-treatment phase, there was a 2% reduction in mean cord area, whereas during treatment the cord area was stable (mean decrease of 0.2% only), but the difference was not significant. This preliminary study indicates the potential of using tissue volume measures in larger cohorts to study the efficacy of neuroprotective agents.

Recent therapeutic trials have evaluated the effect of three immunosuppressive, immunomodulatory agents on atrophy in progressive forms of multiple sclerosis. First, in a placebo-controlled trial of β -interferon in secondary progressive multiple sclerosis, there was steady decrease of cerebral volume in the placebo group with a mean reduction of 3.9% after 3 years. The interferon group also exhibited significant reduction reaching a mean of 2.9% at 3 years; the difference between the groups was not significant (Molyneux *et al.*, 2000). There were significant correlations between increasing cerebral atrophy and deteriorating neuropsychological function in the placebo ($r = 0.32$, $P = 0.002$) but not the interferon arm ($r = 0.14$, $P = 0.31$). Some of the atrophy in the interferon group may have been due to the anti-inflammatory effect of interferons leading to a decrease in brain water content because: (i) reduction in the first 6 months was greater in the interferon than the placebo group; and (ii) those without enhancing lesions, in whom volume measures should be less susceptible to anti-inflammatory effects, exhibited more atrophy in the placebo group. Anti-oedema effects might also have uncoupled the relationship between cerebral volume and neuropsychological function in the interferon arm. Notwithstanding, the effect of interferon on atrophy, if present, would appear to be modest.

Secondly, an open study of 27 secondary progressive patients, treated with the anti-lymphocyte antibody Campath 1-H, revealed a marked reduction in enhancing lesions over 18 months, whilst an increase in disability accompanied by increasing cerebral atrophy was seen in 50% patients (Coles *et al.*, 1999). The increase in cerebral atrophy was related to the volume of enhancing lesions accumulated over 3 months prior to the start of treatment. This observation, as in relapsing–remitting multiple sclerosis (Gasperini *et al.*, 2002), suggests that inflammation leads to later atrophy.

Thirdly, a placebo-controlled trial of cladribine, an immunosuppressive agent, in primary and secondary progressive multiple sclerosis revealed marked suppression of enhancing lesions but no effect of therapy on progressive disability or brain atrophy (Filippi *et al.*, 2000; Rice *et al.*, 2000).

Current role for atrophy measurement in multiple sclerosis

The most compelling reason for measuring atrophy is that it provides a measure of axonal loss, which, if progressive, is likely to result in irreversible disability. The observation that almost a half of normal white matter bulk consists of axons

implies that significant atrophy in multiple sclerosis does indeed reflect axonal/neuronal loss. This conclusion is supported by the significant correlations observed between atrophy and other MR axonal/neuronal markers. Nevertheless, loss of myelin, glial and vascular elements may contribute. Furthermore, tissue bulk may be added by astrogliosis and other inflammatory cells, and changes in brain water due to inflammation, therapeutic effects or other intercurrent factors can all affect brain volume measures. With these caveats, the quantification of sustained and irreversible atrophy of CNS structures over a reasonable period, not less than 6 months and preferably at least 12 months, is a largely reliable surrogate for axonal/neuronal loss.

There is a large range of techniques available for the measurement of tissue volumes. Many methods satisfy the essential requirements of being reproducible, sensitive to change and stable over time. Accuracy is less easy to evaluate as there is no gold standard; the experience of markedly different absolute changes in brain volumes using different acquisition and segmentation methods indicates that not all methods are accurate. This makes it very difficult to combine or compare results from studies that use different methods. Accuracy is an important consideration in natural history studies but is less critical in treatment trials where a treatment and control arm are being compared.

In deciding which methodological approach to use for a study of atrophy, several points should be borne in mind.

- (i) The acquisition sequence should be kept constant (in terms of repetition time, echo time, echo train length, voxel size, slice thickness, field strength, etc.).
- (ii) A T_1 -weighted sequence is preferred as it gives good contrast between parenchyma and CSF. CSF suppression with an inversion recovery prepared pulse may be useful.
- (iii) 3D acquisition sequences have greater resolution and sensitivity to small changes and are the sequence of choice for measuring spinal cord size. Some brain studies report better reproducibility using multislice 2D acquisitions, possibly due to shorter acquisition times; both 3D and 2D sequences appear adequate in the brain.
- (iv) The segmentation algorithm should be as automated as possible to maximize objectivity, reproducibility and efficiency.
- (v) Acceptable automated segmentation algorithms currently available include multispectral classification and boundary-based methods to calculate differences in brain volume. Such automated algorithms routinely include image homogeneity correction and noise filtering.
- (vi) The measurement may be of an absolute volume or of a normalized volume, most commonly a fraction of brain volume related to total intracranial volume, e.g. BPF, WBV, etc. The latter display less inter-subject variability and are superior for cross-sectional studies. Both approaches are sensitive to change over time, but normalized volumes may still be preferred, as they remove scaling errors.

(vii) A quality assurance programme should be in place to confirm the stability of measurements over time. This should include each centre in a multicentre study, and may involve repeat scanning of healthy controls or suitable phantoms. It is especially important before and after a major scanner upgrade (hardware and software), since volume measurements are susceptible to changes in scanner performance.

(viii) Serial studies of atrophy should try to control for factors such as age, gender, fluid intake and nutritional status, alcohol intake, or anti-oedema therapies, all of which might influence brain volume measures.

(ix) When multiple sclerosis patients are being compared with controls, it is essential that the latter are appropriately age- and gender-matched.

From the studies of multiple sclerosis, several consistent observations have emerged.

(i) Accelerated atrophy occurs in the brain and spinal cord in multiple sclerosis, at all stages of the disease from presentation with CIS to advanced progressive forms (Table 2).

(ii) Despite variation between individuals and techniques, a reduction in brain volume by 0.6–1% per annum (p.a.) is seen in most patient cohorts. In comparison, estimated rates of reduction for healthy controls vary from 0.1–0.3% p.a. (Gur *et al.*, 1991; Coffey *et al.*, 1992; Cowell *et al.*, 1994; Pfefferbaum *et al.*, 1994; Courchesne *et al.*, 2000).

(iii) Greater amounts of atrophy are seen when investigation is limited to central brain slices that concentrate on the periventricular region, probably because of the greater pathological involvement of these regions.

(iv) Measures of brain and spinal cord atrophy correlate better with disability than do changes in MRI lesion markers (enhancing, T₂ and T₁ lesion loads). Atrophy is most evident in progressive forms of multiple sclerosis.

(v) The significant correlations found between brain volume and other MR markers of axonal loss (reduced NAA, T₁ hypointense lesions) indicate that axonal loss is an important component of atrophy.

(vi) Inflammatory/demyelinating lesions are followed by progressive atrophy for months or even years. This suggests a potential time window for salvage of axons by an appropriate therapy following acute inflammation. A better understanding and targeting of the mechanisms of post-inflammatory axonal loss (e.g. failure of remyelination) is needed.

(vii) The correlations between lesion load and atrophy are only moderate, suggesting that mechanisms independent of lesions also cause atrophy.

(viii) Whereas a number of immunosuppressive/immunomodulatory therapies, in particular interferons, substantially suppress inflammation (inferred from the extent of enhancing lesions), their effects on atrophy are less evident or undetectable.

(ix) Atrophy provides a sensitive, global measure of the neurodegenerative process of multiple sclerosis, and should be studied in trials of new disease modifying treatments, whether their proposed mechanism of action is anti-inflammatory, remyelinating or neuroprotective.

Recommendations for future research and clinical applications

In considering current applications and results of atrophy studies in multiple sclerosis, several areas for future research have emerged. These relate to both methodological and clinical issues.

Methodological issues

There is a need for more comparative studies of different sequence acquisition and analysis algorithms to determine their relative precision, sensitivity to change and stability over time. Collaborative exchange of image databases and analysis tools would be a first step towards such activities and has already begun. Further work is needed to evaluate algorithms for regional volume measures including infratentorial, CSF, white matter and grey matter volumes. Greater sensitivity is needed to detect small changes especially in small structures such as the optic nerves and spinal cord. Further automation of analysis techniques designed to look at atrophy in brain structures such as the basal ganglia, corpus callosum and major white matter tracts would allow more specific studies to be carried out.

Clinical issues

Clinical questions concern mechanisms, prognosis and treatment monitoring.

Mechanisms

The mechanisms of atrophy are still poorly understood and their elucidation is an important focus for future work. The relationship between lesions and atrophy is complicated by apparent temporal dissociation between lesion appearance and later atrophy. Acute inflammatory and post-inflammatory mechanisms in the lesion may contribute to axonal loss. A potential post-inflammatory mechanism is failure of remyelination with loss of trophic support from oligodendrocytes. Reliable methods for imaging remyelination are still awaited. The evidence of an only moderate relationship between lesions and atrophy suggests that other mechanisms should be explored, e.g. diffuse pathological changes in the NAWM, grey matter pathology. The relationship between cord and brain changes also needs consideration, bearing in mind that disease in one part may, by Wallerian degeneration, cause atrophy in the other. Finally, the importance of cortical adaptive mechanisms in maintaining function in the face of structural pathology including atrophy is unknown but can be explored using functional MRI (Lee *et al.*, 2000; Werring *et al.*, 2000).

Addressing pathogenesis issues should use quantitative techniques such as magnetization transfer imaging, diffusion tensor imaging and MR spectroscopy to study lesions and normal appearing grey and white matter, and functional MRI

to evaluate cortical adaptation. Long-term follow up from the early stages is needed.

Prognosis

Definition of the importance of atrophy as a prognostic marker will require prolonged clinical follow up studies. Preliminary experience suggests a significant role (Fisher *et al.*, 2000a, b; Dalton *et al.*, 2001).

Treatment monitoring

Given the role of atrophy in trials, sample sizes should be calculated for demonstrating a desired treatment effect; this should be achievable using longitudinal data already available.

Atrophy should now be included as an endpoint in trials of disease-modifying agents aimed at preventing disability, whether the proposed mechanism of action is anti-inflammatory, remyelinating or neuroprotective. Because the correlations with disability and atrophy are only moderate, the former should remain the primary outcome measure. However, as a secondary outcome, atrophy is more important than lesion measures when the aim is to prevent disability. Especially in progressive forms of multiple sclerosis, reliance only on enhancing or T₂ lesion load can overestimate treatment benefits. Prolonged follow-up studies are needed to determine the long-term relationship between atrophy and clinical evolution. Such studies will further define the role of atrophy measures in therapeutic monitoring both in trials and individual subjects.

Acknowledgements

This paper is partly based on the proceedings of an International Workshop on the Measurement of Atrophy in Multiple Sclerosis, which was held at The Institute of Neurology, London, UK, on December 7 and 8, 2000. The Workshop received sponsorship support from The National Multiple Sclerosis Society (USA), The Multiple Sclerosis Society of Great Britain and Northern Ireland, The Guarantors of *Brain*, Biogen, Elan Pharmaceutical, Schering AG and Serono. The Workshop participants were: D. Arnold (Montreal), F. Barkhof (Amsterdam), G. Barker (London), P. Brex (London), D. Chard (London), L. Collins (Montreal), C. Dalton (London), N. De Stefano (Siena), M. Esiri (Oxford), M. Filippi (Milan), E. Fisher (Cleveland), N. Fox (London), J. Frank (Washington), P. Freitag (Basel), C. Gasperini (Rome), C. Good (London), C. Guttmann (Boston), S. Hickman (London), C. Jones (Vancouver), N. Kalkers (Amsterdam), L. Layward (London), S. Leary (London), P. Matthews (Oxford), D. Miller (London), P. Molyneux (London), X. Montalban (Barcelona), A. Paolillo (Rome), G. Parker (London), D. Paty (Vancouver), S. Reingold (New York), M. Rovaris (Milan), A. Rovira (Barcelona), R. Rudick (Cleveland), J. Simon

(Denver), S. Smith (Oxford), V. Stevenson (London), L. Stone (Cleveland), A. Thompson (London) and J. Wolinsky (Houston). All participants saw the manuscript prior to submission, and we are grateful for their contributions of data that have been included in this review.

References

- Adams CWM. A colour atlas of multiple sclerosis and other myelin disorders. London: Wolfe; 1989.
- Ashburner J, Friston KJ. Voxel-based morphometry – the methods. [Review]. *Neuroimage* 2000; 11: 805–21.
- Bedell BJ, Narayana PA. Volumetric analysis of white matter, gray matter, and CSF using fractional volume analysis. *Magn Reson Med* 1998; 39: 961–9.
- Bedell BJ, Narayana PA, Wolinsky JS. A dual approach for minimizing false lesion classifications on magnetic resonance images. *Magn Reson Med* 1997; 37: 94–100.
- Bø L, Nyland H, Vedeler C, Mørk S, Trapp BD. Extensive cortical demyelination in multiple sclerosis brains [abstract]. *Ann Neurol* 2000; 48: 446.
- Brex PA, Jenkins R, Fox NC, Crum WR, O’Riordan JI, Plant GT, et al. Detection of ventricular enlargement in patients at the earliest clinical stage of MS. *Neurology* 2000; 54: 1689–91.
- Brex PA, Leary SM, O’Riordan JI, Miszkiel KA, Plant GT, Thompson AJ, et al. Measurement of spinal cord area in clinically isolated syndromes suggestive of multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2001; 70: 544–7.
- Broman T. Blood–brain-barrier damage in multiple sclerosis. Supravital test observations. *Acta Neurol Scand* 1964; 40 Suppl 10: 21–4.
- Camp SJ, Stevenson VL, Thompson AJ, Miller DH, Borrás C, Auriacombe S, et al. Cognitive function in primary progressive and transitional progressive multiple sclerosis. A controlled study with MRI correlates. *Brain* 1999; 122: 1341–8.
- Charcot, J-M. Histologie de la sclerose en plaques. *Gaz Hop (Paris)* 1868; 41: 554–55, 557–8, 566.
- Chard DT, Griffin CMB, Parker GJM, Kapoor R, Thompson AJ, Miller DH. Brain atrophy in clinically early relapsing remitting multiple sclerosis. *Brain* 2002; 125: 327–37.
- Coffey CE, Wilkinson WE, Parashos IA, Soady SA, Sullivan RJ, Patterson LJ, et al. Quantitative cerebral anatomy of the aging human brain: a cross-sectional study using magnetic resonance imaging. *Neurology* 1992; 42: 527–36.
- Coles AJ, Wing MG, Molyneux P, Paolillo A, Davie CM, Hale G, et al. Monoclonal antibody treatment exposes three mechanisms underlying the clinical course of multiple sclerosis. *Ann Neurol* 1999; 46: 296–304.
- Collins DL, Narayanan S, Caramanos Z, De Stefano N, Tartaglia MC, Arnold DL. Relation of cerebral atrophy in multiple sclerosis to severity of disease and axonal injury [abstract]. *Neurology* 2000; 54 Suppl 3: A17.
- Collins DL, Montagnat J, Zijdenbos AP, Evans AC, Arnold DL.

- Automated estimation of brain volume in multiple sclerosis with BICCR. *Lect Notes Comput Sci* 2001; 2082: 141–7.
- Comi G, Filippi M, Barkhof F, Durelli L, Edan G, Fernandez O, et al. Effect of early interferon treatment on conversion to definite multiple sclerosis: a randomised study. *Lancet* 2001; 357: 1576–82.
- Coulon O, Parker GJM, Schnabel JA, Arridge SR. A B-spline active surface method for segmentation of the optic nerve [abstract]. *Proc Int Soc Magn Reson Med* 2000; 8: 1760.
- Coulon O, Parker GJM, Hickman SJ, Arridge SR. Quantification of spinal cord atrophy via an active surface model [abstract]. *Proc Int Soc Magn Reson Med* 2001; 9: 819.
- Courchesne E, Chisum HJ, Townsend J, Cowles A, Covington J, Egaas B, et al. Normal brain development and aging: quantitative analysis at in vivo MR imaging in healthy volunteers. *Radiology* 2000; 216: 672–82.
- Cowell PE, Turetsky BI, Gur RC, Grossman RI, Shtasel DL, Gur RE. Sex differences in aging of the human frontal and temporal lobes. *J Neurosci* 1994; 14: 4748–55.
- Dalton CM, Brex PA, Gordon R, Fox NC, Miszkief KA, Crum WR, et al. Progressive ventricular enlargement in patients with clinically isolated syndromes is associated with the development of multiple sclerosis [abstract]. *J Neurol* 2001; 248 Suppl 2: 26.
- DeCarli C, Maisog J, Murphy DG, Teichberg D, Rapoport SI, Horwitz B. Method for quantification of brain, ventricular, and subarachnoid CSF volumes from MR images. *J Comput Assist Tomogr* 1992; 16: 274–84.
- Edwards SGM, Gong QY, Liu C, Zvartau ME, Jaspan T, Roberts N, et al. Infratentorial atrophy on magnetic resonance imaging and disability in multiple sclerosis. *Brain* 1999; 122: 291–301.
- Esiri MM. Immunoglobulin-containing cells in multiple sclerosis plaques. *Lancet* 1977; 2: 478.
- Evangelou N, Esiri MM, Smith S, Palace J, Matthews PM. Quantitative pathological evidence for axonal loss in normal appearing white matter in multiple sclerosis. *Ann Neurol* 2000a; 47: 391–5.
- Evangelou N, Konz D, Esiri MM, Smith S, Palace J, Matthews PM. Regional axonal loss in the corpus callosum correlates with cerebral white matter lesion volume and distribution in multiple sclerosis. *Brain* 2000b; 123: 1845–9.
- Ferguson B, Matyszak MK, Esiri MM, Perry VH. Axonal damage in acute multiple sclerosis lesions. *Brain* 1997; 120: 393–9.
- Filippi M, Campi A, Colombo B, Pereira C, Martinelli V, Baratti C, et al. A spinal cord MRI study of benign and secondary progressive multiple sclerosis. *J Neurol* 1996; 243: 502–5.
- Filippi M, Colombo B, Rovaris M, Pereira C, Martinelli V, Comi G. A longitudinal magnetic resonance imaging study of the cervical cord in multiple sclerosis. *J Neuroimaging* 1997; 7: 78–80.
- Filippi M, Mastrorlando G, Rocca MA, Pereira C, Comi G. Quantitative volumetric analysis of brain magnetic resonance imaging from patients with multiple sclerosis. *J Neurol Sci* 1998; 158: 148–53.
- Filippi M, Rovaris M, Iannucci G, Mennea S, Sormani MP, Comi G. Whole brain volume changes in patients with progressive MS treated with cladribine. *Neurology* 2000; 55: 1714–8.
- Filippi M, Rovaris M, Barkhof F, Durelli L, Edan G, Fernandez O, et al. Brain volume changes in patients at presentation with clinically isolated syndromes suggestive of multiple sclerosis: results from the ETOMS study [abstract]. *J Neurol* 2001; 248 Suppl 2: 113.
- Fisher E, Cothren RM, Tkach JA, Masaryk TJ, Cornhill JF. Knowledge-based 3D segmentation of the brain in MR images for quantitative multiple sclerosis lesion tracking. *SPIE Med Imag* 1997; 3034: 19–25.
- Fisher E, Rudick RA, Cutter G, Baier M, Miller DM, Weinstock-Guttman B, et al. Usefulness of brain parenchymal fraction for predicting multiple sclerosis disease progression [abstract]. *Neurology* 2000a; 54 Suppl 3: A86.
- Fisher E, Rudick RA, Cutter G, Baier M, Miller D, Weinstock-Guttman B, et al. Relationship between brain atrophy and disability: an 8-year follow-up study of multiple sclerosis patients. *Mult Scler* 2000b; 6: 373–7.
- Fox NC, Jenkins R, Leary SM, Stevenson VL, Losseff NA, Crum WR, et al. Progressive cerebral atrophy in MS. A serial study using registered, volumetric MRI. *Neurology* 2000; 54: 807–12.
- Freeborough PA, Fox NC. The boundary shift integral: an accurate and robust measure of cerebral volume changes from registered repeat MRI. *IEEE Trans Med Imaging* 1997; 16: 623–9.
- Freeborough PA, Fox NC. Modeling brain deformations in Alzheimer disease by fluid registration of serial 3D MR images. *J Comput Assist Tomogr* 1998; 22: 838–43.
- Freeborough PA, Woods RP, Fox NC. Accurate registration of serial 3D MR brain images and its application to visualizing change in neurodegenerative disorders. *J Comput Assist Tomogr* 1996; 20: 1012–22.
- Ganter P, Prince C, Esiri MM. Spinal cord axonal loss in multiple sclerosis: a post-mortem study. *Neuropathol Appl Neurobiol* 1999; 25: 459–67.
- Gasperini C, Paolillo A, Giugni E, Galgani S, Bagnato F, Mainero C, et al. MRI brain volume changes in relapsing remitting multiple sclerosis patients treated with interferon beta-1a. *Mult Scler*. In press 2002.
- Ge Y, Grossman RI, Udupa JK, Wei L, Mannon LJ, Polansky M, et al. Brain atrophy in relapsing–remitting multiple sclerosis and secondary progressive multiple sclerosis: longitudinal quantitative analysis. *Radiology* 2000a; 214: 665–70.
- Ge Y, Grossman RI, Udupa JK, Fulton J, Constantinescu CS, Gonzales-Scarano F, et al. Glatiramer acetate (Copaxone) treatment in relapsing–remitting MS. Quantitative MR assessment. *Neurology* 2000b; 54: 813–7.
- Goldszal AF, Davatzikos C, Pham DL, Yan MX, Bryan RN, Resnick SM. An image-processing system for qualitative and quantitative volumetric analysis of brain images. *J Comput Assist Tomogr* 1998; 22: 827–37.
- Gur RC, Mozley PD, Resnick SM, Gottlieb GL, Kohn M, Zimmerman R, et al. Gender differences in age effect on brain

- atrophy measured by magnetic resonance imaging. *Proc Natl Acad Sci USA* 1991; 88: 2845–9.
- Guttmann CR, Kikinis R, Anderson MC, Jakab M, Warfield SK, Killiany RJ, et al. Quantitative follow-up of patients with multiple sclerosis using MRI: reproducibility. *J Magn Reson Imaging* 1999; 9: 509–18.
- Guttmann CRG, Warfield SK, Guimond A, Kikinis R, Albert MS, Jolesz FA, et al. Variability of brain atrophy estimates in multiple sclerosis [abstract]. *Proc Int Soc Magn Reson Med* 2000a; 8: 1188.
- Guttmann CR, Benson R, Warfield SK, Wei X, Anderson MC, Hall CB, et al. White matter abnormalities in mobility-impaired older persons. *Neurology* 2000b; 54: 1277–83.
- Hickman SJ, Behan CHM, Brex PA, Scolding NJ, Compston DAS, Miller DH. Continuing optic nerve atrophy following optic neuritis: a serial MRI study using a short echo fast-flair sequence [abstract]. *Rev Neurol* 2000; 156 Suppl 3: 3S34.
- Hickman SJ, Brex PA, Brierley CM, Silver NC, Barker GJ, Scolding NJ, et al. Detection of optic nerve atrophy following a single episode of unilateral optic neuritis by MRI using a fat-saturated short-echo fast FLAIR sequence. *Neuroradiology* 2001; 43: 123–8.
- Inglese M, Rovaris M, Giacomotti L, Mastronardo G, Comi G, Filippi M. Quantitative brain volumetric analysis from patients with multiple sclerosis: a follow-up study. *J Neurol Sci* 1999; 171: 8–10.
- Jones CK, Riddehough A, Li DKB, Zhao G, Paty DW. MRI cerebral atrophy in relapsing–remitting MS: results from the PRISMS trial [abstract]. *Neurology* 2001; 56 Suppl 3: A379.
- Kalkers NF, Bergers L, Castelijns JA, van Walderveen MA, Bot JC, Ader HJ, et al. Optimizing the association between disability and biological markers in MS. *Neurology* 2001a; 57: 1253–8.
- Kalkers NF, Barkhof F, Bergers L, Schijndel R van, Polman C. Neuroprotective treatment in primary progressive MS: a phase I/II study with riluzole. In: Comi G, Hommes O, editors. *Topics in neuroscience*. In press 2001b.
- Kidd D, Thorpe JW, Thompson AJ, Kendall BE, Moseley IF, MacManus DG, et al. Spinal cord MRI using multi-array coils and fast spin echo. II. Findings in multiple sclerosis. *Neurology* 1993; 43: 2632–7.
- Kidd D, Barkhof F, McConnell R, Algra PR, Allen IV, Revesz T. Cortical lesions in multiple sclerosis. *Brain* 1999; 122: 17–26.
- Kikinis R, Guttmann CR, Metcalf D, Wells WM 3rd, Ettinger GJ, Weiner HL, et al. Quantitative follow-up of patients with multiple sclerosis using MRI: technical aspects. *J Magn Reson Imaging* 1999; 9: 519–30.
- Lassman H, Suchanek G, Ozawa K. Histopathology and blood-cerebrospinal fluid barrier in multiple sclerosis. *Annal Neurol* 1994; 36: 542–6.
- Leary SM, Parker GJ, Stevenson VL, Barker GJ, Miller DH, Thompson AJ. Reproducibility of magnetic resonance imaging measurements of spinal cord atrophy: the role of quality assurance. *Magn Reson Imaging* 1999; 17: 773–6.
- Lee M, Reddy H, Johansen-Berg H, Pendlebury S, Jenkinson M, Smith S, et al. The motor cortex shows adaptive functional changes to brain injury from multiple sclerosis. *Ann Neurol* 2000; 47: 606–13.
- Leigh R, Ostuni JL, Pham D, Lewis BK, Howard T, Richert N, et al. Estimating cerebral atrophy in ms patients from various MR pulse sequences. *Mult Scler*. In press 2002.
- Lemieux L, Hagemann G, Krakow K, Woermann FG. Fast, accurate, and reproducible automatic segmentation of the brain in T₁-weighted volume MRI data. *Magn Reson Med* 1999; 42: 127–35.
- Liu C, Edwards S, Gong Q, Roberts N, Blumhardt LD. Three dimensional MRI estimates of brain and spinal cord atrophy in multiple sclerosis. *J Neurol Neurosurg Psychiatry* 1999; 66: 323–30.
- Losseff NA, Webb SL, O’Riordan JI, Page R, Wang L, Barker GJ, et al. Spinal cord atrophy and disability in multiple sclerosis. A new reproducible and sensitive MRI method with potential to monitor disease progression. *Brain* 1996a; 119: 701–8.
- Losseff NA, Wang L, Lai HM, Yoo DS, Gawne-Cain ML, McDonald WI, et al. Progressive cerebral atrophy in multiple sclerosis. A serial MRI study. *Brain* 1996b; 119: 2009–19.
- Lovas G, Szilagyi N, Majtenyi K, Palkovits M, Komoly S. Axonal changes in chronic demyelinated cervical spinal cord plaques. *Brain* 2000; 123: 308–17.
- Lycklamaa Nijeholt GJ, van Walderveen MA, Castelijns JA, van Waesberghe JH, Polman C, Scheltens P, et al. Brain and spinal cord abnormalities in multiple sclerosis. Correlation between MRI parameters, clinical subtypes and symptoms. *Brain* 1998; 121: 687–97.
- Mathalon DH, Sullivan EV, Rawles JM, Pfefferbaum A. Correction for head size in brain-imaging measurements. *Psychiatry Res* 1993; 50: 121–39.
- Molyneux PD, Kappos L, Polman C, Pozzilli C, Barkhof F, Filippi M, et al. The effect of interferon beta-1 b treatment on MRI measures of cerebral atrophy in secondary progressive multiple sclerosis. *Brain* 2000; 123: 2256–63.
- O’Riordan JI, Thompson AJ, Kingsley DP, MacManus DG, Kendall BE, Rudge P, et al. The prognostic value of brain MRI in clinically isolated syndromes of the CNS. A 10-year follow-up. *Brain* 1998; 121: 495–503.
- Paolillo A, Coles AJ, Molyneux PD, Gawne-Cain ML, MacManus D, Barker GJ, et al. Quantitative MRI in patients with secondary progressive MS treated with monoclonal antibody Campath 1H. *Neurology* 1999; 53: 751–7.
- Paolillo A, Pozzilli C, Gasperini C, Giugni E, Mainero E, Giuliani S, et al. Brain atrophy in relapsing–remitting multiple sclerosis: relationship with ‘black holes’, disease duration and clinical disability. *J Neurol Sci* 2000; 174: 85–91.
- Pfefferbaum A, Mathalon DH, Sullivan EV, Rawles JM, Zipursky RB, Lim KO. A quantitative magnetic resonance imaging study of changes in brain morphology from infancy to late adulthood. *Arch Neurol* 1994; 51: 874–87.
- Pham DL, Prince JL. Adaptive fuzzy segmentation of magnetic resonance images. *IEEE Trans Med Imaging* 1999a; 18: 737–52.
- Pham DL, Prince JL. An adaptive fuzzy C-means algorithm for

- image segmentation in the presence of intensity inhomogeneities. *Patt Recog Lett* 1999b; 20: 57–68.
- Plummer DL. DispImage: a display and analysis tool for medical images. *Revist Neuroradiol* 1992; 5: 489–95.
- Rao AB, Richert ND, Howard T, Lewis B, Bash CN, McFarland HF, et al. Effect of intravenous methylprednisolone on MRI measures of disease activity in RRMS patients before and after interferon beta-1b. *Neurology*. In press 2002.
- Rice GP, Filippi M, Comi G. Cladribine and progressive MS. Clinical and MRI outcomes of a multicenter controlled trial. Cladribine MRI Study Group. *Neurology* 2000; 54: 1145–54.
- Rovaris M, Inglese M, van Schijndel RA, Sormani MP, Rodegher M, Comi G, et al. Sensitivity and reproducibility of volume change measurements of different brain portions on magnetic resonance imaging in patients with multiple sclerosis. *J Neurol* 2000; 247: 960–5.
- Rovaris M, Comi G, Rocca MA, Wolinsky JS, Filippi M, and the European/Canadian Glatiramer Acetate Study Group. Short-term brain volume change in relapsing–remitting multiple sclerosis: effect of glatiramer acetate and implications. *Brain* 2001; 124: 1803–12.
- Rudick RA, Fischer E, Lee J-C, Simon J, Jacobs L. Use of the brain parenchymal fraction to measure whole brain atrophy in relapsing–remitting MS. *Neurology* 1999; 53: 1698–704.
- Sailer M, Losseff NA, Wang L, Gawne-Cain ML, Thompson AJ, Miller DH. T₁ lesion load and cerebral atrophy as a marker for clinical progression in patients with multiple sclerosis. A prospective 18 months follow-up study. *Eur J Neurol* 2001; 8: 37–42.
- Sandaine AM, Ge Y, Udupa JK, Babb JS, Mannon LJ, Grossman RI. The effect of gadolinium-enhancing lesions on whole brain atrophy in relapsing–remitting MS. *Neurology* 2000; 55: 61–5.
- Silver NC, Tofts PS, Symms MR, Barker GJ, Thompson AJ, Miller DH. Quantitative contrast-enhanced magnetic resonance imaging to evaluate blood–brain barrier integrity in multiple sclerosis: a preliminary study. *Mult Scler* 2001; 7: 75–82.
- Simon JH, Jacobs LD, Campion MK, Rudick RA, Cookfair DL, Herndon RM, et al. A longitudinal study of brain atrophy in relapsing multiple sclerosis. *Neurology* 1999; 53: 139–48.
- Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging* 1998; 17: 87–97.
- Smith SM, De Stefano N, Jenkinson M, Matthews PM. Normalized accurate measurement of longitudinal brain change. *J Comput Assist Tomogr* 2001; 25: 466–75.
- Stevenson VL, Leary SM, Losseff NA, Parker GJ, Barker GJ, Husmani Y, et al. Spinal cord atrophy and disability in MS. A longitudinal study. *Neurology* 1998; 51: 234–8.
- Stevenson VL, Miller DH, Leary SM, Rovaris M, Barkhof F, Brochet B, et al. One year follow up study of primary and transitional progressive multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2000; 68: 713–8.
- Stoll G, Trapp BD, Griffin JW. Macrophage function during Wallerian degeneration of rat optic nerve: clearance of degenerating myelin and Ia expression. *J Neurosci* 1989; 9: 2327–35.
- Tourtellotte WW, Parker JA. Some spaces and barriers in postmortem multiple sclerosis. *Prog Brain Res* 1968; 29: 493–525.
- Trapp BD, Peterson J, Ransohoff RM, Rudick R, Mork S, Bo L. Axonal transection in the lesions of multiple sclerosis. *New Engl J Med* 1998; 338: 278–85.
- Tubridy N, Ader HJ, Barkhof F, Thompson AJ, Miller DH. Exploratory treatment trials in multiple sclerosis using MRI: sample size calculations for relapsing–remitting and secondary progressive subgroups using placebo controlled parallel groups. *J Neurol Neurosurg Psychiatry* 1998; 64: 50–5.
- Udupa JK, Wei L, Samarasekera S, Miki Y, van Buchem MA, Grossman RI. Multiple sclerosis lesion quantification using fuzzy-connectedness principles. *IEEE Trans Med Imaging* 1997; 16: 598–609.
- vanWalderveen MA, Kamphorst W, Scheltens P, van Waesberghe JH, Ravid R, Valk J, et al. Histopathologic correlate of hypointense lesions on T₁-weighted spin-echo MRI in multiple sclerosis. *Neurology* 1998; 50: 1282–8.
- Walters RJ, Fox NC, Crum WR, Taube D, Thomas DJ. Haemodialysis and cerebral oedema. *Nephron* 2001; 87: 143–7.
- Werring DJ, Bullmore ET, Toosy AT, Miller DH, Barker GJ, MacManus DG, et al. Recovery from optic neuritis is associated with a change in the distribution of cerebral response to visual stimulation: a functional magnetic resonance imaging study. *J Neurol Neurosurg Psychiatry* 2000; 68: 441–9.
- Whitwell JL, Crum WR, Watt HC, Fox NC. Normalization of cerebral volumes by use of intracranial volume: implications for longitudinal quantitative MR imaging. *AJNR Am J Neuroradiol* 2001; 22: 1483–9.
- Wolinsky JS, Narayana PA, Noseworthy JH, Lublin FD, Whitaker JN, Linde A, et al. Linomide in relapsing and secondary progressive MS: part II: MRI results. *Neurology* 2000; 54: 1734–41.
- Wolinsky JS, Narayana PA, Johnson KP, Multiple Sclerosis Study Group and the MRI-Analysis Center. United States open-label glatiramer acetate extension trial for relapsing multiple sclerosis: MRI and clinical correlates. *Mult Scler* 2001; 7: 33–41.
- Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging* 2001; 20: 45–57.
- Zijdenbos AP, Dawant BM. Brain segmentation and white matter lesion detection in MR images. [Review]. *Crit Rev Biomed Eng* 1994; 22: 401–65.
- Zijdenbos A, Evans A, Riahi F, Sled J, Chui J, Kollokian V. Automatic quantification of multiple sclerosis lesion volume using stereotaxic space. *Lect Notes Comput Sci* 1996; 1131: 439–48.

Received November 1, 2001.

Accepted December 17, 2001

Appendix I. Brain volume measurement methods

Linear and regional measures

Linear measures of the width of the brain, corpus callosum and third ventricle have been applied in multiple sclerosis for over a decade. Although simple and rapid, these approaches are relatively crude and are influenced by slice thickness and position. Nevertheless, significant atrophic changes have been found over 1 year. Larger changes are seen in third ventricle [+4.5% per annum (p.a.)] and corpus callosum (-5% p.a.) width, with smaller changes in brain width (-0.64% p.a.). While third ventricle enlargement and corpus callosum atrophy may partly reflect effects of adjacent lesions, Wallerian degeneration secondary to more distant lesions may be a more important influence; the third ventricle has both sensory and motor tracts travelling adjacent to it.

In a large cohort of relapsing–remitting patients engaged in a placebo-controlled trial of β -interferon, those with gadolinium-enhancing lesions at baseline exhibited more atrophy over the next 2 years (Simon *et al.*, 1999), but there was no significant effect of treatment on the linear measures of atrophy.

Third ventricular size has been measured using a semi-automated technique that determines the average of multiple linear measurement placements, based on signal intensity and statistical fit (J. Simon, personal communication). Signal peaks from the thalamus on either side are separated automatically from lower signal third ventricle on T₁-weighted images. An advantage of analysing this region, being at the iso-centre of the magnet, is homogeneity of the magnetic and radiofrequency fields. The central third ventricle area (c3VA) has been measured with a coefficient of variance (COV) of 0.86%. In a 2-year follow up of relapsing–remitting patients in a placebo-controlled trial of β -interferon, the amount of c3VA increase correlated with the number of enhancing lesions; there was no significant difference in the rate of change in the active versus placebo treatment arms, although there was a trend towards a treatment effect in year 2. At baseline there was a moderate correlation between c3VA and EDSS ($r = 0.33$) comparable to that between BPF and EDSS ($r = -0.29$). Correlations were also found between c3VA and both T₂ and T₁ lesion load.

Segmentation algorithms can be applied slice by slice to 2D images to provide a reproducible measure of brain volume contained within each slice, thus providing a degree of localization of the atrophy measurement. One such approach classifies brain from non-brain using the X-skull algorithm (Losseff *et al.*, 1996b; Lemieux *et al.*, 1999). Taking four 5-mm thick slices rostral to the velum interpositum, Losseff *et al.* (1996b) found that regional cerebral volume was reproducible to 0.5% with scan-reposition-rescan. Progressive atrophy was detectable over 18 months in a cohort with relapsing–remitting or secondary progressive multiple sclerosis, and was correlated with increasing disability but not with the number of gadolinium-enhancing lesions. The methodology has been successfully applied in clinical trials evaluating β -interferon in secondary progressive multiple sclerosis (Molyneux *et al.*, 2000) and in following large cohorts with primary progressive multiple sclerosis (Stevenson *et al.*, 2000).

Threshold-based measures

Image intensity threshold-based segmentation algorithms have also been applied slice by slice to 2D images (Filippi *et al.*, 1998; Inglese *et al.*, 1999; Rovaris *et al.*, 2000). A thresholding approach

has also been applied to T₂-weighted images to obtain the volume of the intracranial cavity and thus allow normalized measures of parenchymal and ventricular fractions (Kalkers *et al.*, 2001).

One study applied a thresholding approach to 3-mm thick slices in 50 relapsing–remitting patients studied over 18 months (Rovaris *et al.*, 2001). The volumes were calculated and compared for whole brain, the seven slices rostral to the velum interpositum, the single central slice of the brain slab and the infratentorial region. Measurement reproducibility was also evaluated for the four regions. The seven-slice region showed volume decreases in more patients (41 out of 50) than the other regions, provided the most reproducible measure (COV = 1.5%), and provided a volume that correlated strongly with whole brain volume ($r = 0.80$). The time saving of the seven-slice region makes this an attractive approach. The greater sensitivity to change in this region most likely reflects the high frequency of pathological change in lesions in this central brain region.

Segmentation-based methods

CSF volume

Brain atrophy is associated with a compensatory increase in the size of the compartment containing CSF. This can be measured directly by a number of approaches. As with the brain, either the total CSF space or a part thereof can be measured.

Normalized total CSF volume measures have been obtained using multiple contrast mechanisms from a single acquisition (so-called AFFIRMATIVE sequence) in order to improve tissue and CSF segmentation. A dual echo [echo time (TE) 17/102 ms] fast spin echo (FSE) sequence is used with 3-mm thick contiguous axial slices, 1 × 1 mm in plane resolution, and a very long repetition time (TR) (10 s) (Bedell *et al.*, 1997). All scans are acquired on GE 1.5T scanners operating at software version 5.4 or higher. A large multiple sclerosis/MRI database (almost 2000 multiple sclerosis patients, and >4000 scans by May 2000) has been collected from >80 centres (Wolinsky *et al.*, 2000, 2001). Based on tissue signal characteristics and on cluster analysis in a N-dimensional feature space, the scans are segmented into grey matter, white matter, lesions and CSF. A normalized CSF volume (nCSF) is derived from fully processed images in which extra-cerebral tissue has been removed using a largely automated integrated system. The reproducibility of nCSF, both measure-remeasure and scan-rescan proved excellent.

A comparison of subgroups matched for disability has revealed a larger nCSF in relapsing–remitting and secondary progressive than primary progressive patients. A study of monthly scans in 30 relapsing–remitting multiple sclerosis patients found an increase of nCSF of $0.1 \pm 0.07\%$ per month, corresponding to ~1% p.a. In a 6 year follow up of a cohort of patients on long-term treatment with glatiramer acetate, an increase in nCSF was correlated with a higher EDSS (Wolinsky *et al.*, 2001), but not with enhancing lesion volume, supporting the concept that measurement of atrophy (in this case inferred from change in nCSF) provides a valuable measure of disease progression that is different from that obtained using measures of lesion enhancement.

Improved quadruple contrast has recently been developed, making use of additional double inversion pulses (Bedell and Narayana, 1998). Preliminary experience with this sequence suggests improved reliability in lesion detection without compromising the measurement of CSF or tissue volumes.

BPF

One fully automated method calculates the BPF as the ratio of the volume of parenchymal brain tissue to the total volume within the outer surface of the brain [note: other groups have applied the term BPF for alternative methods of normalization; this section refers only to the method developed by the Cleveland group (Fisher *et al.*, 2000a)]. The algorithm consists of a sequence of steps. First, an initial segmentation of the brain is obtained through optimal thresholding, connected-components labelling, and erosion and dilation. Secondly, to remove non-brain structures that are still connected to the brain, a three-dimensional radial search operation is performed to detect the outer surface of the brain, which is then interpolated and smoothed in spherical space. Subsequently, the parenchymal tissue within the outer surface is detected by applying the optimal threshold determined in the first step. The brain parenchymal volume is calculated from the brain mask using a two-compartment mixture model to account for partial volume effects. This method excludes CSF within the sulci but unlike many other methods, not that external to the outer surface of the brain. An advantage is that there is only one segmentation performed to obtain both the numerator and denominator, so differences over time in the field of view or segmentation partially cancel out and improve the reproducibility. The accuracy of the method has been confirmed in digital and physical-known-volume phantoms and by comparing manual outlining with the automated technique. It has been successfully applied to 2D images, with an excellent COV for scan-rescan evaluation (0.2% for BPF, and 0.6% for both the outer contour volume and parenchymal volume). This technique can be applied to any type of MRI of the brain and no special acquisition or repositioning is required, but so far it has only been validated in dual echo spin echo and FLAIR images. A group of 50 adult controls (aged 19–53 years) exhibited a mean BPF of 0.868 (SD 0.008). Significant reduction in BPF has been found in relapsing–remitting multiple sclerosis, and a higher annual rate of atrophy is also seen in relapsing–remitting multiple sclerosis versus controls (mean 0.6 versus 0.1%). Potential drawbacks of the methods are the possibility that the outer volume will change over time (thus normalization is not fixed), insensitivity to increases in the CSF space external to the outer brain volume that occur with progressive atrophy, and very small changes over time due to the global nature of the measurement. Visual inspection of each segmented image is advised to check for unreliable measurements that may arise as a result of patient motion and sequence variations or image artefacts in serial studies.

WBR method

This approach uses standard proton density (PD)/T₂-weighted images that are routinely collected in most serial multiple sclerosis studies; thus it has potential to be applied retrospectively to large databases (C. Jones, personal communication). Anisotropic noise filtering is used, followed by the creation of an ‘angle image’, which is the inverse tangent of the ratio of the filtered long echo image divided by the filtered short echo volume. A histogram of the angle images for all slices is computed and Gaussian distributions are fitted to determine the position and width of the main peak, which represents the brain. From smart thresholding of this histogram, masks of both the intra-dural space (the width of the peak) and CSF space (values greater than the main peak) are determined. A further threshold is applied to extract the brain (~60% of the maximum

signal intensity value). While the method is normally effective in removing scalp and other extradural tissues, manual editing is sometimes required to remove extraneous areas from the intradural mask (e.g. eyes); it is also needed to re-classify lesions into brain (the software sometimes classifies them as CSF). In a serial study of 519 relapsing–remitting patients involved in a multicentre trial only 12 scans could not be analysed because of poor signal-to-noise ratio or an inadequate baseline scan. WBR was calculated as:

$$\text{WBR} = (\text{intradural volume} - \text{CSF volume}) / (\text{intradural volume})$$

The mean (\pm standard deviation) WBR was 0.83 (\pm 0.4), and significant decrease was seen after 1 year. After 2 years, the mean decrease in WBR was –1.4%. The measurement of WBR also demonstrated high reproducibility (COV < 1%).

BICCR

This fully automated method provides a normalized measure of brain volume (Collins *et al.*, 2001). It is applied to conventional dual echo PD- and T₂-weighted image data. The process involves correction of image intensity non-uniformities; stereotactic registration and resampling of image data (which removes extracranial tissues); intensity normalization to a standard average PD- or T₂-weighted volume; removing inferior and superior slices in Talairach coordinates to leave an anatomically equivalent 80-mm thick axial slab in all cases; digital morphometry for intra-cranial cavity classification; and Bayesian tissue classification into grey matter, white matter, lesions, CSF and background voxels within the brain mask. The BICCR is described by the equation:

$$\text{BICCR} = 100 \times (\text{GM} + \text{WM} + \text{lesions}) / (\text{GM} + \text{WM} + \text{lesions} + \text{CSF})$$

In a group of normal controls, mean (standard deviation) BICCR was 86.1 (2.8). Rescanning four controls several months apart yielded a COV of 0.21%. The COV is related to slice thickness, being poorer with 1-mm thick than with thicker slices; this may relate to increased artefacts from increases in image acquisition time. The method eliminates the need for precise repositioning and addresses problems associated with partial head coverage during scanning.

Fuzzy connectedness segmentation

The data acquired from multiple sequences with different signal contrast can be used to segment brain from other structures and white matter from grey matter. One such approach, based on PD and T₂-weighted scans, uses a fuzzy connectedness algorithm (Udupa *et al.*, 1997) and image non-uniformity correction and has exhibited a high sensitivity in detecting atrophy during serial studies both during natural history and treatment trials (Ge *et al.*, 2000a, b; Sandaine *et al.*, 2000).

SPM-based segmentation

SPM software uses prior spatial information from a database of normal brain images to classify voxels according to their location and signal intensity characteristics as grey matter, white matter or CSF (Ashburner and Friston, 2000). It can be applied to 2D or 3D images. Application of SPM99 to 3D T₁-weighted gradient echo images obtained from healthy adults aged 23–55 years revealed

highly reproducible measurements of intracranial volume and brain volume. Image non-uniformity correction improved reproducibility to better than 1% (Chard *et al.*, 2002). General linear model analysis revealed significant age related decreases in BPF with age, and significant gender differences with respect to grey matter fraction (higher in females) and white matter fraction (lower in females). SPM mis-classifies white matter lesions as CSF and grey matter; therefore, they need to be edited manually or by a local threshold method. In early relapsing–remitting multiple sclerosis, significant decreases in BPF have been found, which appears to derive from a reduction in white and grey matter fractions.

Template driven segmentation of the brain

This approach identifies brain tissues and structures by projecting anatomically labelled images of a template brain on to images of the patient's brain (Guttmann *et al.*, 2000b). Template driven segmentation (TDS) puts a standard space template on to an individual patient image, whereas some other algorithms (e.g. SPM) place the patient's scan into stereotactic space; it is likely that the overall result is similar. Global and regional 3D measures of brain parenchyma volumes can be obtained and used in serial studies to evaluate the amount of atrophy. The method has been validated for accuracy and reproducibility and has been applied to 3D T₂-weighted FSE and fast FLAIR sequences. Further definition of normal and abnormal regions (e.g. lesions, NAWM) is made possible by statistical analysis of the distribution of MRI signal on multiple contrast images. Its ability to discriminate normal from abnormal NAWM is less certain.

SIENAX

The longitudinal atrophy rate method SIENA (see section 'Brain surface modelling' below) has been extended to a segmentation-based cross-sectional method, SIENAX. After brain extraction, the input image is registered to a standard brain, using the estimated skull surface to constrain scaling, i.e. to normalize for head size. MRF (MRI random field model)-based segmentation (Zhang *et al.*, 2001) is used to segment the brain image into different tissue types, including partial volume modelling, giving a total normalized brain volume as output. SIENAX gives test-retest reproducibility errors in brain volume of <1%.

Registration-based methods

MIDAS (Medical Image Display and Analysis Software) method

This approach is based on rigid-body registration between pairs of T₁-weighted 3D scans using fast sinc interpolation (MIDAS method; Freeborough *et al.*, 1996). A linear scaling procedure is used to account for variations in voxel size due to scanner performance, and the skull is used as an invariant structure for the registration process. The accuracy of registration is to subvoxel level. A quantitation method, the boundary shift integral (BSI), determines changes in signal intensity along the boundaries of the brain. The volume change measured using BSI has proved highly reproducible (0.1–0.2% differences in absolute brain volumes on repeat measures). Progressive atrophy is readily detected in multiple sclerosis with a mean decrease in brain volume of 0.9% over 1 year.

Atrophy is seen at even shorter intervals in Alzheimer's disease, and in that disorder change in brain volume correlates strongly with progressive cognitive decline. The method has also been used to show that large short-term changes can occur as a result of variation in water content—in patients with chronic renal failure, a 3% decrease in brain volume is seen in a single day following haemodialysis (Walters *et al.*, 2001). Fluid registration is a related approach, now applying a fully non-linear registration, in which voxel compression and enlargement is allowed within the constraints of a viscous fluid model. It allows accurate contrast driven matching of serial images, and provides maps of voxel compression that localize change. The process updates the displacement vector associated with each voxel while constraining the overall deformation to be continuous, topology preserved and one-to-one. Voxel compression maps (VCM) can be derived from the deformation field by computing the determinant of the Jacobian matrix at each point to provide a point estimate of volumetric change. The maps are combined with the structural images in overlay and thereby associate patterns of change with anatomical regions. The VCMs can be integrated numerically to provide a quantitative estimate of expansion of CSF spaces (i.e. brain atrophy) within any given region.

Voxel-based morphometry

Another approach that uses registration methodology is voxel-based morphometry (VBM) (Ashburner and Friston, 2000). This is a fully automated whole-brain technique for characterizing regional volume and tissue 'concentration' differences in structural MRIs. It allows all areas of the brain to be compared in an unbiased way and thus offers an opportunity to assess anatomical differences throughout the brain. Regional differences between groups can be compared, and in serial studies, a measure of volumetric change over time can be obtained. It may be difficult with VBM to detect changes in areas with a high natural variance (e.g. peri-Sylvian region), and aspects of the algorithm such as registration to a template and smoothing will reduce sensitivity to detect longitudinal changes. It has not yet been systematically applied in multiple sclerosis.

Brain surface modelling

A fully automated method for measuring longitudinal percentage changes in brain volume has been developed (SIENA, Structural Image Evaluation using Normalization, of Atrophy) (Smith *et al.*, 2001). The external surface of the skull is used as an invariant constraint on serial images; this is usually clearly visible on T₁-weighted images. The brain is segmented from non-brain, using 3D triangulated mesh modelling to the brain surface. This procedure balances local and global constraints, and uses a local threshold and smoothness factor to reliably detect the brain surface. Having found the brain surface on one scan, the procedure then finds surface point positions to sub-voxel accuracy (between scans at two different time points) using correlation of (differentiated) normal vectors. This is then converted into percentage brain volume change (PBVC). The precision and accuracy of PBVC is ~0.2%; better precision is achieved with thicker slices, perhaps because sequence acquisition time is less, thereby reducing motion artefacts.

The longitudinal method has also been developed to look for regional atrophy variation using standard-space masks. Using this

approach, secondary progressive patients exhibited more periventricular atrophy than relapsing–remitting patients, whereas similar amounts of peripheral atrophy were apparent.

Comparison of different volume analysis techniques and pulse sequences

It is clear that many promising techniques have been developed with which to measure global and regional brain tissue volumes with a high reproducibility and with evident sensitivity to small changes over time. Very little work has focused on comparing different segmentation analysis techniques for determining brain volumes. Such work is important for the interpretation of published studies that have applied different techniques.

In a recent study (Leigh *et al.*, 2002), four different image intensity-based segmentation methods algorithms applied to three different acquisition sequences were used to measure change in brain volume (BV) over 2 years in 10 patients with relapsing–remitting multiple sclerosis and five healthy controls. The segmentation algorithms were: (i) histogram segmentation (Leigh *et al.*, 2002); (ii) adaptive fuzzy C-means algorithm (Pham and Prince, 1999a, b); (iii) adaptive Bayesian segmentation with a K-mean clustering (Goldszal *et al.*, 1998); and (iv) threshold-based segmentation using log-normal modelling of the distribution of pixels (De Carli *et al.*, 1992). The acquisition sequences were T₁-weighted, FLAIR and PD/T₂-weighted. No significant changes over time were seen in the controls. The intra-rater coefficient of variation for the segmentation techniques varied from 0.08 to 0.68%. In multiple sclerosis patients, significant atrophy was seen after 1 and 2 years compared with baseline; however, for various combinations of acquisition sequence and segmentation algorithm, the percentage change in brain volume varied markedly. The reduction ranged from 0.34 to 1.67% comparing month 12 with baseline and from 0.79 to 3.11% comparing month 24 with baseline.

A comparison of two methods for measuring atrophy in 39 primary progressive patients showed that the central slices approach (in this instance six 3-mm thick slices from the velum interpositum rostrally) exhibits a larger percentage volume change than an automated boundary detection method (SIENA) applied to whole brain volume ($-2.3 \pm 2.7\%$ versus $-0.56 \pm 0.57\%$); however, the statistical effect size (mean change/variance) was slightly higher with the SIENA approach (Stevenson *et al.*, 2000).

These studies emphasize the major impact that both the acquisition sequence and the segmentation procedure can have on the actual measures of atrophy obtained. Since precision is more probably important than accuracy—there is no *in vivo* gold standard to determine the latter—the problem may not be crucial as long as the methods applied in longitudinal and multicentre studies are stable and standardized for all scans at all sites. The need for standardization and quality control are, however, emphasized, as is the need to consider all methodological factors that might influence tissue volume measures and their change over time. The observation also suggests the need for caution in attempting to compare results from different studies in which either the acquisition or the segmentation methods were different.

Appendix II. Measurement of spinal cord and optic nerve atrophy

Spinal cord

Since spinal cord disease contributes to much of the locomotor disability in multiple sclerosis, there has been considerable interest in evaluating the cord using MR techniques. While spinal lesions are readily detected using T₂-weighted sequences, they have not been clearly correlated with disability. Spinal cord atrophy is likely to be a consequence of loss of myelin and axons, the likely pathological substrates of irreversible disability in multiple sclerosis.

Early studies of cord size used manual outlining to define the cross-sectional area of the cord at multiple levels in the cervical and thoracic region. While these demonstrated an association of cord atrophy with disability (Kidd *et al.*, 1993; Filippi *et al.*, 1996, 1997), the methods were poorly reproducible (COV ~5%) due to the poor depiction of the cord–CSF interface on 2D T₂-weighted sequences and the manual outlining method. These problems were addressed by the application of a 3D T₁-weighted gradient echo sequence with an inversion recovery pre-pulse to suppress CSF. This provided high-resolution images of the upper cervical cord, with excellent CSF–cord contrast allowing for semi-automated segmentation of the cord on high-resolution axial slices reconstructed from the volume acquired data (Losseff *et al.*, 1996a). The measure–remeasure and scan–rescan reproducibility is ~1%, and quality control with repeat scanning over 1 year or longer has confirmed the stability of the measurements providing there are no major scanner hardware upgrades (Leary *et al.*, 1999). Cross-sectional studies have demonstrated robust correlations between cord atrophy at C2 and EDSS across the different multiple sclerosis subgroups. A significant increase in cord atrophy over 1 year has been seen in both relapsing–remitting and progressive forms of multiple sclerosis, with mean reduction in cord area of 2–3% p.a. (Stevenson *et al.*, 1998, 2000). Progressive cord atrophy is of particular interest in primary progressive multiple sclerosis patients with progressive myelopathy, many of who exhibit little in the way of cerebral disease. No correlation has yet emerged between change in cord area over time and change in disability, but the duration of follow up has been short (1–2 years).

Application of the Cavalieri method with stereological point counting provides an alternative approach to defining the spinal cord cross-sectional area. It has been applied to 3D T₁-weighted gradient echo images of the upper cervical cord, and in a cross-sectional study has shown an association between atrophy from C1–3 and disability (Edwards *et al.*, 1999). There was more atrophy in secondary progressive than relapsing–remitting patients. The technique is less reproducible than the method developed by Losseff, which may limit its use in serial studies.

More global measurements of spinal cord size, e.g. the volume of the whole cervical cord and the atrophy of the cord as a function of position, have yet to be addressed. Such approaches will potentially be more sensitive to change. One method currently under study uses mathematical modelling to the surface of the cord using the assumption that the cord approximates a cylindrically shaped object (i.e. the nerve bundles leaving the spinal cord at each vertebral body are not included in the model) (Coulon *et al.*, 2001). This approach allows cross-sectional areas (and therefore volumes) to be determined as a function of position along the cord (Fig. 2).

Optic nerves

The optic nerve is of interest in multiple sclerosis for several reasons: (i) it is frequently involved clinically (optic neuritis); (ii) the effects of single inflammatory demyelinating lesions on function can be assessed with a high degree of clinical precision (by neuro-ophthalmological assessment) and by evaluating conduction through the lesion using the visual evoked potential. In this context, MRI provides an assessment of the structural integrity of the symptomatic optic nerve lesion. The lesion can be visualized using high resolution, fat suppressed, T₂-weighted sequences and leakage of the blood–nerve barrier is depicted in acute lesions on gadolinium-enhanced images. Measurement of optic nerve size would allow a direct evaluation of the ability of single inflammatory/demyelinating lesions to produce atrophy. However, measuring

the size of the nerve is made difficult by its small dimensions, motion artefact (the eyes may move during a sequence acquisition), and problems in defining the edge of the nerve from its surrounding CSF containing sheath and orbital fat. Partial success has been achieved using a fat saturated fast FLAIR sequence (which suppresses both fat and CSF) with contiguous coronal 3-mm thick slices through the orbit. The mean cross-sectional area of the five slices from the orbital apex forwards has been measured with a measure-remeasure COV of 4.8% and a scan-rescan COV of 6.5%. Further methods are needed to improve the reproducibility of optic nerve atrophy measures: these may involve the use of 3D acquisitions, or mathematical models to detect the surface of the nerve (Coulon *et al.*, 2000).