

Cerebral atrophy as outcome measure in short-term phase 2 clinical trials in multiple sclerosis

I. J. van den Elskamp · B. Boden · V. Dattola · D. L. Knol · M. Filippi · L. Kappos · F. Fazekas · K. Wagner · C. Pohl · R. Sandbrink · C. H. Polman · B. M. J. Uitdehaag · F. Barkhof

Received: 14 October 2009 / Accepted: 3 December 2009 / Published online: 5 January 2010
© Springer-Verlag 2009

Abstract

Introduction Cerebral atrophy is a compound measure of the neurodegenerative component of multiple sclerosis (MS) and a conceivable outcome measure for clinical trials monitoring the effect of neuroprotective agents. In this study, we evaluate the rate of cerebral atrophy in a 6-month period, investigate the predictive and explanatory value of other magnetic resonance imaging (MRI) measures in relation to cerebral atrophy, and determine sample sizes for future short-term clinical trials using cerebral atrophy as primary outcome measure.

Methods One hundred thirty-five relapsing–remitting multiple sclerosis patients underwent six monthly MRI scans

from which the percentage brain volume change (PBVC) and the number and volume of gadolinium (Gd)-enhancing lesions, T2 lesions, and persistent black holes (PBH) were determined. By means of multiple linear regression analysis, the relationship between focal MRI variables and PBVC was assessed. Sample size calculations were performed for all patients and subgroups selected for enhancement or a high T2 lesion load at baseline.

Results A significant atrophy occurred over 6 months (PBVC = -0.33%, SE = 0.061, $p < 0.0001$). The number of baseline T2 lesions ($p = 0.024$), the on-study Gd-enhancing lesion volume ($p = 0.044$), and the number of on-study

I. J. van den Elskamp (✉) · B. Boden · V. Dattola · F. Barkhof
Department of Radiology, MS Center Amsterdam,
VU University Medical Center,
De Boelelaan 1117,
1081 HV Amsterdam, The Netherlands
e-mail: i.vandenelskamp@vumc.nl

V. Dattola
Department of Neurosciences,
Psychiatric and Anaesthesiological Sciences,
University of Messina,
Messina, Italy

D. L. Knol · B. M. J. Uitdehaag
Department of Epidemiology and Biostatistics,
VU University Medical Center,
Amsterdam, The Netherlands

M. Filippi
Neuroimaging Research Unit,
Scientific Institute and University Ospedale San Raffaele,
Milan, Italy

L. Kappos
Department of Neurology, University Hospital,
University of Basel,
Basel, Switzerland

F. Fazekas
Department of Neurology, Medical University of Graz,
Graz, Austria

K. Wagner · C. Pohl · R. Sandbrink
Bayer-Schering Pharma,
Berlin, Germany

C. Pohl
Department of Neurology, University Hospital Bonn,
Bonn, Germany

R. Sandbrink
Department of Neurology, Heinrich-Heine-University Dusseldorf,
Dusseldorf, Germany

C. H. Polman · B. M. J. Uitdehaag
Department of Neurology, MS Center Amsterdam,
VU University Medical Center,
Amsterdam, The Netherlands

PBHs ($p=0.003$) were associated with an increased rate of atrophy. For a 50% decrease in rate of atrophy, the sample size calculations showed that approximately 283 patients per arm are required in an unselected sampled population and 185 patients per arm are required in a selected population.

Conclusion Within a 6-month period, significant atrophy can be detected and on-study associations of PBVC and PBHs emphasizes axonal loss to be a driving mechanism. Application as primary outcome measure in short-term clinical trials with feasible sample size requires a potent drug to obtain sufficient power.

Keywords Multiple sclerosis · Cerebral atrophy · SIENA · Sample size

Introduction

Brain tissue loss is a prominent feature in the pathology of multiple sclerosis (MS) and occurs at a significantly higher rate in patients with MS compared to the normal aging brain [1, 2]. In greater part, this phenomenon has been attributed to neuronal and axonal loss in both lesions [3] and normal-appearing brain tissue [4, 5] and, therefore, cerebral atrophy is recognized as a global marker of the neurodegenerative components of MS [6]. Magnetic resonance imaging (MRI) detects the rate of cerebral atrophy in vivo in a sensitive and reproducible manner which, together with the substantial correlation with later clinical disability [7], makes cerebral atrophy a conceivable outcome measure for clinical trials measuring the efficacy of neuroprotective agents.

Previous studies in untreated patients with relapsing–remitting multiple sclerosis (RRMS) revealed fairly stable annualized brain volume decreases of approximately 0.6% to 1.35%, determined over moderate (1 year) to long (3 years) periods of follow-up [6]. Fewer studies, however, assessed the rate of cerebral atrophy for shorter periods of follow-up, and did so with various measures of atrophy. In a 3-month study of 138 RRMS patients, a significant decrease in atrophy rate was detected using brain parenchymal fraction as atrophy measure [8], whereas in a similar study using 30 RRMS patients, no significant decrease was found [9]. In another study, patients in the placebo arm of a clinical trial with a follow-up duration of 9 months showed a significant decrease in brain volume measured by assessment of seven contiguous brain slices [10].

Although these studies indicate that cerebral atrophy is detectable on the shorter term, little is known about the statistical power and required sample size for detecting significant treatment effects in short-term clinical trials in MS using cerebral atrophy as primary outcome measure. In a recent paper, sample sizes for various MRI brain atrophy

measures in RRMS patients were estimated for longer periods of follow-up (1–3 years) and showed that the so-called SIENA technique, an automated MRI atrophy measurement, yielded the most promising results [11].

Following these findings, in the present study, we aim to assess the feasibility of using SIENA-based cerebral atrophy as primary outcome measure in short-term phase 2 clinical trials in MS. First, we evaluate the rate of atrophy with SIENA in a large cohort of RRMS patients without effective treatment over a 6-month period. Then, the predictive and explanatory value of MRI outcome measures in relation to cerebral atrophy is assessed and, lastly, power calculations based on the detected rate of atrophy are performed to determine the number of patients required in short-term placebo-controlled clinical trials using the rate of cerebral atrophy as primary outcome measure.

Methods

Patients

Our analyses were performed with data derived from the oral interferon beta-1a (IFNB-1a) study [12]. In this study, 173 patients with active RRMS received various doses of IFNB-1a or placebo orally every other day for 6 months. No clinical effect (median expanded disability status scale [EDSS] was 2.0 in all treatment groups at screening and at the end of study; approximately two thirds of patients in each group remained relapse-free) or MRI effect (median cumulative numbers of newly active lesions over 6 months were 4.0 in the placebo and 0.6 MIU groups, compared with 7.5 and 9.0 in the 0.06 and 6 MIU groups [no significant differences]) of any dose could be observed, including low neopterin levels in a subgroup of 21 patients and the absence of neutralizing antibodies in a subgroup of 24 patients; oral IFNB-1a was assumed to be biologically inactive and the cohort is regarded as a cohort representative of the placebo arm of a randomized trial.

MRI acquisition and analysis

MRI of the brain was performed at baseline and on six subsequent monthly MRI scans. Scans were performed, including a dual-echo, T2-weighted, spin-echo, or turbo/fast spin-echo (TR/TE of 2,000–3,000/20–40 and 60–100 ms) and a T1-weighted spin-echo (TR/TE of 400–700/5–25 ms), both after administration of 0.1 mmol/kg gadolinium (Gd)-DTPA intravenously, with a field of view of 25 cm and a 256×256 matrix resulting in roughly 1×1 mm pixel size. Images were acquired in 2×23 interleaved sections with a 3-mm thickness and a 3-mm gap, in accordance with published guidelines for the use of MRI in clinical trials

[13]. In addition to conventional T1- and T2-weighted MRI measures, the baseline normalized brain volume (NBV) and percentage brain volume change (PBVC) over 6 months was assessed using the automated segmentation-based techniques SIENAX and SIENA, respectively [14].

Statistical analysis

Primary outcome measure was the atrophy rate over 6 months, which followed the normal distribution. Comparisons for demographic and MRI characteristics between included and excluded patients were assessed by independent samples *t* tests and Mann–Whitney *U* test for continuous variables and the nonparametric binomial test for proportions. Correlations between variables were assessed with Pearson's *R*. By means of multiple linear regression analysis (SPSS version 13.0; SPSS Inc., Chicago, IL, USA), the predictive and explanatory value of baseline and on-study clinical and MRI variables for the PBVC over 6 months was determined. Independent variables included baseline NBV, presence of a Gd-enhancing lesion at the baseline scan, baseline number and volume of Gd-enhancing lesions, baseline number and volume of T2 lesions, on-study number and volume of Gd-enhancing lesions, on-study number and volume of T2-enhancing lesions, and on-study number of persistent black holes (PBH). A PBH was defined as new enhancing lesions or new T2 lesions (hyperintense on pd/T2, non-enhancing on T1) that appeared at month 1, 2, or 3 and became a black hole at months 4, 5, and 6, respectively [15]. Linearity in relation to PBVC was checked for all variables and natural log transformation was applied if a nonlinear relationship was found (to account for zero lesions, 1 was added prior to transformation). All effects were corrected for age, disease duration, and sex, and statistical testing was performed with

a two-sided test level of 5% with an additional Bonferroni correction for multiple testing.

Sample size calculations

Sample size estimates were based on the standard formula, assuming the rate of cerebral atrophy to be normally distributed:

$$n_1 = n_2 = \frac{2(z_{1-\frac{\alpha}{2}} + z_{1-\beta})^2 \sigma^2}{(u_2 - u_1)^2} \tag{1}$$

with σ^2 as the standard deviation and μ_1 and μ_2 as the mean brain volume atrophy in the placebo and treatment groups, respectively.

Sample sizes were determined for a trial duration of 6 months to detect a treatment effect of 50% to 90% reduction in atrophy rate at 80% and 90% power with and without taking into account a 5% dropout rate and a two-tailed significance level of 5%. Since atrophy rates of healthy controls were unavailable, we assumed a 100% treatment effect to correspond to zero brain volume loss. Treatment effects were assumed immediate and constant. To assess the impact of patient selection at baseline on the required sample size, subgroup analyses were performed for patients selected for the presence of an enhancing lesion at baseline and patients selected for a high T2 lesion load (greater than median) at baseline.

Results

Baseline demographics and MRI characteristics of the included and excluded patients are given in Table 1. From the original 173 patients, four patients lacked MRI assess-

Table 1 Baseline demographics and MRI characteristics.

Characteristic	Included	Excluded	<i>p</i> value
General			
<i>n</i>	135	34	
Sex (female/male)	98/37	25/9	0.44
Age, mean (SD)	35.8 (8.6)	33.7 (7.4)	0.18
Disease duration in years, mean (SD)	6.6 (5.5)	5.3 (4.5)	0.19
Baseline EDSS, median (IQR)	2.0 (1.5–3.5)	2.0 (1.5–2.5)	0.15
MRI			
Number of patients with ≥1 T1 Gd+ lesion (%)	80 (59)	19 (56)	0.07
Number of Gd+ lesions, mean (SD)	2.40 (4.3)	2.0 (3.4)	0.39
Volume of Gd+ lesions in cm ³ , mean (SD)	0.18 (0.6)	0.12 (0.27)	0.54
Number of T2 lesions, mean (SD)	89.1 (74.3)	102.8 (70.0)	0.33
Volume of T2 lesions in cm ³ , mean (SD)	7.39 (7.6)	6.30 (5.7)	0.43
Volume of T2 lesions in cm ³ , median (SD)	5.02 (7.6)	5.0 (5.7)	0.43
NBV in cm ³ , mean (SD)	1,449 (77.5)	1,481 (71.5)	0.49

Baseline demographics and MRI characteristics of the included and excluded patients, with the accompanying *p* value for the difference between both groups
SD standard deviation, *IQR* interquartile range, *EDSS* expanded disability status scale, *Gd+* gadolinium-enhancing

Table 2 On-study MRI measures.

Variable	Baseline selection criteria				
	No selection criteria	Enhancement present	Enhancement absent	High T2 lesion load	Low T2 lesion load
<i>n</i>	135	80	55	67	68
Cumulative number of Gd+ lesions, mean (SD)	10.8 (13.9)	16.2 (15.4)	2.93 (5.3)	16.3 (16.6)	5.4 (7.5)
Cumulative volume of Gd+ lesions in cm ³ , mean (SD)	1.39 (2.07)	2.10 (2.35)	0.35 (0.84)	2.18 (2.54)	0.61 (1.0)
On-study T2 lesion volume change in cm ³ , mean (SD)	+0.18 (2.01)	+0.15 (2.38)	+0.21 (1.33)	+0.16 (2.75)	+0.20 (0.77)
Cumulative number of PBH, mean (SD)	1.71 (42.9)	2.73 (3.4)	0.24 (0.6)	2.6 (3.6)	0.8 (1.5)
PBVC, mean (SD)	-0.33 (0.70)	-0.39 (0.67)	-0.25(0.75)	-0.42 (0.72)	-0.25 (0.69)
Percentage of inactive patients ^a	18	5	36	10	25

On-study MRI measures for whole cohort and subgroups selected for the presence or absence of an enhancing lesion at baseline or a high or low T2 lesion load (defined as above and below the median=5.02 cm³ T2 lesion load, respectively) at baseline

Gd+ gadolinium-enhancing, SD standard deviation, PBH persistent black hole, PBVC percentage brain volume change

^aNo on-study Gd+ lesions

ments after either the baseline or month 1 scan and were not included in the analyses. From the remaining 169 patients, the rate of cerebral atrophy could be assessed in 135 patients. From the 34 excluded patients, 12 patients from the site Graz were not analyzed, and for 22 patients, the available scans were rejected due to image quality. Patients had a mean age of 35.8 years (SD=8.6), a mean disease duration of 6.6 years (SD=5.5), and a median baseline EDSS score of 2.0 (interquartile range=1.5–3.5). No significant differences in disease duration, age, EDSS score, and baseline MRI characteristics were observed between the studied group and the 34 excluded patients. The mean NBV at baseline was 1,494 ml (SD=77.0). NBV was moderately correlated with the EDSS score at baseline ($R=-0.32$, $p<0.0001$) and T2 lesion load and number of T2 lesions at baseline ($R=-0.38$, $p<0001$ and $R=-0.29$, $p=0.01$, respectively) and weakly correlated with age ($R=-0.19$, $p=0.03$; data not shown).

On-study MRI outcome measures are shown in Table 2. For the 6-month follow-up period, a significant loss of brain volume was detected, comparing the baseline scan

with the month 6 scan (PBVC=-0.33%, SD=0.70, $p<0.0001$). Patients selected for the presence of an enhancing lesion at baseline or those selected for a high T2 lesion load at baseline tended to have a more pronounced atrophy (PBVC=-0.39%, SD=0.68 and PBVC=-0.42%, SD=0.73, respectively), compared with patients without an enhancing lesion at baseline or a low T2 lesion load at baseline (PBVC=-0.25%, SD=0.75 and PBVC=-0.25%, SD=0.69, respectively).

Table 3 shows the key results of the regression analyses assessing the relationship between MRI variables and PBVC. With a Bonferroni-corrected alpha of 0.005, the number of PBHs ($p=0.003$) was significantly associated with an increased rate of atrophy, whereas the number of T2 lesions at baseline ($p=0.024$) and the cumulative volume of enhancing lesions ($p=0.044$) failed to predict a higher rate of atrophy over 6 months. The accompanying explained variances are low, indicating that cerebral atrophy is marginally explained by variation in lesional measures alone.

The estimated sample sizes are shown in Table 4. Compared to the estimates based on patients unselected at

Table 3 Regression analysis results.

	Time of measurement	MRI variable	Regression coefficient	<i>p</i> value	<i>R</i> ²
Results of univariate multiple linear regression analysis for the relation between PBVC and predictive/explanatory MRI variables. All models were corrected for age, sex, and disease duration	Baseline	NBV	0.304	0.802	0.004
		Gd+ lesion number	-0.108	0.144	0.018
		Gd+ lesion volume	-0.028	0.252	0.013
		T2 lesion number	-0.175	0.024	0.042
		T2 lesion volume	-0.108	0.069	0.028
NBV normalized brain volume, Gd+ gadolinium-enhancing, PBH persistent black hole	On-study	Gd+ lesion number	-0.073	0.160	0.016
		Gd+ lesion volume	-0.044	0.044	0.033
		T2 lesion number	-0.003	0.220	0.013
		T2 lesion volume	-0.112	0.060	0.028
		PBH	-0.064	0.003	0.067

Table 4 Sample size estimates.

	Baseline selection criteria	Power	Treatment effect size (% reduction in atrophy rate)				
			50%	60%	70%	80%	90%
Number of patients per treatment arm necessary to perform parallel group designed trials with statistical powers of 80% or 90%, with and without taking into account a 5% dropout rate, to detect treatment effects ranging from 50% to 90%	No selection criteria at baseline	80%	283	196	144	110	87
		80% (5% dropout)	297	207	152	116	92
		90%	378	262	193	148	117
		90% (5% dropout)	398	276	203	155	123
	Presence of an enhancing lesion	80%	191	133	97	75	59
		80% (5% dropout)	201	140	103	78	62
		90%	255	177	130	100	79
		90% (5% dropout)	269	187	137	105	83
	High T2 lesion load (greater than median, >5,024 mm ²)	80%	185	128	94	72	57
		80% (5% dropout)	194	135	99	76	60
		90%	247	171	126	96	76
		90% (5% dropout)	260	180	133	101	80

baseline, both selection for enhancement at baseline as well as selection for a high T2 lesion volume at baseline were of favorable influence on the number of required patients. The calculations show that, for a treatment effect of 50% decrease in rate of atrophy, approximately 283 patients per arm are required when no selection criterion is applied and 185 when patients were selected for a high T2 lesion load at baseline, whereas for a 70% decrease in rate of atrophy, 144 patients per arm are required in a unselected cohort and 94 per arm in a selected cohort.

Discussion

As a compound measure of the overall destruction, preservation, and repair of brain tissue in MS patients, cerebral atrophy encompasses both neuroaxonal loss as well as the processes of demyelination, remyelination, gliosis, and edema. The present study shows that the rate of cerebral atrophy can be detected within a 6-month period and, when applied as primary outcome measure in short-term clinical trials with feasible sample size, requires a potent drug to obtain sufficient power.

Compared to sample size estimates for trials using atrophy measured with SIENA over longer periods of follow-up, the present sample sizes prove to be larger. To detect a 50% decrease in atrophy rate, trials for RRMS patients at 90% power showed approximately 69 patients per arm to be required over a 1-year follow-up period and 40 patients per arm over a 3-year follow-up period [11], whereas for trials for secondary progressive multiple sclerosis with 1-year follow-up, 56 patients per arm are required [16]. The lower sample sizes in these studies are explained by the larger atrophy rates due to larger trial durations and the accompanying larger detectable effect

sizes. Also, detection of atrophy over short intervals may be prone to increased measurement errors leading to a greater variability of the atrophy measure, thereby requiring larger subject numbers than would be expected in a longer-interval study. Since SIENA proved to accomplish larger statistical power due to greater measurement precision compared to other measures of atrophy [11], the current sample size estimates are likely the optimal achievable numbers for trials of short duration.

When interpreting the current sample sizes, some considerations should be taken into account. First, the calculations assumed a 100% treatment effect to resemble zero loss of brain volume, whereas healthy controls are known to experience a small amount of brain volume loss. For comparison, a previous study showed a PBVC of 0.11% (0.30) within healthy controls for a follow-up duration of 1 year [11]. When taken into account, a larger sample size will be required to detect a similar treatment effect. Second, the treatment effects are assumed to be effective from onset and constant over time. When this assumption is not met and a compound takes time to become maximally effective, the detected effect sizes will decrease and subsequently increase the required sample size. Wallerian degeneration initiated by axonal damage prior to treatment, for example, might result in a delay of the true effect caused by the already ongoing atrophic processes at initiation of the drug. Another important consideration is the confounding effect of other factors influencing brain volume such as demyelination, remyelination, gliosis, and inflammation. In particular, the resolution of edema and inflammation induced by anti-inflammatory agents, a process known as “pseudoatrophy” [6], can cause loss of brain volume and cloud the measurement of true tissue loss. In order to measure a true neuroprotective effect, especially within a shorter period of time, a future trial might assess the

neuroprotective effect of an experimental treatment as an add-on therapy to immunomodulatory-treated patients in both arms of the trial. Such a design, however, will likely result in higher sample sizes because of decreased rates of cerebral atrophy in the groups compared.

To partly overcome the aforementioned limitations, a more effective trial design would be to perform a run-in period of, e.g., 3 months in which the neuroprotective compound is administered and subsequently perform the short-interval atrophy assessment, thereby providing the opportunity for the applied compound to become maximally effective and confounding processes initiated prior to the trial to wane off.

An advantage of our study is that the present calculations are applicable in multicenter trials since the underlying data were obtained in multiple centers, with the accompanying variability caused by varying scanners and analyses. Also, we found a moderate but highly significant PBVC of -0.33% in a group of untreated RRMS patients within a period of 6 months. When annualized, this rate is well within the range of previous results ($0.6\text{--}1.35\%/year$ [6]) which enhances the generalizability of the sample size estimations since the calculations have not been biased by an atrophy rate at the higher end of the range.

A possible gain in statistical power for shorter-termed trials using cerebral atrophy as primary outcome measure might be achieved by adding multiple scanning time points to a trial. A recent study showed that, by placing additional scans towards the start and end of the trial, reductions in total variance and hence reductions in trial size of 41% could be achieved in patients with Alzheimer disease, using the brain boundary shift integral method for determining the rate of cerebral atrophy [17]. In particular, due to the within-subject variance contributing more to the overall variance at short intervals, acquiring multiple scans has more impact in shorter studies. Although relatively smaller gains in power can be expected from adding time points for more precise measures such as SIENA, the effect on the required sample size should definitely be explored in future multi-time point MS atrophy data.

The on-study association of the rate of atrophy and the number of PBHs suggest axonal loss to be one of the driving mechanisms of brain volume loss in MS patients. Previous studies showed the on-study volume of black holes to be closely related to supratentorial brain volume [18] and the baseline volume of black holes to be significantly correlated with subsequent development of atrophy [10, 19]. In contrast, the present study was not able to show a significant relationship between both Gd enhancement activity and T2 lesion load and changes in brain volume as shown in previous studies [8, 10, 19, 20]. These findings, together with the current results, show that the associations between focal tissue changes and atrophy

are moderate at best over a relatively long period of follow-up and weaken within shorter amounts of time, most probably due to the noise introduced by the increased variability in the measurements and that focal tissue changes in MS only partly explain atrophy development. This also emphasizes the added value of cerebral atrophy as a measure of the overall destruction of neuronal tissue, encompassing not only measurable focal destruction but also unaccounted diffuse destruction.

The detectability of atrophy on the short term has been attributed to the measure of activity of the patients assessed within a cohort. Hardmeier et al. [8], who found a significant atrophy rate within 3 months, stated that their finding most likely reflected the natural history of a very active group of RRMS patients with well-established disease and could explain the absence of atrophy in comparable short-termed studies [9, 21]. The current study population underwent partly MRI-based selection criteria at baseline and can be regarded as an active cohort of MS patients, shown by the increase in T2 lesion load and Gd-enhancing lesion number and volume over the study period and the high proportion of active patients. The influence of the measure of activity of patients on the rate of atrophy is also reflected when the subgroups based on the applied baseline selection criteria are compared (Table 2). Although a more active cohort might influence the generalizability compared to a more random sampled cohort of patients and selection criteria make it more difficult to recruit subjects, the trade-off is a larger sample size being required when using unselected patients.

In conclusion, our finding suggests that the rate of cerebral atrophy is a detectable outcome measure in short-term clinical trials in RRMS and applicable in terms of study power while a potent drug is applied.

Acknowledgements We are grateful to Bayer-Schering Pharma AG, Berlin, Germany and Rentschler Biotechnology GmbH, Laupheim, Germany for allowing us to use the data of the oral interferon study.

Conflict of interest statement We declare that we have no conflict of interest.

References

1. Fox NC, Jenkins R, Leary SM, Stevenson VL, Losseff NA, Crum WR, Harvey RJ, Rossor MN, Miller DH, Thompson AJ (2000) Progressive cerebral atrophy in MS. *Neurology* 54:807–812
2. Kalkers NF, Ameziane N, Bot JC, Minneboo A, Polman CH, Barkhof F (2002) Longitudinal brain volume measurement in multiple sclerosis: rate of brain atrophy is independent of the disease subtype. *Arch Neurol* 59:1572–1576
3. van Waesberghe JH, Kamphorst W, De Groot CJ, van Walderveen MA, Castelijns JA, Ravid R, Lycklama À, Nijeholt GJ, van der Valk P, Polman CH, Thompson AJ, Barkhof F (1999) Axonal loss

- in multiple sclerosis lesions: magnetic resonance imaging insights into substrates of disability. *Ann Neurol* 46(5):747–754
4. De Stefano N, Iannucci G, Sormani MP, Guidi L, Bartolozzi ML, Comi G, Federico A, Filippi M (2002) MR correlates of cerebral atrophy in patients with multiple sclerosis. *J Neurol* 249:1072–1077
 5. Vrenken H, Geurts JJ, Knol DL, van Dijk LN, Dattola V, Jasperse B, van Schijndel RA, Polman CH, Castelijns JA, Barkhof F, Pouwels PJ (2006) Whole-brain T1 mapping in multiple sclerosis: global changes of normal-appearing gray and white matter. *Radiology* 240(3):811–820
 6. Bermel RA, Bakshi R (2006) The measurement and clinical relevance of brain atrophy in multiple sclerosis. *Lancet Neurol* 5(2):158–170
 7. Fisher E, Rudick RA, Cutter G, Baier M, Miller D, Weinstock-Guttman B, Mass MK, Dougherty DS, Simonian NA (2000) Relationship between brain atrophy and disability: an 8-year follow-up study of multiple sclerosis patients. *Mult Scler* 6(6):373–377
 8. Hardmeier M, Wagenpfeil S, Freitag P, Fisher E, Rudlick RA, Kooijmans-Coutinho M, Clanet M, Radue EW, Kappos L, for the European rIFN B-1a in Relapsing MS Dose Comparison Trial Study Group (2003) Atrophy is detectable within a 3-month period in untreated patients with active relapsing remitting multiple sclerosis. *Arch Neurol* 60:1736–1739
 9. Zivadinov R, Bagnato F, Nasuelli D, Bastianello S, Bratina A, Locatelli L, Watts K, Finamore L, Grop A, Dwyer M, Catalan M, Clemenzi A, Millefiorini E, Bakshi R, Zorzon M (2004) Short-term brain atrophy changes in relapsing–remitting multiple sclerosis. *J Neurol Sci* 223(2):185–193
 10. Rovaris M, Comi G, Rocca MA, Wolinsky JS, Filippi M, European/Canadian Glatiramer Acetate Study Group (2001) Short-term brain volume change in relapsing–remitting multiple sclerosis: effect of glatiramer acetate and implications. *Brain* 124(9):1803–1812
 11. Anderson VM, Bartlett JW, Fox NC, Fisniku L, Miller DH (2007) Detecting treatment effects on brain atrophy in relapsing remitting multiple sclerosis: sample size estimates. *J Neurol* 254:1588–1594
 12. Polman C, Barkhof F, Kappos L, Pozzilli C, Sandbrink R, Dahlke F, European Oral Interferon Beta-1a in Relapsing–Remitting MS Study Group et al (2003) Oral interferon beta-1a in relapsing–remitting multiple sclerosis: a double-blind randomized study. *Mult Scler* 9(4):342–348
 13. Miller DH, Albert PS, Barkhof F et al (1996) Guidelines for the use of magnetic resonance techniques in monitoring the treatment of multiple sclerosis. US National MS Society Task Force. *Ann Neurol* 39:6–16
 14. Smith SM, De Stefano N, Jenkinson M, Matthews PM (2001) Normalized accurate measurement of longitudinal brain change. *J Comput Assist Tomogr* 25(3):466–475
 15. van den Elskamp IJ, Lembcke J, Dattola V et al (2008) Persistent T1 hypointensity as an MRI marker for treatment efficacy in multiple sclerosis. *Mult Scler* 14(6):764–769
 16. Altmann DR, Jasperse B, Barkhof F, Beckmann K, Filippi M, Kappos LD, Molyneux P, Polman CH, Pozzilli C, Thompson AJ, Wagner K, Yousry TA, Miller DH (2009) Sample sizes for brain atrophy outcomes in trials for secondary progressive multiple sclerosis. *Neurology* 72(7):595–601
 17. Schott JM, Frost C, Whitwell JL et al (2006) Combining short interval MRI in Alzheimer's disease: implications for therapeutic trials. *J Neurol* 253:1147–1153
 18. Paolillo A, Pozzilli C, Gasperini C, Giugni E, Mainero C, Giuliani S, Tomassini V, Millefiorini E, Bastianello S (2000) Brain atrophy in relapsing–remitting multiple sclerosis: relationship with ‘black holes’, disease duration and clinical disability. *J Neurol Sci* 174(2):85–91
 19. Jasperse B, Minneboo A, de Groot V, Kalkers NF, van Helden P, Uitdehaag BMJ, Barkhof F, Polman C (2007) Determinants of cerebral atrophy rate at the time of diagnosis of multiple sclerosis. *Arch Neurol* 64:190–194
 20. Rudick RA, Fisher E, Lee JC, Simon J, Jacobs L, MS Collaborative Research Group (1999) Use of the brain parenchymal fraction to measure whole brain atrophy in relapsing–remitting MS. *Neurology* 53:1698–1704
 21. Gasperini C, Paolillo A, Giugni E, Galgani S, Bagnato F, Mainero C, Onesti E, Bastianello S, Pozzilli C (2002) MRI brain volume changes in relapsing–remitting multiple sclerosis patients treated with interferon beta-1a. *Mult Scler* 8(2):119–123