Assessing brain atrophy rates in a large population of untreated multiple sclerosis subtypes

N. De Stefano, MD
A. Giorgio, MD
M. Battaglini, MSc
M. Rovaris, MD
M.P. Sormani, PhD
F. Barkhof, MD
T. Korteweg, MD
C. Enzinger, MD
F. Fazekas, MD
M. Calabrese, MD
D. Dinacci, MD
G. Tedeschi, MD
A. Gass, MD
X. Montalban, MD
A. Rovira, MD
A. Thompson, MD
G. Comi, MD
D.H. Miller, MD
M. Filippi, MD

ABSTRACT

Objective: To assess the time course of brain atrophy and the difference across clinical subtypes in multiple sclerosis (MS).

Methods: The percent brain volume change (PBVC) was computed on existing longitudinal (2 time points) T1-weighted MRI from untreated (trial and nontrial) patients with MS. Patients (n = 963) were classified as clinically isolated syndromes suggestive of MS (CIS, 16%), relapsing-remitting (RR, 60%), secondary progressive (SP, 15%), and primary progressive (9%) MS. The median length of follow-up was 14 months (range 12–68).

Results: There was marked heterogeneity of the annualized PBVC (PBVC/y) across MS subtypes ($p = 0.003$), with higher PBVC/y in SP than in CIS ($p = 0.003$). However, this heterogeneity disappeared when data were corrected for the baseline normalized brain volume. When the MS population was divided into trial and nontrial subjects, the heterogeneity of PBVC/y across MS subtypes was present only in the second group, due to the higher PBVC/y values found in trial data in CIS ($p = 0.01$) and RR ($p < 0.001$). The estimation of the sample sizes required for demonstrating a reduction of brain atrophy in patients in a placebo-controlled trial showed that this was larger in patients with early MS than in those with the progressive forms of the disease.

Conclusions: This first large study in untreated patients with multiple sclerosis (MS) with different disease subtypes shows that brain atrophy proceeds relentlessly throughout the course of MS, with a rate that seems largely independent of the MS subtype, when adjusting for baseline brain volume. *Neurology*® 2010;74:1868–1876

GLOSSARY

ANOVA = analysis of variance; **CIS** = clinically isolated syndrome; **DMA** = disease-modifying agent; **EDSS** = Expanded Disability Status Scale; **FSL** = FMRIB Software Library; **Gd** = gadolinium; **MR** = magnetic resonance; **MS** = multiple sclerosis; **NBV** = normalized brain volume; **PBVC** = percent brain volume change; **PP** = primary progressive; **RR** = relapsing-remitting; **SP** = secondary progressive.

Address correspondence and reprint requests to Prof. Nicola De Stefano, Quantitative Neuroimaging Laboratory, Department of Neurological and Behavioral Sciences, University of Siena, Viale Bracci 2, 53100 Siena, Italy.

A number of MRI-based methods for computed estimation of brain volumes[1,2] have prompted the use of brain atrophy as a measure of disease progression in multiple sclerosis (MS). However, the interpretation of brain volume change measurements is not always straightforward in MS and a number of confounding factors such as disease stage and disease-modifying agents (DMA) need to be considered.[2,3]

The natural evolution of global brain volume changes at different MS stages and without the influence of DMA has been investigated for each subtype on patients belonging to the placebo

Supplemental data at www.neurology.org

From the Department of Neurological and Behavioural Sciences (N.D.S., A.G., M.B.), University of Siena; Neuroimaging Research Unit (M.R., M.F.), Institute of Experimental Neurology, Division of Neuroscience, Scientific Institute and University Ospedale San Raffaele, Milan; Biostatistics Unit (M.P.S.), Department of Health Sciences, University of Genoa, Genoa, Italy; Department of Radiology (F.B., T.K.), VU University Medical Centre, Amsterdam, the Netherlands; Department of Neurology (C.E., F.F.), Medical University of Graz, Graz, Austria; The Multiple Sclerosis Centre of Veneto Region (M.C.), First Neurology Clinic, Department of Neurosciences, University Hospital of Padua, Padua; Department of Neurological Sciences (D.D., G.T.), Second University of Naples and Institute Hermitage Capodimonte, Naples, Italy; Department of Neurology (A.G.), University Hospital, Kantonsspital, Basel, Switzerland; Department of Neuroimmunology and Radiology (X.M., A.R.), Hospital Vall d'Hebron, Barcelona, Spain; Department of Brain Repair and Rehabilitation (A.T.), Institute of Neurology, University College London, London, UK; Department of Neurology (G.C.), Institute of Experimental Neurology, Division of Neuroscience, Scientific Institute and University Ospedale San Raffaele, Milan, Italy; and Department of Neuroinflammation (D.H.M.), Institute of Neurology, University College London, London, UK. M.R. is currently affiliated with the Multiple Sclerosis Centre, Scientific Institute S. Maria Nascente, Fondazione Don Gnocchi, Milan, Italy.

Study funding: The Quantitative Neuroimaging Laboratory was supported in part by the Italian MS Foundation. The VU University Center for MS Research (05-358c) and T. Korteweg (00–425 MS) are supported by the Dutch MS Research Foundation.

arms of clinical trials.[4-21] Further, research studies have been performed on patient groups partially treated with DMA or with a limited number of subjects.[22-26] To date, no study has assessed temporal brain volume changes in a large untreated MS population, directly comparing different subtypes.

Thus, we collected a large number of existing MRI data on untreated patients with different MS subtypes and analyzed them with a fully automated method for the estimation of global brain volume changes.[27] We aimed to assess 1) differences in annualized global brain volume changes; 2) potential differences in brain atrophy rates between patients from the placebo arms of clinical trials (trial data) and those who remained untreated for that given follow-up period (nontrial data); 3) the sample sizes required to demonstrate a treatment-related reduction of brain atrophy progression in placebo-controlled MS trials.

METHODS

Study population. This is a European multicenter retrospective study based on the analysis of longitudinal magnetic resonance (MR) datasets (2 time points) of patients with different subtypes of MS who were either collected at different imaging laboratories while not taking any DMA or in the placebo arms of clinical trials. Minimum between-scan interval was 12 months. There was no limitation for the Expanded Disability Status Scale (EDSS)[28] score at study entry. The main inclusion criterion was the complete absence of DMA use during the study period. This did not include the use of steroids, but all the patients with MS had to be corticosteroid-free for at least 1 month before scanning.

A total of 1,160 pairs of T1-weighted MRIs was collected from 193 patients with a clinically isolated syndrome (CIS) suggestive of MS, 642 with relapsing-remitting (RR), 192 with secondary progressive (SP), and 133 with primary progressive (PP) MS.[29] The MRIs were obtained from data of imaging laboratories (Amsterdam, Barcelona, Basel, London, Milan, Naples, and Padua) and placebo arms of clinical trials (ETOMS,[30] CORAL,[12] European/Canadian Glatiramer Acetate Study,[31] and ESIMS[19]).

The median length of follow-up was 14 months (range 12–42), with the exception of 3 RR patients who had the second MRI scan at 48 months and 1 RR patient who had it at 68 months. The median was 24 (range 12–30) for CIS, 14 (range 12–68) for RR, 24 (range 12–40) for SP, and 13 (range 12–42) for PP. Clinical (i.e., disease duration and EDSS score) and demographic (i.e., age and sex) information were collected.

Standard protocol approvals, registrations, and patient consents. The study received approval from the local ethics committee and written informed consent was obtained from all study patients.

MRI data and analysis. All MR scans were acquired at each center using for each patient the same MR procedure/sequences and scanner at both time points. None of the scanners were Conventional T1-weighted images were sent to the Quantitative Neuroimaging Laboratory of the University of Siena for centralized analysis.

Global brain volume changes over time were quantified using the SIENA method,[32] part of the FMRIB Software Library (FSL; www.fmrib.ox.ac.uk/fsl/). This registration-based method uses images from 2 time points to assess brain volume changes by estimating directly the local shifts in brain edges across the entire brain and then converting the edge displacement into a global estimate of percentage brain volume change (PBVC) between the 2 time points.

An automated procedure of brain extraction able to improve removal of eyeballs and remaining nonbrain tissues[33] was implemented in SIENA for a more accurate estimation of brain atrophy.

The scans were all T1-weighted pairs of images obtained either prior to or after injection of gadolinium (pre-Gd and post-Gd). Slice thicknesses (range 1.2–5 mm) were identical in each pair of images. It was shown that use of different image types (pre-Gd and post-Gd T1-weighted images)[2,34] and slice thicknesses[32] does not systematically affect SIENA measurements. However, differences in both T1-weighted image type and slice thicknesses were corrected for during the analysis.

Statistical analysis. Changes in EDSS score (ΔEDSS) and PBVC were annualized (i.e., ΔEDSS/y and PBVC/y) to account for differences in the length of follow-up between the 2 scans.

A univariate analysis of variance (ANOVA) followed by pairwise post hoc comparisons were used to compare PBVC/y across the different MS subtypes. A multivariate ANOVA was used to adjust the comparisons for age, sex, disease duration, data source (trial/nontrial), T1-weighted image type (pre-Gd and post-Gd), and slice thickness. Furthermore, PBVC/y values were compared across the disease subtypes by correcting for the baseline normalized brain volume (NBV) as measured on the T1-weighted image by using the cross-sectional version of SIENA (SIENAX), also part of FSL.

Correlations (unadjusted for baseline values) of PBVC/y with demographic (age and sex) and clinical (disease duration, EDSS score at baseline, ΔEDSS/y) features were analyzed using the Spearman rank correlation coefficient.

SPSS software v11.0 (SPSS Inc., Chicago, IL) was used to perform statistical calculations. A 2-tailed p value of 0.05 was used as the cutoff for significance.

The sample size required to demonstrate a treatment-related reduction in brain atrophy progression in placebo-controlled MS trials was estimated using PBVC/y as the primary outcome for each disease subtype. This was estimated to have a power of 90% at a confidence level of 5% for each disease subtype and to detect a treatment effect of 30%, 50%, and 70%. A nonparametric approach based on Monte Carlo simulations was used. The sample size was estimated by the nonparametric Mann-Whitney U test comparing PBVC/y between the 2 arms.

RESULTS

Clinical and demographic information of the study population as well as MRI features are summarized in table 1.

Out of 1,160 pairs of T1-weighted images, 197 were excluded from the analysis for unsatisfactory quality (n = 143) or incomplete demographic or clinical information (n = 54). The final number of

Table 1 Demographic, clinical, and magnetic resonance features of study patients

	All patients (n = 963)	CIS (n = 157, 16%)	RR (n = 579, 60%)	SP (n = 139, 15%)	PP (n = 88, 9%)
Age, y, mean (SD)	37.88 (10)	31.35 (8.11)	36.18 (8.24)	45.29 (8.58)	48.93 (11)
Sex, n (%)					
Males	333 (35)	58 (37)	160 (28)	63 (45)	52 (59)
Females	630 (65)	99 (63)	419 (72)	76 (55)	36 (41)
Disease duration, y, mean (SD)	7.25 (7.49)	0.64 (1.45)	7 (6.45)	15.44 (8.20)	7.77 (6.88)
Baseline EDSS					
Median	2	1	2	5	5
Range	0 to 8	0 to 4.5	0 to 8	1.5 to 8	1.5 to 7.5
EDSS change/y					
Median	0	0	0	0	0.2
Range	−3 to 3	−3 to 2	−1.5 to 3	−1 to 2.5	−1.5 to 3
T1-weighted image type, n (%)					
Post-Gd	432 (45)	18 (11)	348 (60)	44 (32)	22 (25)
Pre-Gd	531 (55)	139 (89)	231 (40)	95 (68)	66 (75)
Source, n (%)					
Nontrial	390 (40)	74 (47)	173 (30)	54 (39)	100
Trial	573 (60)	83 (53)	406 (70)	85 (61)	NA
Slice thickness					
1.2 mm	64 (7)	21 (14)	37 (6)	3 (2)	3 (4)
1.5 mm	16 (2)	NA	NA	NA	16 (18)
3 mm	446 (46)	13 (8)	416 (72)	NA	17 (19)
4 mm	54 (5)	NA	49 (8)	5 (4)	NA
5 mm	383 (40)	123 (78)	77 (14)	131 (94)	52 (59)

Abbreviations: CIS = clinically isolated syndrome; EDSS = Expanded Disability Status Scale; Gd = gadolinium; NA = not available; PP = primary progressive; RR = relapsing-remitting; SP = secondary progressive.

16%), RR (n = 579, 60%), SP (n = 139, 15%), and PP (n = 88, 9%).

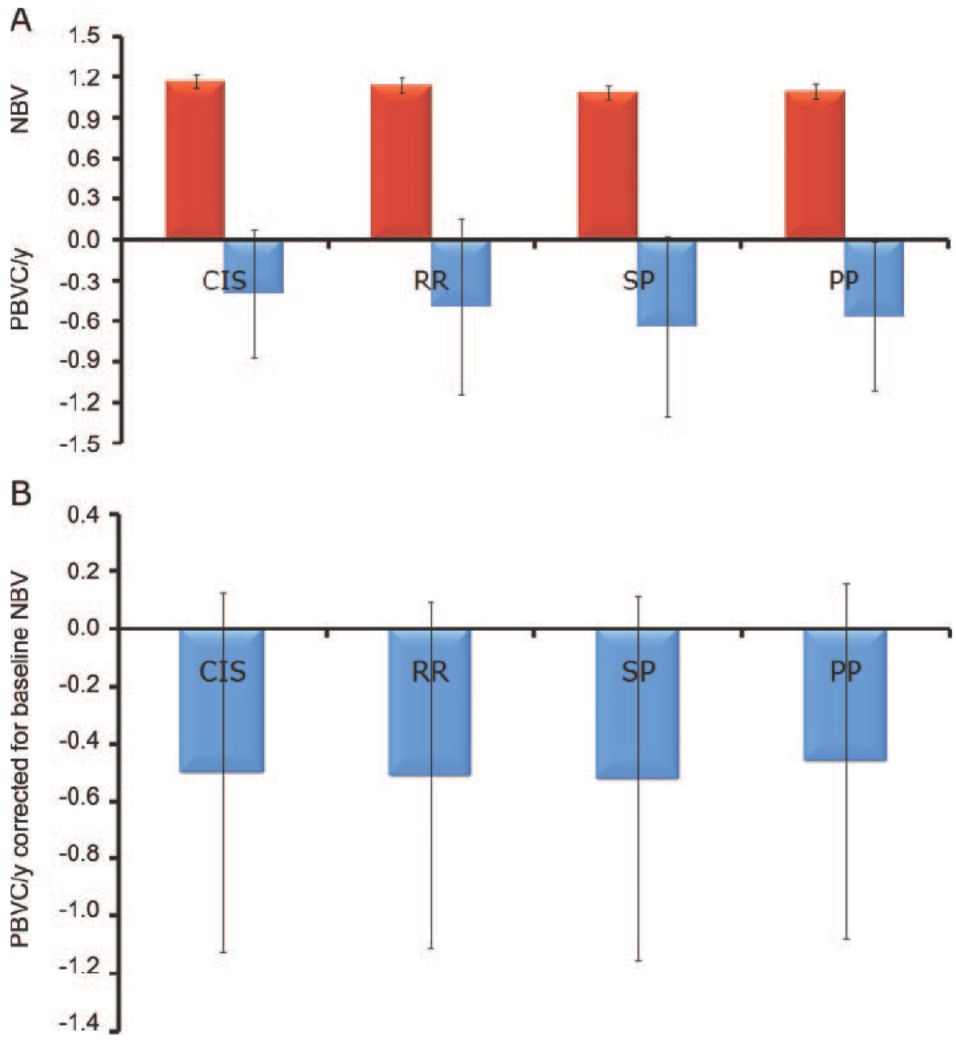
Comparisons across MS subtypes. ***All patients.*** There was heterogeneity of PBVC/y across MS subtypes (PBVC/y, mean ± SD: CIS = −0.40% ± 0.47%, RR = −0.49% ± 0.65%, SP = −0.64% ± 0.68%, PP = −0.56% ± 0.55%, $p = 0.003$), with the pairwise comparisons showing higher PBVC/y in SP patients than in patients with CIS ($p = 0.003$) (figure 1A). The between-group difference was maintained after correcting for age, sex, disease duration, slice thickness, data source (trial/nontrial), and T1-weighted image type (pre-Gd and post-Gd) ($p < 0.001$ at multivariate analysis).

Among the patients with CIS, the 47 subjects who converted to clinically definite MS showed higher PBVC/y than the 110 subjects who did not (−0.51% ± 0.48% vs −0.35% ± 0.47%, $p = 0.04$).

As expected, baseline NBV was different across MS subtypes (NBV, mean ± SD: CIS = 1,169 ± 47 cm^3, RR = 1,140 ± 53 cm^3, SP = 1,089 ± 50 cm^3, comparisons across MS subtypes showing differences ($p < 0.001$) (with the exception of the comparison between SP and PP) (figure 1A). Interestingly, when PBVC/y values were corrected for the baseline NBV, the heterogeneity of PBVC/y across MS subtypes disappeared ($p = 0.90$) (figure 1B). It should be noted that despite significant differences, there was a degree of overlap for baseline NBV across MS subtypes. Figure 2 illustrates this overlap for CIS and SP, which represent the 2 extreme situations. Thus, controlling for baseline NBV should not lead to major extrapolation in comparing PBVC/y values across MS subtypes.

Trial vs nontrial data. The heterogeneity of PBVC/y across MS subtypes detected on the whole population was still present when the analysis was performed on nontrial data (PBVC/y, mean ± SD: CIS = −0.29% ± 0.43%, RR = −0.34% ± 0.45%, SP = −0.65% ± 0.65%, PP = −0.56% ± 0.55%, $p < 0.001$). However, when the analysis was selectively performed on trial data, there was no heterogeneity of PBVC/y across MS subtypes (PBVC/y,

Figure 1 Brain atrophy measures in the different multiple sclerosis (MS) subtypes



(A) Values of percent brain volume change (PBVC)/y (blue) and normalized brain volume (NBV) (expressed in liters, red) in the different MS subtypes. (B) Values of PBVC/y corrected for baseline NBV. Columns and error bars represent means and SDs of atrophy measures. Note the similar PBVC/y in the different MS subtypes when data are corrected for baseline NBV. See Results for details. CIS = clinically isolated syndrome; PP = primary progressive; RR = relapsing-remitting; SP = secondary progressive.

$-0.55\% \pm 0.71\%$, SP $= -0.63\% \pm 0.69\%$, $p = 0.28$) and so remained after correcting for age, disease duration, and baseline NBV ($p > 0.10$ for all). This difference between the 2 datasets (trial/nontrial) was mainly due to the higher PBVC/y values found in trial than in nontrial data in both patients with CIS (PBVC/y, mean $\pm$ SD: $-0.49\% \pm 0.50\%$ vs $-0.29\% \pm 0.43\%$, $p = 0.01$) and RR patients (PBVC/y, mean $\pm$ SD: $-0.55\% \pm 0.71\%$ vs $-0.34\% \pm 0.45\%$, $p < 0.001$) (figure 3). These differences persisted after controlling the PBCV/y values for age, disease duration, and baseline NBV ($p < 0.05$ for all).

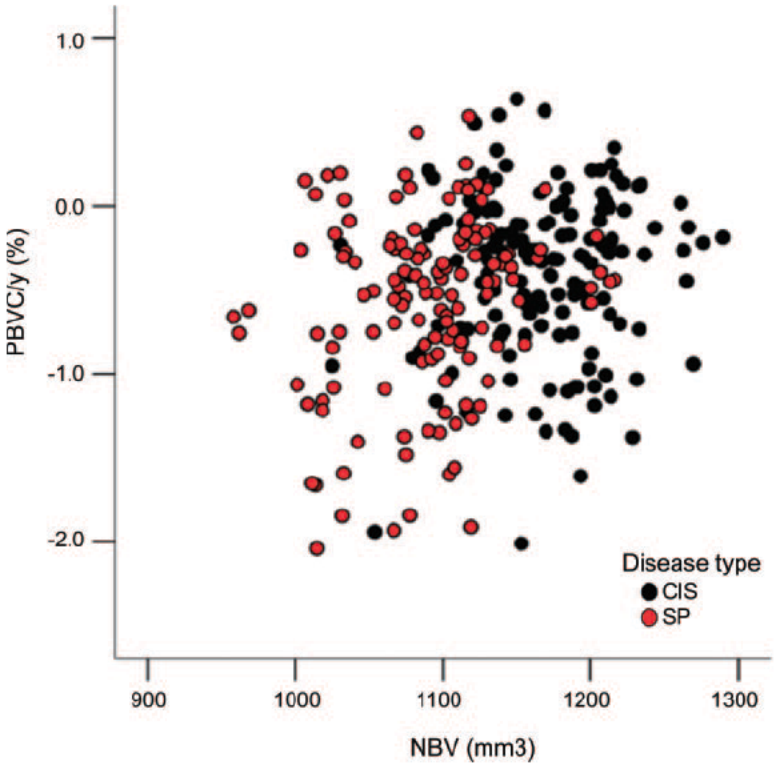
Interestingly, most (64%) of the patients with CIS who converted to clinically definite MS were verted CIS from nontrial data ($-0.58\% \pm 0.55\%$ vs $-0.38\% \pm 0.30\%$, $p = 0.10$).

Relationships of PBVC/y with clinical-demographic features. Overall, the correlations of PBVC/y with both demographic and clinical features were weak or absent.

Values of PBVC/y did not correlate with age, sex, or disease duration in the whole patient population, whereas they correlated weakly with baseline EDSS score ($r = -0.15$, $p < 0.001$) and ΔEDSS/y ($r = -0.10$, $p = 0.003$).

In the different MS subtypes, a weak correlation was found in CIS between PBVC/y and age ($r = 0.20$, $p = 0.01$) and disease duration ($r = 0.18$, $p = 0.02$). In RR,

Figure 2 Plot illustrating overlap of baseline normalized brain volume (NBV) values for patients with clinically isolated syndrome (CIS) and secondary progressive (SP) patients



See Results for details. PBVC = percent brain volume change.

baseline EDSS score ($r = -0.12$, $p = 0.003$). No significant correlations were found in SP and PP.

When data were grouped for source type, both populations of trial and nontrial data showed weak correlations between PBVC/y and baseline EDSS score ($r = -0.21$ and $r = -0.12$, $p < 0.005$ for both). All the other correlations were either weaker or absent (data not shown).

Sample size estimation for PBVC/y. The sample sizes required to give a statistical power of 90% at a significance level of 5% for different treatment effects (30%, 50%, and 70%) are summarized in table 2.

Since there were differences in clinical-demographic characteristics (table e-1 on the *Neurology*® Web site at www.neurology.org) as well as in PBVC/y values across disease subtypes when data were grouped for data source (trial vs nontrial), the sample sizes were estimated in these 2 different datasets. Lower sample size estimation was found for patients with CIS in trial (n = 70) than in nontrial (n = 170) data. By contrast, no differences were found for RR patients and SP patients.

DISCUSSION While measures of brain atrophy rates have been used as an endpoint in many MS plex structural and temporal mechanisms leading to brain atrophy in the different MS subtypes. In particular, the estimates of brain atrophy rates have varied in studies directly comparing different MS subtypes. These have shown higher[25,26] or similar[22,23] atrophy progression in SP patients when compared to those at earlier disease stages. Moreover, atrophy rate in different MS subtypes showed high variability in the placebo arms of clinical trials. Thus, we collected here a large number of longitudinal MRI data with a median follow-up of 14 months from untreated patients with MS with different disease subtypes and tested for differences in annualized brain atrophy rates in such a large population. We found heterogeneity in PBVC/y across MS subtypes, which was particularly evident when comparing patients at earliest with those at later disease stages. Interestingly, however, this heterogeneity disappeared when PBVC/y values were corrected for the baseline NBV. This suggests that, nulling out the differences in atrophy state, atrophy progression rate is very similar in the different MS subtypes and, at late disease stages, does not seem to show either the previously hypothesized nonlinear progression[1,22] or a true acceleration.[25,26] This hypothesis is further supported by the finding that the absolute (in cm^3) unadjusted changes were similar across disease subtypes when estimated, at the first order of approximation, by using the brain volume at the 2 time points as recorded in the so-called halfway space of the SIENA method[32] (data not shown).

Since it is well-known that measures of brain volume changes can be significantly influenced by treatment with DMA,[2,3] we have included in this study only patients who, during the entire follow-up period, were either in the placebo arms of clinical trials or untreated in natural history studies. Interestingly, in the analysis of these patient populations, we found that annualized brain atrophy rates were significantly higher in trial than in nontrial data in both patients with CIS and RR patients. Thus, these data provide direct evidence on how differences in patients' recruitment could significantly influence a measure such as brain atrophy rate even in a presumably homogeneous patient population. Assuming that patients with MS enrolled in a clinical trial are clinically more active than patients who decided to remain untreated, these results also suggest that significant differences in brain atrophy progression may exist between populations of untreated patients with a similar MS subtype and different disease severity. This seems particularly true in patients with CIS converting to clinically definite MS, who showed an almost twice as high brain atrophy rate in trial than

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.