Inter-centre agreement of brain atrophy measurement in Multiple Sclerosis patients using manually edited SIENAX and SIENA

B. Jasperse[1], P. Valsasina[2], V. Neacsu[3], D. L. Knol[4], N. De Stefano[5], C. Enzinger[6], S. M. Smith[7], S. Ropele[6], T. Korteweg[3], A. Giorgio[5], V. Anderson[8], C. H. Polman[9], M. Filippi[2], D. H. Miller[8], M. Rovaris[2], F. Barkhof[3], and H. Vrenken[3]

[1]Department of Neurology, VU Medical Centre, Amsterdam, Netherlands, [2]Neuroimaging Research Unit, Scientific Institute Ospedale San Raffaele, Milan, Italy, [3]Department of Radiology, VU Medical Centre, Amsterdam, Netherlands, [4]Department of Clinical Epidemiology and Biostatistics, VU medical centre, Amsterdam, Netherlands, [5]Department of Neurological and Behavorial Sciences, University of Siena, Siena, Italy, [6]Department of Neurology, Medical University Graz, Graz, Austria, [7]Oxford University Centre for Functional MRI of the Brain (FMRIB), University of Oxford, Oxford, United Kingdom, [8]MS NMR research unit, Department of Neuroinflammation, Institute of Neurology, University College London, London, United Kingdom, [9]Department of Neurology, VU medical centre, Amsterdam, Netherlands

Background: Brain atrophy is an important feature of multiple sclerosis (MS), thought to reflect the neurodegenerative component of the disease. To further elucidate the role of brain atrophy in MS, large multi-centre cohorts need to be studied. SIENAX and SIENA (1) are accurate and robust methods to quantify brain volume and brain atrophy rate respectively. Although these are automated methods, errors may arise in the classification of brain and non-brain tissue, which is frequently corrected by manual adjustment. Assessment of inter-centre agreement is needed to justify exchange of manually edited brain atrophy data between centres. Hence, the goal of this study was to investigate the inter-centre agreement of brain volume measurement using SIENAX and SIENA with and without manual editing.

Methods: Baseline and follow-up T1-weighted MRI scans from a total of 20 MS patients were collected from studies ongoing at two centres (i.e. centre 1 and centre 2). The 10 scan pairs from centre 1 were acquired on a 1.0 Tesla scanner (Siemens, Erlangen, Germany) with an interval ranging from 23 to 36 months, using the following scan parameters: repetition time [TR]: 700 ms, echo time [TE]: 15 ms, 2 excitations, 25 slices with a slice thickness of 5 mm and 10% gap. The 10 scan pairs from centre 2 were acquired on a 1.5 Tesla scanner (Siemens, Erlangen, Germany), with an interval ranging from 12-18 months, using the following scan parameters: [TR]: 600-768 ms, [TE]: 10-14 ms, 2 excitations. Among the scans from centre 2, 5 pairs consisted of 24 slices with a slice thickness of 5.0 mm and 5 consisted of 44 slices with a slice thickness of 3.0 mm. All scans were anonymised, converted to ANALYZE format and dispatched by CD to the five participating centres. *Image analysis:* All centres used the same version of SIENAX and SIENA as provided with FSL 3.2, and performed all analyses with identical parameter settings. Each centre performed fully automated and manually edited analyses for both SIENAX and SIENA, yielding Normalised Brain Volume (NBV) and Percentage Brain volume Change (PBVC) for each subject. *Statistics:* Group differences between fully automated and manually edited NBV and PBVC were tested by mixed model analysis. Variation between centres was analysed using variance component analysis. The extent of absolute agreement between centres was expressed as the Concordance Correlation Coefficient (CCC), which was calculated from the variance components (2). The difference in total variance between fully automated and manually edited SIENAX and SIENA was tested at a p-level of 0.05 (3).

Results: As expected, excellent agreement between centres was observed for both fully automated NBV and PBVC (both CCC=1.0, Table 1). Manual editing decreased agreement between centres (CCC=0.94 for NBV, 0.95 for PBVC, Table 1). Mean NBV values for each centre decreased significantly (difference: 2.8%, P<0.001) after manual editing, probably reflecting the closer resemblance to the 'true' brain volume (Fig. 1:A and B). PBVC values remained similar on average (P=0.88). Interestingly, the total variance for PBVC decreased significantly after manual editing (P<0.05), suggesting an increase in statistical power for SIENA. The latter is probably due to the exclusion of artificial change in non-brain tissue when manual editing is applied (Fig 1:C and D).

Conclusions: NBV and PBVC values from different centres show good agreement, even after manual editing. For SIENA, manual adjustment appears to increase statistical power.

	Total variance	Subject variance	Centre variance	Residual variance	CCC	Overall mean (SD)
NBV in cm3 fully automated	3529.7	3527.6	0.0	2.0	1.00	1651.1 (58.2)
NBV in cm3 manually edited	3776.9	3534.6	51.7	190.7	0.94	1605.3 (60.2)
PBVC fully automated	2.82*	2.81	$7.8 \cdot 10^{-6}$	$2 \cdot 10^{-3}$	1.00	-1.70 (1.64)
PBVC manually edited	1.54*	1.47	0.00	0.07	0.95	-1.74 (1.22)

Table 1. Variance component estimates and Concordance Correlation Coefficients for fully automated and manually edited SIENAX (NBV) and SIENA (PBVC). CCC=Concordance Correlation Coefficient; NBV=Normalised Brain Volume; PBVC=Percentage Brain Volume Change; *: PBVC variance was significantly lower for manually edited than for fully automated analyses (P<0.05).
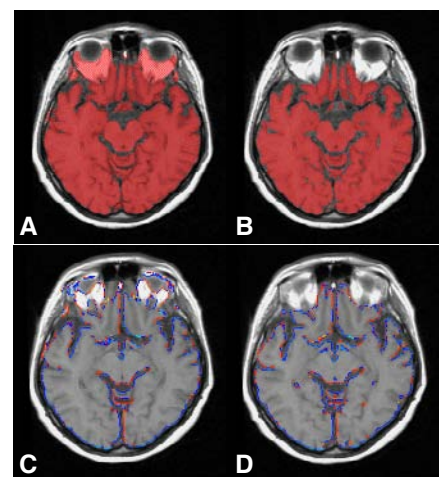

Figure 1. Images showing regions that are classified as brain tissue (red checkerboard) for fully automated (A) and manually edited (B) SIENAX. Images showing positive (red/yellow) and negative (blue/light blue) brain edge displacement for fully automated (C) and manually edited (D) SIENA.

1. Smith et al. NeuroImage 17:479-489, 2002; 2. Carrasco et al. Biometrics 59: 849-858, 2003; 3. Mood AM, Graybill FA, Boes DC. Introduction to the Theory of Statistics, 3rd edition, McGraw-Hill;1974: 438