Improved SIENAX assessment of WM and GM volumes

Marco Battaglini[1], Mark Jenkinson[2], Antonio Giorgio[1], Maria Assunta Rocca[3], Maria Laura Stromillo[1], Massimo Filippi[3], jacqueline Palace[4], and Nicola De Stefano[1]

[1]Medicine, Surgery and Neurosciences, University of Siena, Siena, Tuscany, Italy, [2]Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB), University of Oxford, Oxford, United Kingdom, [3]Neuroimaging Research Unit, Vita-Salute S.Raffaele University, Milano, Italy, [4]Clinical Neurology, University of Oxford, Oxford, United Kingdom

Target audience. This work will interest scientists involved in charting the course of Multiple Sclerosis (MS) by measuring volumetric changes of grey matter (GM) and white matter (WM). **Purpose** – Due to the complex pathophysiologic mechanisms of multiple sclerosis (MS), an accurate in-vivo assessment of cerebral grey matter (GM) and white matter (WM) damage is of paramount importance. New dedicated algorithms, such as the "lesion filling" procedure[(1,2)], have been developed in recent years to face issues that are inherent to the specific nature of MR images from diseased brains (i.e., MS), to improve WM and GM volume assessment. However, despite the number of existing imaging analysis approaches able to perform brain tissue class segmentation, more robust and accurate MR-derived measures of GM and WM are warranted. The purpose of this study was to test an updated version of SIENAX tool of FSL library (depending on the space resolution of the images) for improved MRI-derived assessment of WM and GM volumes in healthy controls (HC) and MS patients. **Methods** The new procedure of SIENAX was tested on two different datasets. The first dataset consisted of a sample from a multisite clinical study (17 different sets of images), including 56 relapsing remitting (RR) MS patients with T1-W, T2-W and PD images (0.97x0.97x3 mm^3) scanned twice over a month. The second dataset included 35 healthy controls (HC) and 19 RR patients scanned twice at 3T scan in Milan within 15 days and 1 month respectively with FFE (0.8x0.8x1 mm^3) and conventional turbo spin-echo (1mm in-plane resolution, 3mm thickness). The patients of both datasets were not clinically active and did not show new or enlarging lesions over the study period On these two datasets, the new version of SIENAX (SIENAX2.0) was tested against the original SIENAX and the SPM8 unified segmentation. On SIENAX2.0, WM and GM maps were obtained automatically for each T1-W image by using FSL tools as follows (see *Table 1*): i) a 2-step procedure for brain extraction: a) creation of a rough brain mask obtained by nonlinearly registering the standard space mask onto the native brain image; and b) inclusion of mislabeled brain tissue voxels to refine the brain masks (for 2D image BET with -S option was used); ii) new inhomogeneity correction performed with fslanat; iii) lesion filling with intensity similar to the surrounding WM; iv) segmentation of the filled brain image by using FAST; v) relabeling of potentially mislabeled GM clusters (performed only on 3D-image); vi) inclusion of the segmented deep GM structures obtained by using FIRST. For each scan pair, measurement errors were derived from percentage difference of GM and WM volumes. Nonparametric comparisons of the absolute error within a region of interest (GM or WM) between different algorithms (SPM8, SIENAx and SIENAX2.0) were done with a paired-sample Wilcoxon signed rank test. A Wilcoxon rank sum test was used for testing differences in the absolute error between patients and controls ($p<0.05$, corrected for Bonferroni).

Pipeline	3D-3T MRI	2D-1.5T MRI
Brain Extraction	FNIRT-based	BET
Inhomogeneity Correction	Fslanat	Fslanat
Filling	FSL	FSL
Segmentation	FAST	FAST
Relabeling	Performed	No Performed
Deep GM segmentation	FIRST	FIRST

Table 1. Steps of SIENAX2.0 for high quality data (3D-3T MRI) and low quality data (2D-1.5 T MRI). Pipeline is similar except for the brain extraction and for the "Relabeling" step.

Results: *2D Patients Dataset.* Five patients were excluded from the analysis for poor data quality (3 for movement artifacts, 2 for problems in nonlinear registration when SPM8 was used). Absolute mean percentage error in GM was 1.88% (±1.4), 1.57% (±1.3) and 0.96% (±0.85) for SPM8, SIENAX and SIENAX2.0 respectively, with significant differences ($p<0.01$, Bonferroni-corrected) between SIENAX2.0 and both SPM8 and SIENAX and no significant differences between SPM8 and SIENAX. Absolute mean percentage error in WM was 2.04% (±1.53), 1.42% (±1.3) and 0.92% (±0.71) for SPM8, SIENAX and SIENAX2.0 respectively, with significant differences ($p<0.05$, Bonferroni-corrected) between SIENAX2.0 and both SPM8 and SIENAX and no significant differences between SPM8 and SIENAX.

3D Healthy Controls Dataset. One patient was excluded from the analysis for poor data quality (movement artifacts). Absolute mean percentage error in GM was 1.12% (±1.4), 1.35% (±0.8) and 1.13% (±1) for SPM8, SIENAX and SIENAX2.0 respectively, with no significant differences ($p>0.5$) between the three methods. Absolute mean percentage error in WM was 1.44% (±1.1), 1.38% (±0.8) and 1.10% (±1.4) for SPM8, SIENAX and SIENAX2.0 respectively, with no significant differences ($p>0.5$) between the three methods.

3D Patients Dataset. Two patients were excluded from the analysis for poor data quality (movement artifacts). Absolute mean percentage error in GM was 2.16% (±1.2), 1.77% (±1.3) and 1.28% (±1.3) for SPM8, SIENAX and SIENAX2.0 respectively, with significant differences ($p<0.05$, Bonferroni corrected) between SPM8 and SIENAX2.0 and no differences between SPM8 and SIENAX and between SIENAX and SIENAX2.0. Absolute mean percentage error in WM was 1.67% (±1.5), 1.97% (±1.2) and 0.89% (±0.5) for SPM8, SIENAX and the SIENAX2.0 respectively, with significant differences ($p<0.05$, Bonferroni- corrected) between SIENAX and SIENAX2.0 and no differences between SPM8 and SIENAX2.0 ($p=0.5$) and between SPM8 and SIENAX ($p=0.1$).

3D Dataset: differences in error in HC and Patients. When mean percentage errors of HC and patients were compared within each method, significant differences in absolute mean percentage error ($p<0.01$, Bonferroni-corrected) were found between the error of GM in HC and patients when SPM8 was used. The other measures showed consistent errors across methods and tissue-types ($p>0.3$)

Discussion: In isotropic high-quality dataset the new SIENAX2.0 allows robust GM and WM volume assessment in both HC and MS patients, halving the error when compared with the original version of SIENAX and reducing it, particularly in MS patients, when compared with SPM8. In non-isotropic dataset, SIENAX2.0 significantly improves the robustness of GM and WM volume assessment compared with traditional SIENAX and SPM8.

References: 1. Chard DT, Jackson JS, Miller DH, Wheeler-Kingshott CA. Reducing the impact of white matter lesions on automated measures of brain gray and white matter volumes. J Magn Reson Imaging. 2010 Jul;32(1):223-8

2. Battaglini M, Jenkinson M, De Stefano N. Evaluating and reducing the impact of white matter lesions on brain volume measurements. Hum Brain Mapp.