Reproducibility of Segmented Brain Volumes Using SPM5: Effects of Changes in Signal-to-Noise Ratio and Scanner Software

B. Shuter[1,2], B. Yeh[1], S. Graham[3], C. Au[1], and S-C. Wang[1,2]

[1]Diagnostic Imaging, National University Hospital, Singapore, Singapore, [2]Faculty of Medicine, National University of Singapore, Singapore, Singapore, [3]Faculty of Arts, National University of Singapore, Singapore, Singapore

Introduction

Accurate and reproducible measurement of *in vivo* brain tissue volume is crucial for reliable trend detection in longitudinal studies of brain atrophy. Using a voxel-based morphometry software package (SPM5; Institute of Neurology, Queen Square, London, UK) we assessed the reproducibility of gray (GM) and white matter (WM) volumes (GMV, WMV) and the effects of changes in signal-to-noise ratio (SNR) and scanner software.

Materials and Methods

61 normal volunteers aged 19 to 60 years (F: n = 37; age: mean ± SD: 29 ± 12 years; M: n = 24; 26 ± 6 years) were recruited as part of a larger neuroimaging study. F & M mean ages were not significantly different (SigDif) ($p>0.1$; *t*-test) but the variances were ($p<0.005$; F-test), reflecting a lack of men aged over 40. 3D-MPRAGE sequences were acquired on a Siemens Symphony 1.5 T MRI scanner with a standard quadrature head coil (TR/TE/TI: 1840/2.11/850ms; flip angle: 10°; matrix size: 256 x 192; voxel dimensions: x & y: 1 mm; slices: 80; thickness: 2 mm; number of averages (**NA**)/scan time: **1**/6 min or **2**/12 min). Subjects were scanned at least twice (St1/St2), each time with the same NA (1 or 2). Inter-scan periods were 135 min (0.09 days; n = 15), 1.1 days (n = 13) or 7─124 days (n = 36). 13 subjects had all scans on original scanner software [Syngo 2002B, (2B)]; 48 had all scans with upgraded software [Syngo 2004A, (4A)]. NA1 and NA2 sequences were acquired on 11 subjects (twice) using 4A software, and on 7 subjects using 2B. Six of these 7 subjects had a third scanning session using 4A software to directly assess the effects of the change in scanner software. No changes were made to the scanner hardware during the study interval.

SPM5 (default parameters) was used to output tissue probability images (TPI) for native space WM, GM and cerebrospinal fluid compartments. All other image processing was performed using in-house software (MATLAB v7.1; MathWorks, Natick, MA). Pixels with $p\geq0.5$ were counted in GM and WM TPI and their sum multiplied by the volume of a native voxel to give GMV and WMV. Peaks in number-intensity histograms for the GMV and WMV, and for regions of interest placed in noise-only regions, were used to measure SNR as ratios of peak intensities. The volume of a head-sized spherical phantom was determined from images acquired with the same MPRAGE sequence (using NA2) at regular intervals during the study. Microsoft Excel was used for all statistical analysis (means: *t*-tests, variances: F-tests, trends: regressions).

Results and Discussion

Differences in mean GMV and WMV between scans using the same software versions and NA (St2 - St1) were less than ±1% and not SigDif (see Table 1). Volume changes for St2 vs St1 as a function of inter-study interval (0.09, 1.1, 7─124 days) were also not SigDif by regression analysis. However, substantial (>~3%) and statistically SigDifs were found in GMV for NA2 vs NA1 and 4A vs 2B comparisons, indicating a significant increase in GMV for greater SNR [SNR(NA2) = 1.41 x SNR(NA1); confirmed by measurement] and for 4A software. A small SigDif depression of WMV was also found for NA2 vs NA1 using 4A software. 4A images were found to have 6 times greater pixel intensity than 2B images, and a SNR increase of about 15% for GM and WM. Reducing the dynamic range of the 4A studies by a factor of 6 and reprocessing with SPM5 produced very little change with all parameter combinations, indicating that the differences were not due solely to changes in intensity scaling. Phantom volumes were within a tight range of 0.5% of the mean result throughout the entire study period of 2 years, suggesting that the hardware was stable.

Tissue, Subject Number & Other Conditions	Tissue Volume Change Ratio	Mean Volume Change (%) ± SD	p (paired *t*-test)
WM; n = 61; NA1,2; 2B,4A	(St2 - St1)/St1	-0.3 ± 1.5	>0.05 (ns)
WM; n = 11 (x2); 4A	(NA2 - NA1)/NA1	-1.5 ± 1.7	<0.02
GM; n = 61; NA1,2; 2B,4A	(St2 - St1)/St1	0.3 ± 1.6	>0.1 (ns)
GM; n = 11 (x2); 4A	(NA2 - NA1)/NA1	2.9 ± 1.4	<<0.001
GM; n = 7; 2B	(NA2 - NA1)/NA1	4.2 ± 2.5	<0.001
GM; n = 6; NA1	(4A - 2B)/2B	3.3 ± 0.7	<0.002
GM; n = 6; NA2	(4A - 2B)/2B	3.3 ± 1.0	<0.01

Table 1: GMV & WMV changes by scan (St1/2), SNR (NA1/2) and software version

M & F absolute volumes in subjects aged <40 years were SigDif (GMV(M - F): +7%, $p<0.01$; WMV(M - F): +12%, $p<<0.001$). GMV differences of >5% were found in 2 subjects. One (10% GMV(St2 - St1) difference) had poorer resolution on the second study, most probably as a result of slight movement. This result was not included in the above as a repeat scan showed improved image quality and good agreement. The other subject had the largest brain capacity and GM was clearly removed by SPM5 in both St1 and St2. Two subjects with ~5% WMV difference are being investigated.

Conclusions

Significant differences in segmented brain volumes can result from interactions between volume software, user determined scanning conditions (SNR changes in our case) and manufacturer-set scanner software changes, which in this work produced changes in SNR, voxel intensity and possibly other unknown alterations in scanning parameters. Such differences may mask subtle volume changes during the course of longitudinal studies, implying that uncompromising and constant vigilance is required in scanning protocols and in the handling of data from control subjects and patients throughout the study period. With good technique, volume changes of ~1% can be reliably assessed in groups of as few as 15 subjects using SPM5.