Intra-Observer, Inter-Observer and Inter-Scanner Variations in Brain MRI Volume Measurements in Multiple Sclerosis

Marco Rovaris\(^1\), Claudio Gasperini\(^2\), Maria Pia Sormani\(^3\), Stefano Bastianello\(^4\), Carlo Pozzilli\(^4\), Giancarlo Comi\(^5\), Massimo Filippi\(^1\)

\(^1\)Neuroimaging Research Unit, Dept. of Neuroscience, IRCCS H San Raffaele, Via Olgettina, 60, Milan, Italy; \(^2\)Dept. of Neurology, Ospedale San Camillo-Forlanini, Rome, Italy; \(^3\)Neuroimaging Research Unit, Dept. of Neuroscience, IRCCS H San Raffaele, Milan, Italy; \(^4\)Dept. of Neuroscience, University La Sapienza, Rome, Italy; \(^5\)Clinical Trials Unit, Dept. of Neuroscience, IRCCS H San Raffaele, Milan, Italy.

Introduction

Several cross-sectional and longitudinal studies with magnetic resonance imaging (MRI) showed that brain atrophy can develop in the early phases of multiple sclerosis (MS) [1, 2] and that the amount of tissue loss is more pronounced in patients with more severe disability [3-5]. As a consequence, measurement of brain volume is regarded as an objective marker of MS severity with the potential to monitor treatment efficacy in MS accurately [6]. In this scenario, it is necessary to develop and use MR techniques with low intrinsic variability in order to increase the chances of detecting relatively small treatment effects. Several studies have shown that intra-observer variability in measuring brain volume is relatively low [1-8]. However, since scanner performance and accuracy in repositioning are among the major contributors to MRI measurement variability [9, 10], the validation of brain volume measurement should include evaluation of scan-rescan variability on the same and, when possible, on different scanners. In this study, we assessed the effect of scanner-related variation on brain MRI volume measurements in MS.

Methods

Nine MS patients were included in the study. All had a relapsing-remitting disease course; their mean age was 29 years, median duration of the disease was 6 years and median Expanded Disability Status Scale (EDSS) score [11] was 1.5. During two sessions separated by a 24-hour interval, patients were imaged using two MR systems operating at 1.5 T (Siemens Vision -scanner A- and Philips Gyroscan NT - scanner B). Fast fluid-attenuated inversion recovery (fast-FLAIR) images were obtained from all the patients. All images consisted of 44 contiguous, interleaved, 3-mm-thick axial sections with a rectangular field of view of 250x188 mm. The following acquisition parameters were used: scanner A; TR/TE/TI=9999/105/2200, echo train length=7; raw-data matrix = 256x182, in-plane resolution = 0.98x0.99 mm; number of acquisitions=2. The two fast-FLAIR sequences resulted in similar contrast between brain tissue and surrounding cerebro-spinal fluid (CSF) at the two centers.

Whole brain volumes were measured using a seed growing technique for brain parenchyma segmentation [8]. At the end of the segmentation process, the tissue volume was calculated by multiplying the number of pixels included in the regions of interest (ROIs) for the voxel size. To evaluate the intra-observer, inter-observer and inter-scanner coefficients of variation (COVs), the MR brain volume computed from each image was measured three times by three blinded observers in a random order. An interval of one month separated each of the measurement sessions. Inter-scanner variability was defined as the variability between mean estimates of brain volume determined by the same single observers when images of one patient obtained using the two different MR scanners were evaluated. Thus, it includes not only the intra-observer variation but also the repositioning variability and the variation in observed brain volume caused by different scanner hardware and sequence implementations. The variability was expressed as coefficients of variation (COVs= square root of the variance/mean). The components of variance (intra- and inter-observer, and inter-scanners) were estimated using a random effect ANOVA model. The standard errors (SE) for the components of variance were estimated using the bootstrap resampling technique.

Results

The overall means and SE of the lesion volumes obtained for the entire sample and those for each observer and scanner separately are presented in Table 1. There was no statistically significant difference in brain volume between observers or between scanners. The mean intra-observer COV was 1.2% (SE=0.8%) and the mean inter-observer COV was 2.4% (SE=1.2%). The COVs for each observer and scanner are reported in Table 2. The intra-observer variance was not statistically different from those found between observers or scanners.

Discussion

In this study, we measured the volume of the whole of the brain tissue in a group of patients with MS, using fast-FLAIR images with high spatial resolution. When determining brain volume, it is important that the cerebral tissue is well-delineated from the surrounding CSF. The fast-FLAIR sequences used in this study are designed to null the signal from CSF. We also obtained 3-mm thick slices in order to reduce partial volume effect at the brain's edge and, as a consequence, improve further brain/CSF delineation. The segmentation technique used is based on local thresholding and, therefore, took advantage of sharp delineation of brain/CSF edges, requires only a modest amount of human intervention and is characterized by a high intra-observer reproducibility when measuring the volume of different brain portions [8]. This study was designed to assess the variability introduced by the use of different scanners on brain volume measurements, in a situation that resembles closely that of MS clinical trials. The difference in acquisition parameters between the two fast-FLAIR sequences is also within the ranges usually accepted in multicentre MR trials to accomplish with different MR scanners’ constraints. The main result of the present study was to show that the use of different raters and MR scanners has only a small influence on the measured brain volume and does not affect the reproducibility of this measurement greatly. This suggests that using MRI to measure brain volume as part of longitudinal studies of MS should be relatively robust, and not be affected greatly by the inevitable use of different scanners and sequences.

References

10. Filippi1, Intra-Observer, Inter-Observer and Inter-Scanner Variations in Brain MRI Volume Measurements in Multiple Sclerosis.

Table 1. Mean (SE) brain volumes (ml) obtained for different observers and scanners

<table>
<thead>
<tr>
<th>Observer</th>
<th>All observers</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1204 (32)</td>
</tr>
<tr>
<td>B</td>
<td>1181 (30)</td>
</tr>
</tbody>
</table>

Table 2. Mean (SE) COVs (%) for brain volume measurements obtained for different observers and scanners

<table>
<thead>
<tr>
<th>Observer</th>
<th>All observers</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1.5 (0.5)</td>
</tr>
<tr>
<td>B</td>
<td>0.6 (0.1)</td>
</tr>
</tbody>
</table>