

Reproducibility of Magnetization Transfer Ratio Histogram–Derived Measures of the Brain in Healthy Volunteers

Maria P. Sormani, Giuseppe Iannucci, Maria A. Rocca, Giovanna Mastronardo, Mara Cercignani, Luca Minicucci, and Massimo Filippi

Summary: Using two MR scanners, we evaluated the intraobserver, interobserver, image-reimage, and interimager variabilities in the assessment of magnetization transfer ratio (MTR) histograms obtained monthly on four occasions from five healthy volunteers. With multiple observers, the mean coefficients of variations ranged from 2.2% to 8.2% for “pure” image-reimage variability, from 1.2% to 4.9% for interobserver variability, and from 2.1% to 4.9% for image-reimage variability. The mean intraobserver coefficients of variations were always lower than 1%. The mean coefficients of variations ranged from 10.2% to 14.6% for pure interimager variability and from 8.6% to 14.3% for interimager variability with multiple observers. Interimager variability accounted for 96.0% of the overall variability of average MTR, for 96.7% of peak location, and for 41.1% of the peak height. The use of different MR scanners is the main source of variability when obtaining MTR histograms.

For cases of multiple sclerosis (MS), conventional MR imaging is a very sensitive tool that can show the formation of new lesions and their subsequent evolution. Nevertheless, conventional MR imaging is not without limitations; perhaps the most important of which is its lack of specificity to the many pathologic substrates of individual MS lesions, which range from edema and inflammation to severe demyelination and axonal loss (1). In addition, conventional MR imaging may not reveal the full extent of the disease activity (1).

Magnetization transfer ratio (MTR) histograms may, at least partially, meet some of these demands

(2). A recent preliminary postmortem report found a correlation between MTR and the percentage of residual axons and severity of demyelination in MS lesions (3). Animal studies have also shown that low MTR correlates with histopathologic findings of myelin loss and axon destruction (4–6). Whole-brain magnetization transfer (MT) histogram–derived measures provide a complete assessment of both macro- and microscopic disease burden in cases of MS and are correlated with the clinical manifestations of the disease (7, 8).

It has previously been shown that for other MR-derived measures of MS lesion burden (9), the use of different imagers and image evaluators introduces undesirable measurement error. In this study, we evaluated the intraobserver, interobserver, image-reimage, and interimager variabilities in the assessment of MTR histograms obtained serially from five healthy volunteers examined by two imagers.

Technique

Methods

Five healthy volunteers (three women and two men; age range, 27–40 years; mean age, 35 years) entered the study after providing informed consent. Approval from the local ethical committee was also obtained before study initiation. None of the participants had a previous history of neurologic diseases, and all were normal at the time of neurologic examination.

MR images of the brain were obtained, using a 1.5-T machine (Magnetom SP63, Siemens; this machine will be referred as imager A), from each of the participants every 28 (± 5) days on four separate occasions (ie, the follow-up was of 3 months' duration). For imager A, the maximum available gradient strength was 10 mTm^{-1} , with a maximum slew rate of $10 \text{ Tm}^{-1}\text{s}^{-1}$. A birdcage head coil with a diameter of approximately 300 mm was used for both RF transmission and for signal reception. On each imaging occasion, we obtained the following images: 1) dual-echo conventional spin-echo images (2400/30–80/1 [TR/TE/excitations]; 24 contiguous, interleaved, 5-mm-thick, axial sections; matrix, 256×256 ; field of view, 250 mm), which were normal throughout the entire follow-up period for all of the participants; and 2) 2D gradient-echo images (600/12/2; $\alpha = 20^\circ$; 20 contiguous, interleaved, 5-mm-thick, axial sections; matrix, 256×256 ; field of view, 250 mm), with and without a saturation pulse. The saturation pulse was an off-resonance RF pulse centered 1.5 kHz below the water frequency, with a gaussian envelope of a duration of 16.4 milliseconds, a bandwidth of 250 Hz, and a flip angle of 850° . After the final MR session, we obtained a pair of addi-

Received March 9, 1999; accepted after revision July 12.

From the Neuroimaging Research Unit (M.P.S., G.I., M.A.R., G.M., M.C., L.M., M.F.), Department of Neuroscience, Scientific Institute Ospedale San Raffaele, University of Milan, Milan, and the Unit of Clinical Epidemiology and Trials (M.P.S.), National Institute for Cancer Research, Genoa, Italy.

Dr. M.P. Sormani is supported by a grant from TEVA Italy. Dr. G. Iannucci is supported by a grant from the Neurology School, University of Chieti, Chieti, Italy.

Address reprint requests to Dr. Massimo Filippi, Neuroimaging Research Unit, Department of Neuroscience, Scientific Institute Ospedale San Raffaele, Via Olgettina, 60, 20132 Milan, Italy.

tional gradient-echo images (with and without the saturation pulse) of each participant, using a different 1.5-T machine (Vision, Siemens; this machine will be referred to as imager B). For imager B, the maximum available gradient strength was 21 mTm^{-1} , with a maximum slew rate of $167 \text{ Tm}^{-1}\text{s}^{-1}$. The head coil used was identical to that used for imager A. The gradient-echo sequence on this second imager was the same as on the first, except that the saturation pulse was slightly different. Its offset frequency and envelope shape were the same, but the duration was 7.68 milliseconds and the flip angle was 500° .

For follow-up images, the image planes were carefully repositioned according to published guidelines (10). Both imagers were on a course of regular maintenance throughout the study.

The gradient-echo images were transferred to a workstation (Sun Sparkstation; Sun Microsystems, Mountain View, CA) for postprocessing. From the two gradient-echo images (with and without the saturation pulse), quantitative MTR images were derived pixel-by-pixel using an in-house developed software according to the following equation: $\text{TR} = (M_0 - M_S) / M_0 \times 100\%$, in which M_0 is the signal intensity for a given pixel without the saturation pulse and M_S is the signal intensity for the same pixel when the saturation pulse is applied. Signal intensities in the calculated images represent the MTR values.

From the MTR images, we obtained two sets of MTR histograms. The first was a set of histograms from the whole imaged tissue without any human intervention (the air around the scalp was removed by means of fully automated homemade software). This allowed us to assess the pure image-reimage and interimager variabilities. The second set consisted of brain MTR histograms obtained by three observers, who were blinded to the image details. They first segmented the brain from the surrounding tissue by using a semiautomated segmentation technique based on local thresholding, and then followed the postprocessing method described in detail by Rovaris et al (8) to obtain the brain MTR histogram. One of the observers repeated the whole evaluation procedure after an interval of 1 month, and for this second evaluation, the observer was blinded to the results obtained previously. This allowed us to assess the intraobserver and interobserver variabilities and the image-reimage and the interimager variabilities with multiple observers. The intra- and interobserver variabilities in the number of segmented pixels were also calculated.

To correct for the between-patient differences in brain volume, each histogram was normalized by dividing it by the total number of pixels included. For each histogram, the following measures were derived: the relative peak height (proportion of pixels at the most common MTR value); peak position (most common MTR); and mean brain MTR. For each of these, the variability introduced by each of the factors considered was calculated.

Intraobserver variability was defined as the variability between brain MTR histogram-derived measures obtained by the observer who evaluated the same images on two separate occasions. Interobserver variability was defined as the variability between brain MTR histogram-derived measures obtained by the three observers who evaluated all of the images (ie, the images of each participant at each time point). Pure image-reimage variability was defined as the variability between MTR histogram-derived measures obtained from the whole imaged tissue (ie, without any human intervention) for each participant at each time point. Image-reimage variability with multiple observers was defined as the variability between brain MTR histogram-derived measures obtained by each observer for each participant at each time point. Pure interimager variability was defined as the variability between MTR histogram-derived measures obtained from the whole imaged tissue when images from each patient were compared between the two MR imagers. Interimager variability with multiple observers was defined as the variability between brain MTR histogram-derived

TABLE 1: Mean image-reimage and inter-imager COV for the MTR histogram-derived measures obtained without human intervention

	Average		Peak
	MTR	Peak height	Position
Mean image-reimage COV (SE)%	2.2 (0.3)	8.2 (2.0)	3.0 (0.3)
Mean inter-imager COV (SE)%	12.2 (1.2)	14.6 (4.2)	10.2 (1.1)

TABLE 2: Means and ranges of the MTR histogram-derived parameters obtained with and without human intervention from images of the two imagers

	Average	Peak	Peak
	MTR (%)	height	Position
Mean (ranges) obtained without human intervention/imager A	49.0	48.9	44
Mean (ranges) obtained without human intervention/imager B	47.0–51.3	35.4–56.8	40–45
Mean (ranges) obtained with human intervention/imager A	38.3	58.0	36
Mean (ranges) obtained with human intervention/imager B	37.2–39.5	54.2–60.5	35–37
Mean (ranges) obtained without human intervention/imager A	46.0	63.7	44
Mean (ranges) obtained with human intervention/imager A	44.3–48.0	55.2–76.7	41–45
Mean (ranges) obtained without human intervention/imager B	39.6	71.7	36
Mean (ranges) obtained with human intervention/imager B	38.4–41.0	66.3–78.8	35–37

derived measures obtained by the three observers when images from each patient were compared between the two MR imagers. The contribution of each of the components of variance to the overall variance was calculated using a random-effects model. In this model, only the main effects were included. The interaction terms were also calculated but not included in the model because they were always small and not significant. Thus, the final model included contributions from all of the tested factors plus a residual variance, which represents the variance not explained by the model. Then, the variance estimate for each factor was obtained by adding the variance of each individual factor to the residual variance. Coefficients of variation were calculated to assess these variabilities. The coefficient of variations is defined as the SD of a random variable divided by its mean value. The components of variance were estimated using a random-effects model, and the SD related to each of the sources of variability was calculated by taking the square root of the corresponding component of variance. The standard errors (SE) of the coefficients of variations were estimated using the bootstrap resampling technique.

MTR Histogram Variabilities from the Whole Imaged Tissue

In Table 1, the pure image-reimage and interimager variabilities are reported for the average histogram MTR, peak height, and peak position. The variability introduced by the use of multiple imagers was much higher than that due to repeat imaging on multiple occasions (image-reimage) for all three measures. The mean and ranges of average MTR, histogram peak height, and location obtained from the whole imaged tissue without human intervention are reported in Table 2.

MTR Histogram Variabilities from the Brain

The intraobserver coefficients of variations for the number of segmented pixels and for all of the MTR histogram-derived measures were less than 1%. The mean interobserver coefficient of variations for the number of the segmented pixels was 3.3% (SE, 0.7%), the mean image-reimage coefficient of variations was 2.1% (SE, 0.2%), and the mean interimager coefficients of variations was 5.5% (SE, 0.9%). This had an impact on the interobserver, image-reimage, and interimager variability of all three MTR histogram-derived measures, as reported

TABLE 3: Mean inter-observer, image-reimage and inter-imager COV for the MTR histogram-derived measures obtained after brain segmentation by observers

	Average		Peak
	MTR	Peak Height	Position
Mean inter-observer COV (SE)%	1.2 (0.3)	4.9 (1.1)	1.9 (0.6)
Mean image-reimage COV (SE)%	2.1 (0.3)	4.9 (1.1)	2.5 (0.6)
Mean inter-imager COV (SE)%	11.7 (0.4)	8.6 (1.8)	14.3 (0.7)

TABLE 4: Absolute variances and relative contributions (in brackets) to the overall variance from each source of variance

Sources of Variance	Average MTR	Peak Height	Peak Position
Subject	0.19 (0.7%)	20.66 (38.3%)	0.11 (0.3%)
Observer	0.04 (0.1%)	0.53 (1.0%)	0.00 (0.0%)
Image-reimage	0.61 (2.1%)	0.92 (1.7%)	0.53 (1.4%)
Imager	27.0 (96.0%)	22.17 (41.1%)	36.40 (96.7%)
Residual	0.25 (0.8%)	9.68 (17.9%)	0.62 (1.6%)
Total	28.13	53.96	37.66

in Table 3. Again, the variability introduced by the use of multiple imagers was much higher for all three measures than were the variabilities owing to multiple observers and repeated imaging. The mean and ranges of average MTR, histogram peak height, and location of the brain obtained using the two different imagers are reported in Table 2.

Relative Contributions of the Different Components of Variance to Overall Variance

Table 4 reports (as a percentage) the relative contributions to the overall variance from the various components, such as variance owing to participant variability, multiple observers, and multiple imagers. The ratio of each individual variance to the overall variance gives the relative contributions of each factor. The interparticipant variability accounted for only 0.7% of the total variance of the average histogram MTR and 0.3% of the peak position variance, whereas it accounted for the 38.3% of the total variance of the peak height.

The main source of variability for all of the MTR histogram-derived measures was the interimager variability. It accounted for 96.0% of the total variance for average histogram MTR, for 96.7% of the total variance of the peak location, and for 41.1% of the total variance of the peak height. The relative contribution of interimager variability to the total variance of the peak height was smaller than for the other MTR histogram-derived measures because of the large between-participant variability of this parameter. For all of the MTR histogram-derived measures, the interobserver variability accounted for small percentages of the total variance (0.1% for the average MTR, 1.0% for the peak height, and 0.0% for the peak location).

Discussion

As is already known for conventional MR measures (9), MTR histogram-derived measures may be influenced by the use of different imagers, sequences, and observers, as well as by the day-to-day variability of both the imager and the image evaluator. We studied healthy volunteers to assess the variability without the confounding factor of

MS biological variation. Because the MT histograms from patients with MS have broadly the same characteristics as control subjects (7, 8) but with, in general, a lower mean MTR, peak height, and peak position, we think that our findings are transferable to patient studies.

This study indicates that variability in MTR histograms coming from the use of the different MR imagers, with their different MT pulse characteristics, is much higher than the intraobserver, interobserver, and image-reimage variabilities. Because the machines had the same field strength and general pulse sequence parameters, it is likely that it is the different off-resonance MT pulse properties that account for the differences in the MT histograms. The higher flip angle of the pulse of imager A, which resulted in a higher average MTR and peak position than imager B, supports this fact. Nevertheless, this kind of variation between different manufacturers and imager types is typical of that found when conducting multicenter studies. The inclusion of imagers with different field strengths would add another source of variability, because the tissue relaxation properties vary with field strength, altering the MT behavior (11).

Intra- and interobserver variabilities were very low for all of the MTR-derived measures. This is unlikely to change in real-life MS studies, because the postprocessing steps used to create MTR histograms do not require significant amounts of subjective human intervention and, as a consequence, operator training can be less intensive than is required for measuring T2 lesion load (12). Also, image-reimage variability did not contribute significantly to the overall variability. This suggests that standardized repositioning procedures and regular maintenance of the imagers will ensure relatively good stability, at least for short-term follow-up studies. For longer-term studies (eg, 2–3 years), the collection of MTR data from normal control subjects throughout the study, for the purposes of normalization, would be prudent.

In our study, the MTR histogram peak height has the largest intra- and interparticipant variability. This is likely because of participant motion during the acquisition of the two gradient-echo images, with and without the saturation pulse. The effect of motion is likely to be even higher for patients with MS because of their disabilities and because MT imaging is acquired in the context of imaging sessions that are usually longer than those of the present study. Nevertheless, it is conceivable that the effect of participant motion on the MTR histogram peak height might be successfully reduced by the coregistration of the two gradient-echo images.

References

1. Filippi M. The role of non-conventional magnetic resonance techniques in monitoring evolution of multiple sclerosis. *J Neurol Neurosurg Psychiatry* 1998;64[suppl]:S52–S58
2. van Buchem MA, McGowan JC, Kolson DL, Kolson DL, Polansky M, Grossman RI. Quantitative volumetric magnetization

- transfer analysis in multiple sclerosis: estimation of macroscopic and microscopic disease burden.** *Magn Reson Med* 1996; 36:632–636
3. van Waesberghe JHTM, van Walderveen MAA, de Groot C, et al. **Postmortem correlation between axonal loss, MTR, and hypointensity on T1 SE in MS.** *Proc Int Soc Magn Reson Med* 1998;2:1334
 4. Lexa FJ, Grossman RI, Rosenquist AC. **MR of Wallerian degeneration in the feline visual system: characterization by magnetization transfer rate with histopathologic correlation.** *AJNR Am J Neuroradiol* 1994;15:201–212
 5. Dousset V, Brochet B, Vital A, et al. **Lysolecithin-induced demyelination in primates: preliminary in vivo study with MR and magnetization transfer.** *AJNR Am J Neuroradiol* 1995;16: 225–231
 6. Kimura H, Meaney DF, McGowan JC, et al. **Magnetization transfer imaging of diffuse axonal injury following experimental brain injury in the pig: characterization by magnetization transfer ratio with histopathologic correlation.** *J Comp Assist Tomogr* 1996;20:540–546
 7. Filippi M, Iannucci G, Tortorella C, et al. **Comparison of MS clinical phenotypes using conventional and magnetization transfer MRI.** *Neurology* 1999;52:588–594
 8. Rovaris M, Filippi M, Falautano M, et al. **Relation between MR abnormalities and patterns of cognitive impairment in multiple sclerosis.** *Neurology* 1998;50:1601–1608
 9. Filippi M, vanWaesberghe JH, Horsfield MA, et al. **Interscanner variation in brain MRI lesion load measurements in MS: implications for clinical trials.** *Neurology* 1997;49:371–377
 10. Miller DH, Barkhof F, Berry I, Kappos L, Scotti G, Thompson AJ. **Magnetic resonance imaging in monitoring the treatment of multiple sclerosis: concerted action guidelines.** *J Neurol Neurosurg Psychiatry* 1991;54:683–688
 11. Bottomley PA, Foster TH, Argersinger RE, Pfeifer LM. **A review of normal tissue hydrogen NMR relaxation times and relaxation mechanisms from 1–100 MHz: dependence of tissue type: NMR frequency, temperature, species, excision, and age.** *Med Phys* 1984;11:425–448
 12. Rovaris M, Rocca MA, Sormani MP, Comi G, Filippi M. **Reproducibility of brain MRI lesion volume measurements in multiple sclerosis using a local thresholding technique: the effects of formal operator training.** *Eur Neurol* 1999;41:226–230