MR Imaging Quantitation of Gray Matter Involvement in Multiple Sclerosis and Its Correlation with Disability Measures and Neurocognitive Testing

Isabelle Catalaa, Jennifer C. Fulton, Xiang Zhang, Jayarm K. Udupa, Dennis Kolson, Murry Grossman, Lougang Wei, Joseph C. McGowan, Marcia Polansky, and Robert I. Grossman

BACKGROUND AND PURPOSE: Multiple sclerosis (MS) is the most common inflammatory disease of the central nervous system and manifests both physical and neurocognitive disabilities. Although predominantly a disease of the white matter, MS is also characterized by lesions in the gray matter. Previous pathologic studies have found that cortical and deep gray matter lesions comprised 5% and 4%, respectively, of total lesions. Using software for lesion detection and quantitation, our study was designed to determine MS involvement in the cortical and deep gray matter and to correlate gray matter lesion load with neurocognitive function and the Kurtzke Expanded Disability Status Scale.

METHODS: Using a semiautomated segmentation algorithm that detected and delineated all possible brain MS lesions on MR images, we investigated gray matter lesion volume in 18 patients with untreated relapsing-remitting MS. Cortical and deep gray matter lesions then were correlated with the neurocognitive and physical disability measurements.

RESULTS: We found that cortical gray matter lesions comprised approximately 5.7% of the total lesion volume, whereas deep gray matter lesions comprised another 4.6% in this patient cohort. No strong correlations were found between gray matter lesions and disability status or neurocognitive function.

CONCLUSION: These results are similar to those found in previous pathologic studies. The cortical lesion load in cases of relapsing-remitting MS, as measured by MR imaging, represents less than 6% of the total lesion volume and does not correlate with disability measures or neurocognitive tests.
al (19) found significant correlations between total lesion volume (TLV) and nonverbal memory, information processing speed, and attention.

We developed and validated a technique for lesion detection based on the concept of fuzzy connectedness. When compared with previous methods, this technique seems to be more robust (20). We have used this method to quantitate the deep and cortical gray matter lesion volumes (CGMLVs) over the whole brain and to determine whether cortical gray matter lesion load correlated with neuropsychologic tests.

Methods

A cohort of 18 patients (16 women and two men), ranging in age from 22 to 50 years, was identified for this project. All participants met the clinical criteria of having relapsing-remitting MS as defined by Poser et al, and none had been treated with β-interferon medications. The patients’ disease duration varied from 3 months to 10 years (6 years ± 1.5 year). Each patient was followed up regularly at 6-month intervals for a minimum of 2 years with neurologic, cognitive, and MR imaging examinations.

Neurologic Assessment

The degree of physical disability was determined by a Neurologist (D.K.) specializing in MS and using the EDSS and the Ambulation Index (AI).

Neurocognitive Assessment

A battery of neurocognitive tests was performed on each patient by trained technicians supervised by a neuropsychologist. Four different categories of functions were assessed, including executive function and processing speed, verbal short- and long-term memory, and visuospatial skills. Executive function and information-processing speed includes the paced auditory serial addition test (1.6 and 2.0 seconds), the controlled oral work association test (phonetic fluency), the category fluency test (semantic fluency), the Bells test, and the symbol-digit modalities test. Each test is described below.

For the paced auditory serial addition test, the participant is asked to complete a sequence of serial additions. A sequence of consecutive numbers is presented on an audiotape and read at two speeds, a 1.6- and a 2.0-second interval. The participant is asked to provide the sum of each pair of numbers (21).

For the controlled oral word association test (phonetic fluency), the participant is asked to recite words beginning with given letters (eg, as E, A, and S) for 1 minute each (22).

For the category fluency test (semantic fluency), the participant is asked to recite the names of as many animals as possible in 1 minute (23).

The Bells test involves the identification of specific-target icons that are distributed randomly on a page with similar, distracting icons (24).

The symbol-digit modalities test involves the decoding of a series of visual symbols by naming the numbers corresponding to the symbol in a defined time limit (90 seconds) (25).

For the Rey auditory verbal learning test (long-term memory), the patients are presented orally a list of 15 words. The examination consists of five learning trials, retrieval of the target list after presenting a distracter list, memory retrieval after a half-hour delay, and verbal recognition of the words from a list also containing semantic and phonetic foils (26).

For the word-span test (short-term memory), the participants are asked to recall an increasingly longer list of one-syllable words presented at a 1-second interstimulus rate. The words consist of three different types: imaginable nouns from a similar superordinate category, semantically unrelated nouns, and semantically unrelated verbs.

For the roadmap test (visuospatial and perceptual processing skills), the participant is asked to describe the turns necessary to follow a path marked on a map (27). For each test, the number of correct responses given was scored and reported. Using these tests in a larger neuropsychological test battery, Grossman et al (28) showed that patients with MS differ from age- and education-matched control participants in their neuropsychological performance.

MR Imaging Assessment

Cranial MR imaging was performed on a 1.5-T scanner. Fast spin-echo, proton density-, and T2-weighted images were acquired as 3-mm contiguous sections of the brain. The protocol had an echo train length of 8, an echo time of 16/60 milliseconds, and a repetition time of 2500 milliseconds.

A three-dimensional software system, operating on a workstation, was used to detect, delineate, and calculate the number and volume of MS lesions. This software system, using the theory of fuzzy connectedness, identifies and delineates all potential lesions automatically (20, 29). True lesions are selected from all the possible lesions identified by the program by a yes/no response from a trained reader. This method has been tested extensively and validated for precision, accuracy, and efficiency (20). Based on 15 whole-brain studies, its inter- and intraobserver variability has been found to be approximately 1% (20). Because delineation is performed taking into account the graded constitution (fuzziness) and not based on drawing a boundary around the lesions, its delineation of the lesion region has a false-negative lesion volume of approximately 1.3% (20). The variability in the volume comes essentially from the variability in the reader’s ability to distinguish true from false lesions. Its sensitivity to detect very subtle lesions, such as those of the U-fibers, has also been proven by Miki et al (29). We think that our algorithms are more sensitive for distinguishing lesions from CSF than are other reported algorithms because of the use of the fuzzy connectedness principle (20). They may well be comparable with the fluid-attenuated inversion-recovery MR techniques in sensitivity, although we have not conducted a formal study to verify this.

Using this method, each image was read three separate times. First, the program was used to delineate all the MS lesions in the brain, giving the TLV. The second reading was completed by isolating those lesions that lie in the deep gray matter regions. The last reading was performed separately by two neuroradiologists (I.C. and X.Z.). At this second step, those lesions that were in the white matter region were removed manually by the click of a button, leaving only the lesions with gray matter involvement. This removal step was performed to avoid the overestimation of gray matter lesions, knowing that these lesions are usually small (Fig 1). Finally, deep gray matter and white matter lesions were removed, leading only to cortical gray matter lesions.

The criterion of classification for cortical gray matter lesion was the location of the lesions in a part of the cerebral cortex. The criterion of classification for deep gray matter lesions was the location of the lesions in the deep gray matter of the cerebrum (ie, caudate, putamen, pallidum, and thalamus). Lesion loads were computed separately for each parameter. Three volumetric measurements were calculated; TLV, CGMLV, and deep gray matter lesion volume (DGMVLV). The percentages of both DGMVLV (%DGMVLV) and CGMLV (%CGMLV) over the TLV were calculated. The results obtained separately from the two readers (DGMVLV, CGMLV, %DGMVLV, %CGMLV) were then compared to calculate the variability of the method.

Statistical Tests

Intraclass correlations were obtained to see the variability between the volumetric measures (DGMVLV, CGMLV,
Results

The percentage of CGMLV over TLV averaged 5.75 ± 4.8% (mean ± SD) for the first reader (Table 1), with similar results (5.6 ± 5.3%) obtained from the second reader.

The percentage of DGMLV over TLV was also similar between both readers (4.88 ± 4.3% for the first reader and 4.33 ± 3.3% for the second reader).

The DGMLV averaged 355 mm$^3$ for the first reader and 325 mm$^3$ for the second reader. The CGMLV averaged 522 mm$^3$ for the first reader and 570 mm$^3$ for the second reader.

The average %DGMLV between the two readers’ results was 4.6%. The average %CGMLV between the two readers’ results was 5.7%. The average DGMLV between the two readers’ results was 340 mm$^3$. The average CGMLV between the two readers’ results was 546 mm$^3$.

These results, based on the detection of gray matter lesions via T2-weighted imaging and 3D software, are similar to the previous pathologic results presented by Brownell et al (1).

The intrareader repeatability for the percentages of deep and cortical gray matter volumes was 85% and 89%, respectively. The intrareader repeatability for DGMLV and CGMLV was 61% and 77%, respectively.

The volume variability between the two readers’ measures may have been a direct result of the three-step reading, principally the manual removal of white matter lesions. We attribute the small size of the gray matter lesions analyzed to account for the reader variability. The deep gray matter lesions were found predominantly in the caudate nuclei and in the thalami (Fig 2). The cortical lesions were found essentially in the temporal lobes (Fig 3) and in the upper part of the frontal lobes.

We found no correlation between the physical measures of disability (EDSS, AI), and the volumetric measurements (TLV, CGMLV, DGMLV, %DGMLV, %CGMLV). Additionally, no correlation was found between volumetric measurements and neurocognitive tests, although a suggestion of an association was detected between the DGMLV (average between the results of both readers) and the roadmap test ($c = .65; P < .005$). The latter correlation was not strong enough to satisfy Bonferroni criteria because of the multiple comparisons performed in our analysis. Theses criteria would
have necessitated a $P$ value $< .001$ according to the Bonferroni correction. There still exists the possibility that weaker correlations were not detected in this study because of the relatively small sample size, with which a correlation must be at least approximately 0.67 to be reliably detected. We did not find any correlation between age or disease duration and neuroradiologic volumetric measure or clinical data.

Discussion

This is the first study, to our knowledge, to quantitate gray matter lesion volume in the entire brain, in patients with relapsing-remitting MS. Our results indicate the following: 1) the cortical gray matter lesion load as detected by MR methods was similar to that previously reported in studies based on histopathologic analysis (5.7% versus 5%) (1); 2) the DGMLV was similar to cortical gray matter lesion load (ie, approximately 4.5%); and 3) the gray matter lesion volume did not exhibit strong correlation with physical disability, as judged by the EDSS and AI scoring system, or with results of neurocognitive testing.

Quantitation of gray matter lesions on proton density- and T2-weighted images produced using the three-dimensional software gives reliable and repeatable numbers. The absence of correlation between gray matter lesion load and neurologic or neuropsychological data is interesting and has several possible explanations.

Comi et al (31) suggested that the involvement of spinal cord could be a determining factor in physical disability, which could serve as an explanation for our results. Ikuta and Zimmerman (32) found spinal cord damage in 70 autopsy cases of MS. Thus, a lack of correlation with measures of disability may, in part, be related to spinal cord disease that was not a consideration of this study (33).

Our results showed only the suggestion of possible correlation between DGMLV and neurocognitive tests. Most likely, the weak correlation between gray matter lesion load and cognitive impairment in MS represents both the heterogeneity of MS lesions and the relatively small proportion of lesion load involving gray matter. High-intensity lesions are known to have a variable histopathologic substrate ranging from demyelination and edema to neuronal loss. Furthermore, gray matter MS lesions may have distinct histopathologic features when compared with those of white matter. Disconnection of gray matter from white matter may occur as a result of either gray matter or juxtacortical lesions and may play a role in cognitive
impairment in MS patients. The junction of the cortex and white matter, where the U-fibers are located, demonstrates a predilection for MS lesions. In a recent study, Miki et al (29) showed that U-fibers may contribute to memory and executive control impairment. In this study, three of the 11 executive control and memory tests demonstrated statistically significant differences between the 45 patients with no or one U-fiber involvement compared with the eight patients with multiple U-fiber lesions. De-myelination and axonal degeneration processes can interrupt the connections between associative cortical areas and subcortical structures and have been implicated in subcortical dementia (4, 14, 16, 35). Maurelli et al (8) suggested that in patients with MS with extended periventricular demyelination, memory impairment could be explained by the disconnection between prefrontal and limbic structures.

Cognitive impairment is not only a function of lesion load but also a function of the lesion location. A smaller lesion, strategically placed, can cause significant cognitive damage (eg, bilateral lesions in both medial temporal lobes can cause severe memory disturbance) (18, 36). In a study presented by Arnett et al (37), patients with MS with frontal white matter lesions were more impaired on a conceptual reasoning task (Wisconsin card sorting test). Swirsky-Sacchetti et al (16) found distinct patterns of cognitive impairment, related to different lesion locations. Frontal lobe involvement impaired abstract problem solving, memory, and word fluency. Left parieto-occipital lesion involvement best predicted deficits in verbal learning and complex visual integrative skills.

Gray matter lesions have been underestimated by visual inspection of conventional MR pulse sequences because of a lack of conspicuity between lesions (high intensity) and CSF. Our image processing technique enables identification of each individual lesion and thus detects cortical and deep gray matter lesions. This type of research software is not readily available for detection of gray matter lesions on conventional T2-weighted images. Fluid-attenuated inversion-recovery images are more sensitive than conventional T2-weighted images to gray matter lesions because of increased conspicuity between brain (white) and CSF (black) (38–42). Boggild et al (43), in an MR imaging study, identified 45% more high-signal lesions with fluid-attenuated inversion-recovery imaging than with T2- or proton density-weighted imaging and 60% more lesions within the cortex and the gray-white interface; 8% of these lesions were cortical. Because a direct comparison of lesion detection by fluid-attenuated inversion-recovery imaging versus our computer-assisted T2-weighted method was not intended nor performed, we do not know their relative sensitivity.

Our study found that T2-weighted lesion load in cortical gray matter was approximately equivalent to what has been estimated pathologically to be approximately 6%. Our results indicate that the gray matter lesion load is not a strong predictor of cognitive impairment in MS. This suggests that gray matter lesion load measurements alone may not be a sensitive indicator of functional neuronal impairment. Other concepts, such as quantitation of whole-brain lesions or correlation with lesion location combined with more specific methods (magnetization transfer, spectroscopy) in assessing MS lesions, are indicated to understand further the effect of MS lesions on neurons and neuronal function.

References