Prediction of Neuropsychological Impairment in Multiple Sclerosis

Comparison of Conventional Magnetic Resonance Imaging Measures of Atrophy and Lesion Burden

Ralph H. B. Benedict, PhD; Bianca Weinstock-Guttman, MD; Inna Fishman, MA; Jitendra Sharma, MD; Christopher W. Tjoa; Rohit Bakshi, MD

Background: Cognition and magnetic resonance imaging correlations are well established in patients with multiple sclerosis (MS), but it is unclear whether lesion burden or atrophy accounts for most of the predictive variance. These indices have been directly compared in only a few studies. No such study included measurement of the third ventricle, which was strongly predictive of neuropsychological competence in the early literature. Furthermore, few studies accounted for the influence of age, premorbid intelligence, or depression.

Objective: To determine if conventional measures of lesion burden or atrophy predict cognitive dysfunction in MS while accounting for age, premorbid intelligence, and depression.

Methods: We studied 37 patients with MS and 27 controls matched according to demographic variables. Correlations between neuropsychological tests and the following magnetic resonance imaging indices were considered: T1 hypointense lesion volume, fluid-attenuated inversion recovery hyperintense lesion volume, third ventricle width, bicaudate ratio, and brain parenchymal fraction. Regression models predicting neuropsychological performance controlled for the effects of age, premorbid intelligence, and depression. We included only those tests discriminating patients with MS from controls.

Results: In each regression model, third ventricle width was the sole magnetic resonance imaging measure retained. When this variable was removed from consideration, brain parenchymal fraction was retained in all analyses.

Conclusions: Brain atrophy accounts for more variance than lesion burden in predicting cognitive impairment in MS, and central atrophy in particular is strongly associated with neuropsychological morbidity. This finding may be explained in part by atrophy of the thalamus, a deep gray-matter structure that mediates cognitive function via cortical and subcortical pathways. Enthusiasm for the clinical utility of third ventricle width is tempered by modest intraobserver and interobserver reliability.

Arch Neurol. 2004;61:226-230

A previous study by Rao et al showed that a linear measurement of the third ventricle derived from computed tomography was significantly correlated with cognition in multiple sclerosis (MS). More recent magnetic resonance imaging (MRI) studies demonstrated an association between lesion burden and cognitive impairment. In one study, T2 lesion area was a more significant predictor of most cognitive variables than either ventricular-brain ratio or corpus callosum size. Recently, T1 hypointense and fluid-attenuated inversion recovery (FLAIR)-defined lesions as well as whole-brain and bicaudate ratio atrophy measurements were correlated with cognitive impairment. Although cognitive dysfunction is correlated with measures of lesion burden and atrophy, it is unclear which dimension is most predictive.

In this study, MRI and cognition correlations were investigated to determine whether measures of lesion burden or atrophy account for most variance in MS-associated cognitive disorder. We measured 5 MRI variables including third ventricle width and used cognitive tests recommended in a recent position paper. In contrast to recent research, regression models predicting cognitive dysfunction controlled for the influence of age, premorbid intelligence, and depression.

METHODS

SUBJECTS

The participants included 37 patients with MS and 27 healthy volunteers. To facilitate gen-
eralization of results to the MS population, patients were recruited consecutively from an MS clinic. Informed consent was obtained per institutional review board requirements. Exclusion criteria were (1) a current or past medical or psychiatric disorder (including major depressive disorder and bipolar disorder) according to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV)9 other than MS that could affect cognitive function, (2) substance abuse, (3) neurologic impairment that might interfere with psychometric testing, and (4) MS relapse or corticosteroid use within the past 6 weeks.

Mean ± SD age was 41.8 ± 9.1 years with a range of 23 to 61 years. There were 28 (75%) women, and 35 (95%) were white. Mean ± SD level of education was 15.1 ± 2.2 years. Mean ± SD disease duration was 11.2 ± 8.0 years. The median Expanded Disability Status Scale10 score, determined by an experienced neurolologist specializing in MS, was 2.5 (range, 0-7.5). Thirty patients had relapsing-remitting MS, and the remainder had a secondary progressive course.

Healthy volunteers were enrolled to identify abnormal test results because we decided a priori that MRI variables would be regressed only for tests that discriminate patients from controls. The controls were matched with patients for age (mean ± SD, 40.8 ± 8.7 years; range, 21-57 years), level of education (mean ± SD, 14.5 ± 1.9 years), race, and sex.

**MRI PROTOCOL AND ANALYSIS**

Axial T1-weighted (repetition time/echo time: 400/10 ms), T2-weighted (3000/120 ms), and fast FLAIR images (repetition time/echo time/inversion time: 8000/120/2200 ms) were obtained using a scanner (Gyroscan ACS-NT 1.5-T; Philips Medical Systems, Andover, Mass). Analysis was performed on a workstation (Unix; Sun Microsystems) by a trained technician who was blind to the clinical data. As previously described,6,11,12 T1 lesions, defined as having a reduced signal compared with white matter and being at least partially hyperintense on FLAIR images, were segmented using an edge-finding tool. Total brain T1 hypointense lesion volume was calculated as the sum of the area of all lesions multiplied by the section thickness; FLAIR lesion volume was determined using a thresholding procedure as recently described.11 Bicaudate ratio and third ventricle width were measured from FLAIR images.11,13,11 The former is the ratio of the intercaudate distance to the brain width along the same line. To obtain third ventricle width, a line region of interest was drawn through the long axis of the ventricle, parallel to the interhemispheric fissure in the section where the third ventricle was most visible. The width was measured by drawing a second line perpendicular to the first at its midpoint and recording its length (**Figure**). Brain parenchymal fraction was based on our semi-automated technique,11 defined as the ratio of brain parenchymal volume (tissue compartment) to the total brain volume within the surface contour. The reliability of these measures as assessed by intraobserver and interobserver coefficients of variation, as reported previously,6,11,12 are as follows: 1.7% and 4.9% for T1 hypointense lesion volume, 1.2% and 3.1% for FLAIR lesion volume, 3.2% and 7% for third ventricle width, 2.3% and 4.2% for bicaudate ratio, and 0.31% and 0.34% for brain parenchymal fraction. Test-retest coefficients of variation were 0.41% for brain parenchymal fraction,11 4.36% for bicaudate ratio,6 and 2.4% for third ventricle width (R. Bakshi, unpublished data, 2003). These methods are significantly correlated with each other and with clinical parameters.6,11,12

**NEUROPSYCHOLOGICAL TESTING**

Neuropsychological tests, based on another recent position paper,7 were administered by a technician supervised by a board-certified neuropsychologist, both blind to the MRI findings. Premorbid intelligence was estimated using the North American Adult Reading Test (NAART).24 Attention (or working memory) was assessed using the adaptations by Rao et al26 of the Paced Auditory Serial Addition Test (PASAT) and the Symbol Digit Modalities Test28 (SDMT). The PASAT (3-second and 2-second presentation rates) required participants to monitor audiotaped digits while adding each consecutive digit to the preceding one. On the SDMT, participants were presented with symbol-number pairings at the top of an 8.5 × 11-inch page and asked to voice the digit for each unpaired symbol as quickly as possible. The number of correct responses was tabulated for each test (not available in 6 cases).

The second edition of the California Verbal Learning Test19 (CVLT-II) and the Brief Vusuospatial Memory Test—Revised20 (BVMT-R) were administered to assess memory in the verbal and spatial domains. Material was presented successively during repeated learning trials. After a 25-minute delay interval, patients were asked to recall the information and then recognize it in a yes/no format. The stimuli for the CVLT-II were 16 words and for the BVMT-R a matrix of 6 designs presented for 10 seconds. The variables of interest were recall summed across all learning trials (learning) and recall after a 25-minute delay interval (delayed recall).

Tests emphasizing spatial perception, language, and executive function were also administered, including the Judgment of Line Orientation Test,25 Boston Naming Test,22 Controlled Oral Word Association Test,21 and Wisconsin Card Sorting Test.23 The Judgment of Line Orientation Test required participants to match the angles of lines to a model depicting 10 of them. The dependent measure was the number correct of 30 stimuli. On the Boston Naming Test, participants orally retrieved the names for 60 drawings; the number...
WARD: stepwise procedure, with the Center for Epidemiologic Studies Depression Scale (CES-D-10). Patients rated the frequency of perseverative responses. Finally, depression was quantified based on feedback. We considered the total number of perseverative responses and the CES-D-10 score in the second. The MRI measures (T1 hypointense, FLAIR lesion volume, bicaudate ratio, third ventricle width, brain parenchymal fraction) were then entered in block 2.

**STATISTICAL ANALYSIS**

The groups were compared using 1-way analysis of variance and the χ2test. Regression analysis was limited to neuropsychological variables that discriminated between patients and controls; significance was set at P<.01. The approach was a forward stepwise procedure, with P<.05 to enter and P=.10 to exit. Covariates entered and retained in block 1 included age and NAART score in the first model and age, NAART score, and CES-D-10 score in the second. The MRI measures (T1 hypointense lesion volume, FLAIR lesion volume, bicaudate ratio, third ventricle width, and brain parenchymal fraction) were then entered in block 2.

**RESULTS**

There were no group differences for age, level of education, sex, race, or NAART score. As expected, patients with MS produced significantly higher CES-D-10 scores (P=.001).

The MRI measures were intercorrelated (Table 1). The strongest Pearson correlation (r=0.89; P<.001) was between T1 hypointense lesion volume and FLAIR lesion volume. Correlations among atrophy measures were strong, ranging from −0.68 (bicaudate ratio/brain parenchymal fraction; P<.001) to 0.80 (bicaudate ratio/third ventricle width; P<.001). Correlations between atrophy and lesion burden were moderate or not significant (brain parenchymal fraction/FLAIR lesion volume).

Patients performed more poorly than controls on the Judgment of Line Orientation Test, CVLT-II Learning, CVLT-II Delayed Recall, BVMT-R Learning, BVMT-R Delayed Recall, PASAT, and SDMT (Table 2). After controlling for age and premorbid intelligence (NAART) in regression models (Table 3), third ventricle width accounted for significant variance in CVLT-II Learning, CVLT-II Delayed Recall, BVMT-R Delayed Recall, BVMT-R Delayed Recall, PASAT, and SDMT performance. Similar results were obtained when depression (CES-D-10) was added to the covariates.

When the significant models were repeated without third ventricle width (Table 3), the CVLT-II analyses were not significant. For BVMT-R, the model accounting for age and premorbid intelligence retained brain parenchymal fraction, as did the model also accounting for CES-D-10 score. In addition, brain parenchymal fraction was retained in the model predicting PASAT score that accounted for age and premorbid intelligence. Both brain parenchymal fraction and FLAIR lesion volume were retained in the models predicting SDMT score, with brain parenchymal fraction entering first.

**COMMENT**

This study was designed to determine whether conventional measures of lesion burden or atrophy account for more variance in cognitive performance. All cognitive variables were predicted by third ventricle width in regression models controlling for age, premorbid intelligence, and depression. When third ventricle width was excluded, brain parenchymal fraction accounted for most variance. We therefore conclude that central and whole brain atrophy account for more variance in MS cognition than lesion burden.

We obtained stronger correlations between third ventricle width and cognitive function than in an early computed tomographic study.1 This may be attributed to reduction in measurement error with the advent of MRI. In addition, as proposed by Rao et al,1 the anatomy of the third ventricle may be less variable across patients than other areas of the ventricular system, further strengthening correlations with clinical measures. The predictive power of third ventricle width may also result from its proximity to the thalamus. Evidence for injury to the thalamus in MS includes increased water diffusion,26 T2 hypointensity (a marker of iron deposition), and decreased N-acetylaspartate level, neuronal degeneration, and macro-

---

### Table 2. Group Comparisons of Neuropsychological Data

<table>
<thead>
<tr>
<th>Test</th>
<th>Patients</th>
<th>Controls</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>North American Adult Reading Test</td>
<td>107.4 ± 9.6</td>
<td>108.6 ± 7.6</td>
<td>.58</td>
</tr>
<tr>
<td>Boston Naming Test</td>
<td>55.2 ± 5.6</td>
<td>55.6 ± 4.1</td>
<td>.73</td>
</tr>
<tr>
<td>Controlled Oral Word Association Test</td>
<td>38.6 ± 11.2</td>
<td>40.8 ± 12.4</td>
<td>.46</td>
</tr>
<tr>
<td>Judgment of Line Orientation Test</td>
<td>21.7 ± 4.8</td>
<td>26.4 ± 3.0</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>CVLT-II Learning</td>
<td>47.0 ± 10.4</td>
<td>57.8 ± 10.1</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>CVLT-II Delayed Recall</td>
<td>10.0 ± 3.6</td>
<td>13.0 ± 2.9</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>BVMT-R Learning</td>
<td>18.5 ± 6.4</td>
<td>27.9 ± 6.9</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>BVMT-R Delayed Recall</td>
<td>7.7 ± 2.7</td>
<td>10.8 ± 1.7</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Paced Auditory Serial Addition Test</td>
<td>33.5 ± 12.9</td>
<td>43.7 ± 8.8</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Symbol Digit Modalities Test</td>
<td>45.9 ± 13.9</td>
<td>63.6 ± 8.4</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

**Abbreviations:** BVMT-R, Brief Visuospatial Memory Test–Revised; CVLT-II, California Verbal Learning Test, second edition.

*Data are presented as mean ± SD.

---

### Table 1. Correlation Matrix

<table>
<thead>
<tr>
<th>T1 Hypointense Lesion Volume</th>
<th>FLAIR Lesion Volume</th>
<th>Bicaudate Ratio</th>
<th>Third Ventricle Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1 hypointense lesion volume</td>
<td>0.69†</td>
<td>0.49†</td>
<td>0.80†</td>
</tr>
<tr>
<td>FLAIR lesion volume</td>
<td>0.50†</td>
<td>0.47†</td>
<td>0.80†</td>
</tr>
<tr>
<td>Bicaudate ratio</td>
<td>0.49†</td>
<td>0.80†</td>
<td>0.79†</td>
</tr>
<tr>
<td>Third ventricle width</td>
<td>−0.82</td>
<td>−0.68</td>
<td>−0.79†</td>
</tr>
<tr>
<td>Brain parenchymal fraction</td>
<td>−0.42†</td>
<td>−0.26</td>
<td>−0.68†</td>
</tr>
</tbody>
</table>

**Abbreviation:** FLAIR, fluid-attenuated inversion recovery.

*Data are presented as Pearson correlation coefficients.

†P<.001.

‡P<.01, indicating significance.

---

©2004 American Medical Association. All rights reserved.
Table 3. Results of Significant Linear Regression Analyses

<table>
<thead>
<tr>
<th>Neuropsychological Test</th>
<th>Block 1 Covariates</th>
<th>Retained MRI Variables</th>
<th>Partial $r$ for Primary Predictor</th>
<th>Multiple $R^2$</th>
<th>Change in $R^2$ From Block 1</th>
<th>$P$ Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVLT-II Learning</td>
<td>Age, NAART score</td>
<td>Third ventricle width</td>
<td>-0.45</td>
<td>0.42</td>
<td>0.14</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>CVLT-II Delayed Recall</td>
<td>Age, NAART score</td>
<td>Third ventricle width</td>
<td>-0.44</td>
<td>0.26</td>
<td>0.15</td>
<td>&lt; .02</td>
</tr>
<tr>
<td>BVMT-R Delayed Recall</td>
<td>Age, NAART score</td>
<td>Third ventricle width</td>
<td>-0.45</td>
<td>0.44</td>
<td>0.14</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>PASAT</td>
<td>Age, NAART score</td>
<td>Third ventricle width</td>
<td>-0.47</td>
<td>0.42</td>
<td>0.17</td>
<td>&lt; .01</td>
</tr>
<tr>
<td>SDMT</td>
<td>Age, NAART score</td>
<td>Third ventricle width</td>
<td>-0.57</td>
<td>0.43</td>
<td>0.16</td>
<td>&lt; .01</td>
</tr>
</tbody>
</table>

Abbreviations: BVMT-R, Brief Visuospatial Memory Test–Revised; CES-D, Center for Epidemiologic Studies Depression Scale; CVLT-II, California Verbal Learning Test, second edition; FLAIR, fluid-attenuated inversion recovery; MRI, magnetic resonance imaging; NAART, North American Adult Reading Test; PASAT, Paced Auditory Serial Addition Test; SDMT, Symbol Digit Modalities Test.

Accepted for publication September 19, 2003.

**Author contributions:** Study concept and design (Drs Benedict, Weinstock-Guttman, and Bakshi); acquisition of data (Drs Benedict, Sharma, and Bakshi, Ms Fishman, and Mr Tjoa); analysis and interpretation of data (Dr Benedict and Bakshi and Mr Tjoa); drafting of the manuscript (Drs Benedict, Weinstock-Guttman, and Bakshi and Ms Fishman); critical revision of the manuscript for important intellectual content (Drs Benedict and Sharma and Mr Tjoa); statistical expertise (Dr Benedict); administrative, technical, and material support (Drs Benedict, Weinstock-Guttman, Sharma, and Bakshi, Ms Fishman, and Mr Tjoa); study supervision (Drs Benedict and Bakshi).

This study was supported in part by grants NIH-NINDS 1 K23 NS42379-01 (Dr Bakshi) from the National Institutes of Health, Bethesda, Md; RG 3258A2/1 (Drs Weinstock-Guttman and Bakshi) from the National Multiple Sclerosis Society, New York, NY; and DBI-0234895 (Drs Weinstock-Guttman and Bakshi) from the National Science Foundation, Arlington, Va.

Corresponding author: Ralph H. Benedict, PhD, Department of Neurology, Suite D-6, SUNY at Buffalo School of Medicine and Biomedical Sciences, Buffalo General Hospital, 100 High St, Buffalo, NY 14203 (e-mail: benedict@buffalo.edu).

linear measure of third ventricle width explained most of the variance in neuropsychological outcomes. The strength of the correlation may be attributed to the careful assessment of cognitive functions, presumed validity of the third ventricle width measure, and anatomical significance of the thalamus. This study highlights the close relationship between brain atrophy and cognitive impairment after accounting for the effect of plaque load.

scopic volume loss. A recent study also showed that third ventricular width is inversely correlated with thalamic volume. Retrograde (wallerian) degeneration, which has been demonstrated in MS, could contribute to atrophy of subcortical structures. Moreover, the thalamus has reciprocal connections with widespread areas of the neocortex and other deep gray-matter structures. Because these interconnections are mediated by long white-matter tracts, such axonopathy could contribute to dysfunction and atrophy of this structure.

One weakness of this study is a lack of recently developed MRI measures such as diffusion tensor imaging, magnetization transfer, and spectroscopy. A more important limitation is the low test-retest reliability of third ventricle width measurements as compared with 3-dimensional quantitative MRI measurements of brain atrophy, which are less prone to variability because of changes in head position between imaging procedures. Future studies should investigate the utility of highly reproducible 3-dimensional measurement of third ventricle size, such as that proposed recently. The methodological strengths of this study are the use of computer-assisted quantitative MRI analysis methods, controlling for premorbid intelligence and depression, and application of reliable and well-validated measures of cognitive function. The battery used, an adapted version of the Minimal Assessment of Cognitive Function in MS, represents a consensus opinion of neuropsychology experts based on psychometric and validity criteria. Use of such measures may optimize power to detect associations between imaging and cognition.

In conclusion, we compared measures of lesion burden and atrophy to determine which index explained more variance in the cognitive performance of patients with MS. Of the 5 measures investigated, an easily calculated
REFERENCES


