Medulla Oblongata Volume: A Biomarker of Spinal Cord Damage and Disability in Multiple Sclerosis

BACKGROUND AND PURPOSE: While brain MR imaging is routinely performed, the MR imaging assessment of spinal cord pathology in multiple sclerosis (MS) is less frequent in clinical practice. The purpose of this study was to determine whether measurements of medulla oblongata volume (MOV) on routine brain MR imaging could serve as a biomarker of spinal cord damage and disability in MS.

MATERIALS AND METHODS: We identified 45 patients with MS with both head and cervical spinal cord MR imaging and 29 age-matched and sex-matched healthy control subjects with head MR imaging. Disability was assessed by the expanded disability status scale (EDSS) and ambulation index (AI). MOV and upper cervical cord volume (UCCV) were manually segmented; semiautomated segmentation was used for brain parenchymal fraction (BPF). These measures were compared between groups, and linear regression models were built to predict disability.

RESULTS: In the patients, MOV correlated significantly with UCCV ($r = 0.67$), BPF ($r = 0.45$), disease duration ($r = -0.64$), age ($r = -0.47$), EDSS score ($r = -0.49$) and AI ($r = -0.52$). Volume loss of the medulla oblongata was $-0.008 \text{ cm}^3/\text{year}$ of age in patients with MS, but no significant linear relationship with age was found for healthy control subjects. The patients had a smaller MOV (mean $\pm$ SD, 1.02 $\pm$ 0.17 cm$^3$) than healthy control subjects (1.15 $\pm$ 0.15 cm$^3$), though BPF was unable to distinguish between these 2 groups. MOV was smaller in patients with progressive MS (secondary-progressive MS, 0.88 $\pm$ 0.19 cm$^3$ and primary-progressive MS, 0.98 $\pm$ 0.30 cm$^3$) than in patients with relapsing-remitting MS (1.08 $\pm$ 0.15 cm$^3$). A model including both MOV and BPF better predicted AI than BPF alone ($P = .04$). Good reproducibility in MOV measurements was demonstrated for intrarater (intraclass correlation coefficient, 0.97), interrater (0.79), and scan rescan data (0.81).

CONCLUSION: MOV is associated with disability in MS and can serve as a biomarker of spinal cord damage.

Received February 28, 2008; accepted April 1.

From the Center for Neurological Imaging (Z.L., A.M.B., M.P.S., A.C., P.H., R.B., C.R.G.G.), and Partners Multiple Sclerosis Center (S.J.K., H.L.W.), Brigham and Women’s Hospital, Harvard Medical School, Boston, Mass.; Massachusetts Institute of Technology (O.F.), Cambridge, Mass.; Brigham and Women’s Hospital (B.C.H.), Harvard Medical School, Boston, Mass.; and Lahey Clinic (P.H.), Burlington, Mass.

This study was supported in part by funding provided by the National Multiple Sclerosis Society grant as well as the National Institutes of Health. Z. Liptak and A.M. Berger contributed equally to the study.


Please address correspondence to Charles R.G. Guttmann, MD, Center for Neurological Imaging, 221 Longwood Ave, RF 394, Boston, MA 02115; e-mail: guttmann@bwh.harvard.edu

DOI 10.3174/ajnr.A1162

Copyright 2008 by American Society of Neuroradiology.
Materials and Methods

Patient Selection
For this retrospective analysis, patients were selected from the general patient population and the Comprehensive Longitudinal Investigation of Multiple Sclerosis at the Brigham and Women’s Hospital study at the Partners MS Center in Boston. A data base search was conducted in October 2004. At that time, the data base contained 251 patients who fulfilled our inclusion criteria: both a head and spinal cord MR imaging examination performed on the same day, paired with a clinical visit that was less than 30 days before or later than the MR imaging scan. Of these 251 patients, 50 with good quality imaging studies (head MR images encompassing the cranial cavity from the foramen magnum to upper convexity without major image artifacts) were chosen randomly to be the subjects of this retrospective study. Of 50 patients, 4 were excluded because they received corticosteroids up to 30 days before the MR imaging examination, a potential confounder of atrophy measures. One patient was excluded because of failure to fulfill the diagnostic criteria of MS according to the International Panel criteria. The group of 45 patients with MS was composed of 32 with relapsing-remitting (RR) MS, 8 with secondary-progressive (SP) MS, and 5 with primary-progressive (PP) MS.

Within 1 month of the MR imaging examination, every patient underwent a full neurologic examination and a clinical evaluation, including the expanded disability status scale (EDSS) and ambulation index (AI), by a neurologist who specializes in MS. Although all patients were participating in various research studies and drug trials, 13 patients were receiving no disease-modifying treatment during the period of the study because of their early or mild disease course. Thirteen patients were treated with glatiramer acetate, 11 with interferon beta 1a, 3 with corticosteroids (but not within the last month), 3 with cyclophosphamide, 1 with interferon beta 1b, and 1 with intravenous immunoglobulin. The patients were compared with an age-matched and sex-matched group of 29 healthy subjects. To ascertain the reproducibility of the measurement technique, we also evaluated the MR images of 20 patients from a separate study on segmentation reproducibility.

Brain Morphometry
For the measurement of the ICC and brain parenchymal fraction (BPF), we used template-driven segmentation plus partial volume effect correction (TDS+), which is a segmentation method that uses both signal intensity characteristics and anatomic locations to subdivide the brain into WM (normal-appearing WM and lesions), gray matter (GM), and CSF. TDS+ carried out on the dual-echo images of the brain. The BPF was defined as BPF = (WM + GM)/(WM + GM + CSF) = 1 - CSF/ICC.
**Statistical Analysis**

We calculated the correlation coefficients among MOV, UCCV, BPF, disease duration, and age using the Pearson correlation coefficient. As the EDSS and AI are ordinal variables, we used the Spearman rank correlation coefficient to measure the relationship between these variables and MOV, UCCV, and BPF. The linear relationship between all of the factors was confirmed by the appropriate plots (data not shown). To assess the significance of differences between correlation coefficients, we used the 2-sided Meng test \(^1^1\) for all Pearson correlation coefficients and the test proposed by Choi \(^1^2\) for Spearman correlation coefficients. The Dunn-Sidak method was used to correct for multiple comparisons. \(^1^3\) In addition, we examined whether MOV or UCCV significantly improved the model fit for the relationship between clinical disability landmarks and BPF by use of logistic regression models. Cutoff values of >4 for EDSS and >3 for AI were used to classify patients into 2 groups because these cutoff values reflected clinically relevant changes in the patient’s ambulatory capacity. The likelihood ratio test was used to select the best model.

We compared the age-related MOV change in patients with MS and healthy control subjects using the following linear regression model:

\[
MOV = \beta_0 + \beta_1 \times \text{age} + \beta_2 \times \text{MS} + \beta_3 \times \text{age} \times \text{MS} + e,
\]

where MS is an indicator for disease status. If \(\beta_1\) is significantly different from 0, the rates of change in the healthy control subjects and the patients with MS are significantly different from each other.

Although the patients were mostly in the RR phase, the presence of a small sample of SPMS and PPMS patients allowed for a preliminary post hoc analysis. Because of the limited sample size, the Kruskal-Wallis and Mann-Whitney tests were used for the overall and pairwise comparisons. As above, to correct for multiple comparisons, we used the Dunn-Sidak method.

Finally, we calculated the intraclass correlation coefficient (ICCcoeff) \(^3^4\) to determine the reproducibility of the MOV measurement. For the computation of the intrarater agreement and reproducibility of MOV measurement on scan-rescan images, a 1-way random model was used. \(^3^4\) For interrater agreement, we fit a 2-way random model using 1 observation per rater per scan. Although different guidelines exist for the interpretation of ICCcoeff, we applied a scale \(^3^5\) in which an ICCcoeff value of less than 0.40 indicates poor reproducibility; ICCcoeff values in the range 0.40 to 0.75 indicate fair to good reproducibility; and an ICCcoeff value of greater than 0.75 shows excellent reproducibility. We performed all statistical analyses using SPSS (version 14.0.1; SPSS, Chicago, Ill), StatXact (version 6; Cytel, Cambridge, Mass), and Matlab (version 2006b; Math Works, Natick, Mass).

**Results**

**Demographics**

Demographic and clinical characteristics of the 45 patients with MS and 29 healthy control subjects are shown in on-line Table 1. Patients and control subjects had similar distributions for age and sex. On-line Table 1 also compares demographic and clinical characteristics of the study’s patient sample and the overall MS patient population of our clinic at the time of patient selection (October 2004). The patient sample seems to match our clinic’s patient population reasonably well, with perhaps a slight underrepresentation of patients with SP compared with those with PP (all differences not significant, \(P > .110\)).

**Does MOV Reflect Spinal Cord Pathologic Damage?**

The correlation coefficient between MOV and UCCV was \(r = 0.67\), whereas that between MOV and BPF was \(r = 0.45\). This difference did not reach statistical significance (Meng test: \(P = .086;\) difference in the correlation coefficients = 0.22, confidence interval –0.05 to 0.67). To put this finding in the context of measurement variability within the cervical spinal cord, we subdivided UCCV into 2 volumes of 3 sections each. The first one contained the volume of spinal cord at the level of C2; the second was the volume of the cord at the level of C3. The correlation between these 2 volumes was \(r = 0.85\).

**MOV and Motor Disability**

MOV and UCCV showed similar degree of correlation with EDSS (\(r = −0.49\) and \(r = −0.48\), respectively). The correlation coefficient between EDSS and BPF was \(r = −0.59\). The test proposed by Choi \(^3^2\) did not demonstrate a significant difference between these correlation coefficients. The correlation of AI with MOV, UCCV, and BPF were \(r = −0.52, r = −0.52,\) and \(r = −0.58\), respectively (on-line Table 2). Again, none of these correlation coefficients were significantly different.

In terms of clinical landmarks of disability, the optimal model for EDSS of more than 4 included only BPF; including MOV (\(P = .287\)) or UCCV (\(P = .125\)) did not significantly improve the model. The odds ratio for 0.01 U decrease in BPF was 1.53. For AI of more than 3, including MOV in addition to BPF significantly improved model fit (\(P = .040\)), but the addition of UCCV did not (\(P = .066\)). For the final model with MOV and BPF, the odds ratios for a 0.01 decrease in MOV and BPF, holding the other factor constant, were 1.09 and 1.11, respectively.

**Age and Disease Effects on MOV and UCCV**

In our patients, both age and disease duration showed statistically significant inverse correlations with MOV (\(r = −0.47\) and \(r = −0.64\), respectively). In a similar fashion, the correlations between UCCV and disease duration (\(r = −0.61\)) or BPF and disease duration (\(r = −0.50\)) also showed moderate correlation (on-line Table 2).

**Difference between Healthy Control Subjects and Patients with MS on MOV**

Patients with MS had a significantly smaller MOV (mean ± SD, 1.02 ± 0.17 cm³) than healthy control subjects (1.15 ± 0.15 cm³) (\(P = .012\)). In the patients, age-related medulla oblongata atrophy was −0.008 cm³/year of age (\(P = .001\)), whereas for healthy subjects, no significant linear relationship between MOV and age was found (0.002 cm³/year of age, \(P = .324\)) with use of the linear regression model. Age-related changes of MOV were significantly different between patients and healthy control subjects (\(P = .003\)). There was no significant difference between the BPF of the patients (mean ± SD, 0.86 ± 0.05 cm³) and the BPF of the healthy control subjects (mean ± SD, 0.87 ± 0.04 cm³).

**Post Hoc Analysis**

As this study was primarily designed to validate the measurement of MOV, subjects were not selected according to clinical subtypes. The patients with MS in the study population were...
mostly in the RR phase \((n = 32)\); 8 had SPMS and 5 had PPMS. For all 3 MR imaging measures, the Kruskal-Wallis test showed significant differences between the groups \((P < .001\) for MOV, \(P = .028\) for UCCV, and \(P < .001\) for BPF). A preliminary post hoc analysis was performed to evaluate the differences in MOV between the established clinical phenotypes \(\text{(on-line Table 3 and 4).}\) The largest mean MOV was found in patients with RRMS \((1.08 \text{ cm}^3)\), whereas patients with SPMS or PPMS had a lower mean MOV \((0.88 \text{ cm}^3)\) and PPMS, \(0.95 \text{ cm}^3\). The difference between the MOV of RRMS and patients with progressive MS \((\text{SPMS or PPMS})\) was statistically significant \(\text{(on-line Table 4).}\) BPF was also able to differentiate between RRMS and patients with progressive MS, whereas UCCV was not. No significant difference was found between patients with SPMS and those with PPMS or between those with RRMS and healthy control subjects \(\text{(on-line Table 4).}\)

Reproducibility Analysis

We assessed the intrarater, interrater, and scan-rescan reproducibility of MOV using the ICCoeff. For the intrarater and interrater reproducibility analysis, ICCoeff values of 0.97 and 0.79 were obtained, indicating excellent reproducibility.\(^{35}\) The scan-rescan measurements were also highly reproducible with an ICCoeff of 0.81.

Discussion

We propose a relatively simple method to measure MOV that could serve as a biomarker of spinal cord damage and clinical disability in MS. The correlation between MOV and UCCV \((r = 0.67)\) underscores the relative sensitivity of this metric to spinal cord atrophy. In comparison with our results, a previous study\(^{36}\) reported smaller correlation coefficients between the whole brain stem volume and the upper cervical cord area in patients with RRMS \((r = 0.35)\) and patients with SPMS \((r = 0.38)\) possibly because of the inclusion of brain stem gray matter. In aggregate, these findings support the role of MOV as a biomarker of spinal cord damage in MS.

To assess the strength of associations between the MR imaging measures \((\text{UCCV and MOV})\) and clinical disability, we compared our results with those of previous studies. These studies have reported correlation coefficients between the upper cervical cord area and EDSS ranging between \(-0.34\) and \(-0.70^{12,15,27,37}\) and a correlation of \(-0.60\) between the upper cervical cord area and AI.\(^{27}\) Our findings are consistent with these results, as the correlation of UCCV with EDSS and AI were \(-0.48\) and \(-0.52\), respectively. MOV showed a similar relationship with clinical disability: the correlation coefficients between MOV and EDSS or AI were \(r = -0.49\) and \(-0.52\), respectively. In summary, the associations of MOV and clinical scores emulate those with previous spinal cord metrics.

Many studies have investigated the use of BPF as a predictor of EDSS.\(^{38-41}\) One previous MR imaging study reported a correlation between BPF and EDSS \((r = -0.391)\) and found that BPF was the best predictor of EDSS.\(^{42}\) Therefore, we included BPF in all prediction models and tested whether the addition of UCCV or MOV would add predictive value to those models. The logistic regression analysis showed that the models predicting EDSS did not improve with the addition of either MOV or UCCV to BPF. However, BPF and MOV predicted AI better than BPF alone, whereas including UCCV instead of MOV did not significantly improve the predictive model of AI. However, it should be noted that our spinal cord measurements were not done on MR images with a single optimized MR imaging protocol, whereas the brain measurements were. This might be an explanation for the slightly better results of MOV compared with UCCV. To assess the fidelity of our UCCV measurements, we compared them to findings reported in the literature. In our study, the mean cross-sectional area was 70.64 mm\(^2\) for PPMS, 60.00 mm\(^2\) for SPMS, and 74.52 mm\(^2\) for RRMS. Previous studies have reported means of cross-sectional areas of the cervical spinal cord \(\text{(at the level of C2) ranging from 67.82 to 73.1 mm}\(^2\) for PPMS, 61.2 mm\(^2\) for SPMS, and 72.65 to 85.6 mm\(^2\) for RRMS.\(^{12,26,43}\) These results are consistent with our results despite some variability in the acquisition parameters of our MR images of the spinal cord.

MR imaging–based studies have reported a correlation of \(r = -0.281\) between disease duration and BPF\(^{42}\) and a correlation of \(r = -0.75\) between disease duration and spinal cord measures.\(^{37}\) In our study, the correlations of disease duration with BPF and UCCV \((r = -0.50\) and \(-0.61\), respectively) were consistent with those of published findings. A similar relationship was found between disease duration and MOV \((r = -0.64)\). In summary, MOV findings parallel findings in the spinal cord in the detection of degenerative aspects of MS.

To further document the atrophy of the medulla oblongata in MS, we compared the MOV of the patients with that of healthy control subjects and found a significant difference between these groups. In a similar fashion, histopathologic\(^{27,8,44}\) and MR imaging studies\(^{12,15,37}\) have shown that spinal cord measures are smaller in patients with MS than in healthy control subjects. Age-related brain atrophy is considered a normal aspect of aging and is a confounder in the assessment of MS-specific degeneration with BPF. In contrast, previous studies and our results on MOV suggest that the medulla oblongata, the whole brain stem, as well as the spinal cord, do not show significant age-related atrophy in healthy people.\(^{24-26,45}\) Because MOV does not seem to change significantly with age in healthy volunteers but shows disease-dependent atrophy in patients with MS, the decline in MOV might provide a valuable complementary measure of degenerative processes in MS that is not as prone to covariance with age as BPF.

Our preliminary subgroup analysis demonstrates that patients with a progressive disease course \((\text{SPMS or PPMS})\) had a smaller MOV than patients with RRMS or healthy control subjects. Other studies have also demonstrated an increased reduction of the spinal cord in patients with progressive MS compared with patients with RR or healthy control subjects.\(^{12,15,37,43}\) As our results in the MS subgroup comparisons are based on small numbers of patients with SP and PPMS, future studies are required to confirm these preliminary results.

Good intraobserver and interobserver agreement was demonstrated for the measurement of MOV. The ICCoeff value was 0.97 for intrarater and 0.79 for interrater reproducibility. An ICCoeff value of greater than 0.75 denotes excellent agreement.\(^{30}\) In this work, we studied the interobserver agreement only between 2 observers. Future studies should assess the interobserver agreement among a larger number of observers to
Further assess the robustness of this measure. Note that in this work, all medulla oblongata and upper cervical cord measurements were divided by the subject/population ratio of the ICC to normalize these measures. A similar methodology was applied in several previous studies. 

Although our MOV measurement technique seems to be adequate for assessing group differences in MOV, further improvement in the measurement technique may be required to assess longitudinal changes in individual patients. The use of higher-resolution sequences could provide a more accurate and robust measurement of MOV. We also did not try to account for partial volume effects, which might improve reproducibility and sensitivity. In future studies, we plan to address these methodological challenges.

In this study, only patients who received both a brain and spinal cord MR imaging examination were included. Because not every patient at our center undergoes spinal cord MR imaging examination, this may be a potential source of selection bias. To address this issue, we compared the demographic and clinical characteristics of the patients in our study to the overall MS patient population at our center. Because no differences were found between these groups on several characteristics, the study patient population was a good representative sample of our MS patient population. Despite the fact that most patients were undergoing treatment, we were able to find a difference between patients with MS and healthy control subjects. However, the study was not designed to observe treatment effects, and we did not control for this factor. To be considered as a surrogate marker, future studies should assess whether MOV is able to capture specific treatment effects.

A significant correlation between axonal loss and N-acetyl-aspartate levels was found in spinal cord lesions of patients with MS with severe disability. Axonal loss results from both primary (axonal transection) and secondary (wallerian/retro-synaptic degeneration) mechanisms. Axonal degeneration is likely to be the most significant factor causing spinal cord atrophy, rather than the loss of tissue related to individual cord lesions. In addition, local cord atrophy typically does not develop at the site of lesions. In our study, we found 14 patients with lesions in the portion of the medulla oblongata that we measured. No significant difference was found when patients with or without lesions were compared (data not shown). It seems reasonable to assume that both the spinal cord and the medulla oblongata are affected in a similar way by the pathologic destructive processes of MS, as they are predominately composed of the same ascending and descending fiber tracts as the ones passing through the spinal cord. It is not too surprising that we were unable to demonstrate a relationship between spinal cord lesion load and atrophy (UCCV or MOV; data not shown) but did not deem the quality of our images sufficient to conclusively disprove such a relationship. The assessment of spinal cord lesions in clinical practice is still considered challenging and might be a contributing factor to the lack of correlation between spinal lesion load and EDSS, even when spinal cord atrophy and EDSS were correlated. MOV, as UCCV, may currently represent the only viable alternative to assess spinal cord involvement in clinical practice.

Furthermore, we suggest that the medulla oblongata may provide a unique opportunity to study the neurodegenerative component of MS. One plausible mechanism of reduction in MOV is linked to primary or secondary degeneration of axons superior and inferior to the medulla. The segmental/inter-nodal organization of axonal myelination by oligodendrocytes makes it unlikely that demyelination outside of the medulla would directly contribute to medullary atrophy. Therefore, we suggest that MOV might also provide a simple way to assess axonal degeneration in MS.

MOV is an easily obtainable measure that presents nonnegligible advantages, especially in the context of large-scale clinical trials in MS. Obtaining measures of MOV is relatively easier than obtaining measures of spinal cord volumes because the contrast between the surrounding CSF and the medulla oblongata is high on T2-weighted images. Furthermore, the medulla oblongata has a larger diameter than that of the spinal cord, which facilitates its cross-sectional outlining.

Conclusion
We have proposed a straightforward method for MOV measurement on routine clinical head MR imaging. Preliminary results have demonstrated the use of MOV as a potential biomarker of spinal cord damage in MS.

References
presentation to predict conversion to clinically definite multiple sclerosis. Brain 1997;120(Pt 11):2059–69