Semiautomated Evaluation of the Primary Motor Cortex in Patients with Amyotrophic Lateral Sclerosis at 3T


ABSTRACT

BACKGROUND AND PURPOSE: Amyotrophic lateral sclerosis is a neurodegenerative disease involving the upper and lower motor neurons. In amyotrophic lateral sclerosis, pathologic changes in the primary motor cortex include Betz cell depletion and the presence of reactive iron-loaded microglia, detectable on 7T MR images as atrophy and T2*-hypointensity. Our purposes were the following: 1) to investigate the signal hypointensity-to-thickness ratio of the primary motor cortex as a radiologic marker of upper motor neuron involvement in amyotrophic lateral sclerosis with a semiautomated method at 3T, 2) to compare 3T and 7T results, and 3) to evaluate whether semiautomated measurement outperforms visual image assessment.

MATERIALS AND METHODS: We investigated 27 patients and 13 healthy subjects at 3T, and 19 patients and 18 healthy subjects at 7T, performing a high-resolution 3D multiecho T2*-weighted sequence targeting the primary motor cortex. The signal hypointensity-to-thickness ratio of the primary motor cortex was calculated with a semiautomated method depicting signal intensity profiles of the cortex. Images were also visually classified as “pathologic” or “nonpathologic” based on the primary motor cortex signal intensity and thickness.

RESULTS: The signal hypointensity-to-thickness ratio of the primary motor cortex was greater in patients than in controls ($P < .001$), and it correlated with upper motor neuron impairment in patients ($\rho = 0.57$, $P < .001$). The diagnostic accuracy of the signal hypointensity-to-thickness ratio was high at 3T (area under the curve $= 0.89$) and even higher at 7T (area under the curve $= 0.94$). The sensitivity of the semiautomated method (0.81) outperformed the sensitivity of the visual assessment (0.56–0.63) at 3T.

CONCLUSIONS: The signal hypointensity-to-thickness ratio of the primary motor cortex calculated with a semiautomated method is suggested as a radiologic marker of upper motor neuron burden in patients with amyotrophic lateral sclerosis. This semiautomated method may be useful for improving the subjective radiologic evaluation of upper motor neuron pathology in patients suspected of having amyotrophic lateral sclerosis.

ABBREVIATIONS: ALS = amyotrophic lateral sclerosis; HS = healthy subjects; MI = primary motor cortex; ROC = receiver operating characteristic; SH/Thk = signal hypointensity-to-thickness ratio; UMN = upper motor neuron

Am}yotrophic lateral sclerosis (ALS) is a progressive and clinically heterogeneous neurodegenerative disease involving both upper and lower motor neurons, having different prognoses and, perhaps, different responses to possible therapies, even in the experimental scenario. Different from the lower motor neuron impairment that can be carefully investigated with electrophysiologic tests, the evaluation of upper motor neuron (UMN) burden is mainly clinical and is partially confounded by signs related to lower motor neuron degeneration. Moreover, at the time of the diagnosis, the UMN impairment can range widely from faint to severe, and the variability in signs and symptoms at disease onset makes early diagnosis and correct phenotypic characterization of the disease difficult.

In ALS, the main pathologic changes in the primary motor...
cortex (M1) include the loss of Betz cells in the layer V\textsuperscript{o} and the presence of reactive iron-loaded microglia,\textsuperscript{6,7} visible on MR images as cortical atrophy\textsuperscript{8-10} and T2\textsuperscript{\*} hypointensity,\textsuperscript{7} respectively. In recent years, several conventional and nonconventional MR imaging techniques have been used to look for a biomarker of UMN impairment at both cortical and subcortical levels with variable results, and quantitative measurements of cortical atrophy were performed at a group level.\textsuperscript{11} Despite such effort, a definite and reliable marker of UMN degeneration is not yet available. As a result, while MR imaging of the brain is currently used to exclude mimic pathology and the detection of the T2 hypointensity of M1 can support suspicion of ALS, the specific search for this abnormality is not recommended for ALS diagnosis.\textsuperscript{12}

The first attempt to move toward the radiologic diagnosis at the single-subject level was recently performed with an ultra-high-field MR imaging system (7T).\textsuperscript{13} Taking advantage of the very high sensitivity of ultra-high-field strength to the magnetic susceptibility of microglial ferritin, the authors localized pathologic cortical thinning and T2\textsuperscript{\*} hypointensity in the deep layers of M1, and they were shown to significantly correlate with the clinical UMN burden. In light of these results, the T2\textsuperscript{\*} hypointensity of M1 was suggested as a possible marker of neuroinflammation and UMN impairment in patients with ALS rather than a marker of the disease. Unfortunately, the T2\textsuperscript{\*} sensitivity to microglial ferritin depends on the magnetic field strength, and M1 assessment in patients with ALS can be a challenge in clinical practice using MR imaging systems up to 1.5T also in patients with a severe UMN impairment. On the contrary, 3T scanners may change the radiologic approach to patients with pyramidal symptoms and signs in motor neuron diseases. Therefore, our main aim was to investigate the signal hypointensity-to-thickness ratio (SH/Thk) of the deep layers of the M1 as a radiologic marker of UMN burden in patients with ALS with a semiautomated method at 3T. Secondary aims were to evaluate whether the results obtained with a clinical scanner (3T) were comparable with those achieved with a research scanner (7T) and whether the semiautomated measurements improved the sensitivity of the visual radiologic assessment.

**MATERIALS AND METHODS**

**Patients with ALS and Healthy Subjects**

Twenty-seven patients with ALS (18 men and 9 women; mean age, 58 ± 11 years) and 13 healthy subjects (HS; 6 men and 7 women; mean age, 56 ± 15 years) underwent brain MR imaging with a 3T system.

Nineteen patients with ALS (14 men and 5 women; mean age, 63 ± 10 years), different from the patients with ALS investigated at 3T, and 18 HS (9 men and 9 women; mean age, 56 ± 13 years) underwent brain MR imaging with a 7T system.

Five of the above-mentioned HS underwent both examinations to compare the performance of the 3T and 7T systems in the assessment of signal intensity and thickness of the deep layers of M1.

All patients had a diagnosis of definite ALS according to the revised El Escorial criteria\textsuperscript{1} and were consecutively enrolled by the neurology unit of the University Hospital of Pisa. They were clinically evaluated by an experienced neurologist on the day of enrollment, and the UMN impairment was quantified for each limb using a clinical composite semiquantitative arbitrary score of UMN burden (UMN score), according to that previously used by Cosottini et al.\textsuperscript{11} For each patient, the total UMN score (range, 0–33) and the UMN score of each limb (range, 0–8) were recorded; then, the average UMN score was calculated as the mean of UMN scores of both arms and legs. Clinical and demographic data of patients are reported in On-line Tables 1–3. Exclusion criteria for enrollment were the presence of neurologic comorbidities. HS were enrolled from among relatives and spouses of patients with ALS and radiology department staff; none had any history of neurologic or psychiatric diseases.

All patients and controls gave their written informed consent for the enrollment. This study was performed as part of the experimental protocol called “Clinical Impact of Ultra-High-Field MRI in Neurodegenerative Diseases Diagnosis,” RF-2009-1546281, approved by the Italian Ministry of Health and by the local ethics committee. The project was founded by the Italian Ministry of Health and cofunded by the Health Service of Tuscany.

**MR Imaging Acquisition**

The MR imaging protocol at both 3T and 7T included a 3D multiecho T2\textsuperscript{\*}-weighted sequence prescribed axially and covering the brain from the vertex to the splenium of the corpus callosum.

MR imaging examinations at 3T were performed with a Discovery MR 750 (GE Healthcare, Milwaukee, Wisconsin) scanner equipped with an 8-channel head coil. Acquisition parameters of the 3D multiecho T2\textsuperscript{\*}-weighted sequence were the following: TR = 68.3 ms; TEs = 13, 18.6, 24.3, 29.9, 35.5, 41.2, 46.8, 52.4, 58.1, 63.7 ms; flip angle = 15°; NEX = 0.70; acquisition matrix = 448 × 384; FOV = 20 × 20 cm; spatial resolution of reconstructed images = 0.39 × 0.39 × 1 mm\textsuperscript{3}; scan duration = 4 minutes 22 seconds.

MR imaging examinations at 7T were performed with a Discovery MR 950 scanner (GE Healthcare) equipped with a 2CH-TX/32CH-RX head coil (Nova Medical, Wilmington, Massachusetts). Technical parameters of the 3D multiecho T2\textsuperscript{\*}-weighted sequence were the following: TR = 54.1 ms; TEs = 5.6, 12, 18.3, 24.7, 31.1, 37.5, 43.9 ms; flip angle = 15°; NEX = 0.70; acquisition matrix = 448 × 448; FOV = 22.4 × 22.4 cm; spatial resolution of reconstructed images = 0.5 × 0.5 × 1 mm\textsuperscript{3}; scan duration = 6 minutes 59 seconds.

**Semiautomated Image Assessment**

In each hemisphere of all subjects, the thickness and signal intensity of the regions of M1 corresponding to Penfield areas of the upper\textsuperscript{14} and lower\textsuperscript{15} limbs were assessed with an in-house-developed, semiautomated tool for image processing. Given the wide cortical extension of the upper and lower limb motor areas, to increase the reliability of the ROI position among subjects, we selected 2 smaller M1 subregions for the ROI positioning; thus, ROIs of the upper limbs were positioned in the hand knob, whereas ROIs of the lower limbs were positioned in the most cranial and lateral part of the paracentral lobule. In each set of images, a neuroradiologist blinded to the clinical diagnosis identified the sections that best represented each M1 target region and an additional section including the splenium of the corpus callo-
sum, which served to obtain 1 region of reference for the M1 signal intensity measures. The interactive image-processing tools were run according to the steps described in detail in On-line Fig 1. The observer is prompted to draw a polygonal ROI (not < 5 mm²) in the splenium of the corpus callosum, whose average intensity was retained to normalize the intensity values of the cortex ROIs. For each M1 target region, the observer manually draws the profile of the interface between the M1 and the neighboring subarachnoid space, which is interpolated by the software with a spline function to make it smoother. The directions normal to the spline are computed, and the trend of the signal intensity (y-axis) is reported as the function of the distance along the normal direction to the cortex (x-axis) in physical units (millimeter). The signal intensity profiles are then averaged and the absolute value is considered. A baseline intensity value corresponding to the spline is computed, and the trend of the signal intensity is retained to normalize the intensity values of the deep layers of M1 compared with those of neighboring cortices or “pathologic” if they were markedly more hypointense and thin. After 1 month from the first reading, the same neuroradiologists were asked to again assess the whole set of images.

Statistical Analysis
Quantitative and semiquantitative data were analyzed using non-parametric statistical tests with the significance level set to .05. In more detail, 3T and 7T data concerning HS who underwent both MR imaging examinations were compared using the Wilcoxon test. The SH/Thk comparison between patients with ALS and HS was performed using the Friedman test, whereas all other intergroup comparisons were investigated using the Mann-Whitney U test. Relationships between variables were investigated with the Spearman rank test, and ROC analysis was used as a binary classifier system to evaluate the performance of the SH/Thk in distinguishing patients from HS.

Using the clinical diagnosis as the criterion standard, we calculated the sensitivity, specificity, positive predictive value, negative predictive value, and diagnostic accuracy of the visual assessment of images for both reading sessions of both observers (for each reader, data shown in the “Results” section refer to the reading session with the best diagnostic accuracy). The Cohen κ statistic was used to calculate intra- and interrater reliability.

RESULTS
Epidemiologic and Clinical Data Analysis
The age of patients was not significantly different from that of HS at both 7T and 3T (P = .12 and P = .64, respectively). Total UMN scores and mean UMN scores did not significantly differ between patients investigated at 3T and patients assessed at 7T (P = .87 and P = .86, respectively).

Semiautomated Image Assessment
SH/Thk of the Primary Motor Cortex.
3T. SH/Thk was significantly higher in patients than in HS (mean, 0.11 ± 0.09; 95% CI, 0.07–0.14 for patients; and mean, 0.03 ± 0.02; 95% CI, 0.02–0.03 for HS; P < .001) (Fig 1A).
7T. SH/Thk was significantly higher in patients than in HS (mean, 0.25 ± 0.12; 95% CI, 0.20–0.31 for patients; and mean, 0.08 ± 0.04; 95% CI, 0.07–0.09 for HS; P < .001) (Fig 1B).
In the subgroup of HS who underwent MR imaging at both 3T and 7T, the SH/Thk of the deep layers of the M1 was significantly higher at 7T than at 3T.

Correlation of the SH/Thk with UMN Score.

3T. In patients with ALS, the SH/Thk of the deep layers of M1 significantly correlated with the UMN score of the corresponding limbs ($p = 0.57$, $P < .001$; 95% CI, 0.42–0.69) (Fig 2A).

7T. In patients with ALS, the SH/Thk measured in the M1 deep layers showed a significant correlation with the UMN score of the corresponding limbs ($p = 0.57$, $P < .001$; 95% CI, 0.39–0.71) (Fig 2B).

The measures of signal intensity and thickness of M1 subregions of HS overlapped those recorded in cortical subregions of patients corresponding to limbs with UMN scores of up to 3. On the contrary, the dataset of HS was more clearly distinguishable from that concerning cortical subregions of patients corresponding to limbs with UMN scores of $\geq 3$. This observation was seen with both 3T and 7T data (On-line Figs 2 and 3).

Diagnostic Performance of SH/Thk.

3T. ROC analysis of 3T data showed a good accuracy (area under the curve $= 0.89$) in differentiating pathologic from nonpathologic data, with a sensitivity of 0.81 and a specificity of 0.84 (Fig 3A).

7T. ROC analysis of 7T data showed an excellent accuracy (area under the curve $= 0.94$) as binary classifiers of data as pathologic or nonpathologic, with a sensitivity of 0.89 and a specificity of 0.86 (Fig 3B).

Visual Image Assessment

3T. Sensitivity, specificity, and diagnostic accuracy were, respectively, 0.56, 0.92, and 0.68 for the first reader and 0.63, 0.85, and 0.70 for the second reader (Table). The intrarater agreement was good for the first reader and very good for the second reader (0.74 and 0.85, respectively), and the interrater agreement was good for both reading sessions (0.80 and 0.70). The mean values of the total UMN scores of patients correctly diagnosed and patients misdiagnosed were, respectively, 11 and 5 for the first reader and 12 and 3 for the second reader.

7T. Sensitivity, specificity, and diagnostic accuracy were, respectively, 0.68, 0.89, and 0.78 for the first reader and 0.63, 1.00, and 0.81 for the second reader (Table). The intrarater agreement was very good for both readers (0.83 and 0.81), and the interrater agreement was good for both reading sessions (0.65 and 0.77). The mean values of the total UMN scores of patients correctly diagnosed and patients misdiagnosed were, respectively, 13 and 2 for the first reader and 13 and 4 for the second reader.

DISCUSSION

SH/Thk Changes in Patients with ALS and Correlation with Clinical UMN Impairment

The magnetic susceptibility of the deep layers of the M1, revealed by 3D multiecho T2*-weighted images and related to cortical content of nonheme iron, can be used to assess both their signal intensity and thickness. The distinction between the superficial and deep layers of the M1 is also often detectable on 3T images of patients and HS (On-line Fig 1 and Fig 4) and depends on the amount of myelinated fibers, deep layers being more myelinated than superficial ones. The gray-white matter junction in the M1 is sometimes hardly distinguishable in HS and in patients with mild UMN impairment because of the heavily myelinated deep layers of M1. On the contrary, in patients with ALS with moderate-to-severe UMN impairment, the deep layers of the M1 appear mark-
ity9,10,21 or UMN impairment.13 Our results agree with previous investigations such a correlation using scores of functional disability,13 findings at ultra-high-field MR imaging13 and confirm the link between the degree of focal cortical atrophy and hypointensity in the corresponding limbs: the greater the SH/Thk, the greater the clinical limb impairment. In other words, the cortical hypointensity ranged from being very pronounced to being seemingly indistinguishable from that of an unaffected cortex in patients with severe or light UMN impairment, respectively.

The hypothesis of a direct link between the location of cortical atrophy within the motor homunculus and clinical signs of UMN impairment was proposed 20 years ago on the basis of pathologic studies,17 but until now, only a few MR imaging studies investigated such a correlation using scores of functional disability8,10,21 or UMN impairment.13 Our results agree with previous findings at ultra-high-field MR imaging13 and confirm the link between the degree of focal cortical atrophy and hypointensity in the motor homunculus and the degree of signs of UMN degeneration in the corresponding limbs. Such results explain why neuroradiologists can correctly diagnose patients with ALS with moderate-to-severe UMN impairment, whereas patients with low or very low UMN burden are misdiagnosed. Moreover, the pattern of T2* hypointensity can be different among patients. In fact, according to the UMN burden, the extension of M1 signal hypointensity ranges from being localized to a small region of the M1 to bilaterally involving its full length, from the interhemispheric fissure to the lateral sulcus.

The marked T2* hypointensity of the deep layers of M1 in some patients with ALS compared with HS was demonstrated to be the expression of the greater magnetic susceptibility related to the abundant intracortical deposition of iron in the form of microglial ferritin.7,16 With the magnetic susceptibility having a positive and strict dependence on the magnetic field strength, T2* sensitivity to paramagnetic substances is lower at 3T than at 7T, as demonstrated in the subgroup of HS who underwent MR imaging examinations at both magnetic fields, thus reducing the ability to detect small collections of intracortical ferritin in patients with ALS with moderate-to-low UMN impairment. Such findings explain why the sensitivity, negative predictive value, and diagnostic accuracy in the visual imaging assessment were lower at 3T than at 7T. In line with this result, the performance of ROC analysis was also slightly better at 7T than at 3T. However, the accuracy in distinguishing pathologic and nonpathologic images using the semiautomated method was also high at 3T.

ROC data can be used in the assessment of subjects referred from neurologists with suspected motor neuron disease. In this scenario, the choice of a cutoff that improves sensitivity though affecting specificity can result in a more accurate M1 evaluation of patients. In fact, in our population, the use of the semiautomated method showed an increased sensitivity in evaluating M1 morphologic changes compared with visual imaging assessment. Therefore, besides visual image evaluation, after data collection from healthy subjects and the definition of a cutoff value, the use of the SH/Thk in clinical practice can contribute to the radiologic evaluation of images, mainly in patients with mild UMN burden, confirming morphologic changes that are only slightly visible on visual inspection. More interesting prospects are to increase the sensitivity of neuroimages in the detection of UMN pathology, allowing the identification of very small M1 changes and reducing the false-negative rate, and to estimate the UMN burden, thus supporting the clinical evaluation of patients and contributing to their phenotypic classification. To further support this hypothesis, one could confirm, in a larger sample, that the UMN score of patients correctly classified at visual assessment is higher than that of patients in whom visual and semiautomated assessments disagree in the classification of images.

One should note the following aspects of the present study. First, different from studies investigating cortical thickness, which assessed the cortex in a full-thickness fashion,8-10 here we evaluated only the deep layers of M1, known to be the location of pathologic changes. Second, contrary to cortical thickness and
voxel-based morphometry studies, after having collected some data from HS, the presented method could be applied not only at the group level but also at the single-subject level to estimate the UMN burden in each patient with pyramidal signs and symptoms. Third, until now, MR imaging studies investigated, individually, cortical atrophy or hypointensity. To the best of our knowledge, this is the first study assessing the combination of both parameters, thus improving the radiologic evaluation of M1. Compared with the assessment of a single parameter (SH or Thk), the use of the SH/Thk gives us 3 main advantages: 1) to assess simultaneously 2 different radiologic features related to the cortical neurodegeneration; 2) to find a radiologic tool that correlates with clinical UMN impairment; and 3) to reduce the false-positive ratio related to the increase in T2* hypointensity of M1 in the elderly. Furthermore, because the SH/Thk is semi-quantitative data measured in each single subject, it could be used for the phenotypic stratification of UMN involvement in longitudinal studies aiming at investigating the spread of cortical changes across time or the efficacy of neuronal and non-neuronal therapies.

Methodologic Considerations
T2* signal features within the cortex allow distinguishing superficial and deep layers, thus measuring only the thickness of M1 deep layers, where the atrophy seems to be localized. By contrast, sequences commonly used for cortical thickness measurements, such as inversion recovery T1-weighted sequences, provide better gray-white matter contrast but are used only for full-thickness cortical measurements and do not allow more targeted measurements.

For evaluation of cortical thinning and signal hypointensity of the deep layers of the primary motor cortex at 7T, 2D gradient recalled sequences with high in-plane resolution have previously been used. However, in clinical settings on high-field MR imaging systems (3T), 2D gradient recalled-echo sequences are not as efficient as 3D multiecho T2*-weighted techniques, which were proved to be the most sensitive in the detection of the low signal intensity in the precentral cortex of patients with ALS, due to the higher sensitivity of multiecho T2*-weighted imaging to iron in the form of ferritin; hence, their use was preferred in this study.

The sequence used in this study has often been used in the assessment of brain iron deposits; however, it has recently been demonstrated that techniques that rely on the signal phase, namely quantitative susceptibility mapping, are more accurate than transverse relaxation times in terms of iron quantification. Nevertheless, the production of susceptibility maps requires particular acquisition settings: in most quantitative susceptibility mapping implementations such as ours, the MR imaging system is programmed to provide the complex MR imaging data divided into real and imaginary parts, which are of no radiologic use. In fact, quantitative susceptibility mapping requires additional scan time in addition to the conventional 3D T2* multiecho sequence that is included in the clinical protocol. Quantitative susceptibility mapping also requires time-consuming postprocessing. A limitation of this study was that patients with ALS investigated at 3T and 7T were not the same; hence, a direct comparison of the diagnostic accuracies obtained with the systems working at different magnetic field strengths would be unfair. However, on the basis of the absence of significant differences in total and mean UMN scores between the 2 groups of patients and on the significant correlation of signal intensity and thickness with UMN scores, we could hypothesize that MR imaging morphologic changes of M1 are comparable between groups, and a cautious comparison of MR diagnostic accuracy between different magnetic fields could be made. A further limitation is the number of subjects enrolled. Considering that ALS is a rare disease, the population we investigated is quite large, but the potential clinical applications described above need to be tested on a larger group of subjects or, at least, on a different cohort of patients to confirm the feasibility and reproducibility of results.

CONCLUSIONS
The SH/Thk of the deep layers of the M1 measured with a semi-automated method at 3T seems to be a radiologic marker of upper motor neuron burden in patients with ALS, though with less accuracy than at 7T.

Despite the heterogeneous magnitude of the UMN burden of patients, the combination of visual imaging assessment and the use of a semiautomated algorithm able to assess both thickness and T2* hypointensity of the deep layers of M1 could increase the sensitivity in evaluating images of patients referred with suspected motor neuron disease.

Disclosures: Mirco Cosottini—RELATED: Grant: Italian Ministry of Health. Comments: This study has received funding (RF-2009-154628) to Mirco Cosottini as Principal Investigator by the Italian Ministry of Health (www.salute.gov.it) and the Health Service of Tuscany (www.regione.toscana.it/sst). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES
their phenotypes: a neuroimaging study. J Neurol 2013;260:2856–64 CrossRef Medline