Sonographic and 3T-MRI-based evaluation of the tongue in ALS

Nathalie Hensiek¹, Frank Schreiber¹, Frank Schreiber¹, Thomas Wimmer¹, Jörn Kaufmann¹, Judith Machts¹,², Laura Fahnsbusch³, Cornelia Garz¹,², Susanne Vogt¹,², Johannes Prudlo¹,², Reinhard Dengler⁵, Susanne Petri⁵, Peter J. Nestor⁶, Stefan Vielhaber¹,⁷,¹, Stefanie Schreiber¹,²,⁷,¹

¹Department of Neurology, Otto-von-Guericke University, Magdeburg, Germany
²German Center for Neurodegenerative Diseases (DZNE) within the Helmholtz Association, Magdeburg, Germany
³Department of Neurology, Rostock University Medical Center, Germany
⁴German Center for Neurodegenerative Diseases (DNZE) within the Helmholtz Association, Rostock, Germany
⁵Department of Neurology, Hannover Medical School, Hannover, Germany
⁶Queensland Brain Institute, University of Queensland, Brisbane 4072, Australia
⁷Center for behavioral brain sciences (CBBS), Magdeburg, Germany

ARTICLE INFO

Keywords:
Amyotrophic lateral sclerosis
Tongue
Ultrasound
MRI
Biomarker
Intensity
Prognosis

ABSTRACT

A few systematic imaging studies employing ultrasound (HRUS) and magnetic resonance imaging (MRI) have suggested tongue measures to aid in diagnosis of amyotrophic lateral sclerosis (ALS). The relationship between structural tongue alterations and the ALS patients’ bulbar and overall motor function has not yet been elucidated. We here thus aimed to understand how in-vivo tongue alterations relate to motor function and motor function evolution over time in ALS.

Our study included 206 ALS patients and 104 age- and sex-matched controls that underwent HRUS and 3T MRI of the tongue at baseline. Sonographic measures comprised coronal tongue echointensity, area, height, width and height/width ratio, while MRI measures comprised sagittal T1 intensity, tongue area, position and shape. Imaging-derived markers were related to baseline and longitudinal bulbar and overall motor function.

Baseline T1 intensity was lower in ALS patients with more severe bulbar involvement at baseline. Smaller baseline coronal (HRUS) and sagittal (MRI) tongue area, smaller coronal height (HRUS) and width (HRUS) as well as more rounded sagittal tongue shape predicted more rapid functional impairment - not only of bulbar, but also of overall motor function - in ALS.

Our results suggest that in-vivo sonography und MRI tongue measures could aid as biomarkers to reflect bulbar and motor function impairment.

1. Introduction

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disorder with considerable variation in presenting features and progression (Al Chalabi et al., 2016). Heterogeneity of clinical presentation may stem from several factors such as functional reserve or compensating capacities, genetic susceptibility and environmental factors (Al Chalabi and Hardiman, 2013). Prognostic biomarkers of ALS are needed, particularly of bulbar involvement, which is, of course, not the only, but a key determinant of long-term prognosis and survival in this disorder (Rosenbohm et al., 2017). The establishment of a non-invasive marker of bulbar involvement may be of value in prevention aspiration pneumonia and managing nutrition, feeding and rehabilitation. Currently, there is no such marker or any technique applicable in the clinical setting that would have the potential to assess the progression of bulbar involvement at the individual patient level.

Several case reports and some systematic imaging studies employing ultrasound and magnetic resonance imaging (MRI) have suggested

Abbreviations: ALS, amyotrophic lateral sclerosis; ALSFRS-R, revised ALS functional rating scale; CON, controls; FLIRT, FMRIB’s Linear Image Registration Tool; FMRIBI, functional magnetic resonance imaging of the brain; FSL, FMRIB’s Software Library; HRUS, high-resolution ultrasound; ICC, intraclass correlation coefficient; LMND, lower motor neuron dominant; MPRAIG, magnetization prepared rapid gradient echo; MRI, magnetic resonance imaging; NIV, non-invasive ventilation; PLs, primary lateral sclerosis; ROI, region of interest; RR, reference ROI; SOD1, superoxide dismutase – 1; SD, standard deviation; TIV, tracheostomy – invasive ventilation; TR, tongue ROI; UMND, upper motor neuron dominant

Corresponding author.
E-mail address: stefanie.schreiber@med.ovgu.de (S. Schreiber).

1Equally contributed.

https://doi.org/10.1016/j.nicl.2020.102233
Received 4 November 2019; Received in revised form 18 February 2020; Accepted 29 February 2020
Available online 02 March 2020
2213-1582/ © 2020 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/BY-NC-ND/4.0/).
structural tongue measures to aid in the differentiation between ALS and patients with other neurological conditions or healthy controls (Cha and Patten, 1989; Fox and Cohen, 2012; Lee et al., 2018; Misawa et al., 2011; Nakamori et al., 2016). The relationship of quantitative tongue measurements and ALS progression is unknown.

In consideration of these uncertainties, we conducted a large cross-sectional and longitudinal study measuring several tongue features using both, high-resolution ultrasound (HRUS) and 3T MRI. Both techniques provide different advantages and disadvantages. HRUS is widely available and suitable for serial studies even in advanced stages. However, user-dependence of HRUS may limit its reliability (Nijholt et al., 2017). MRI, in contrast, provides more standardized results and hence reliability, but is less well suited to serial studies due to its cost, restricted availability and requirement for patient cooperation (Jenkins et al., 2013). This study was designed to compare the usefulness of MRI and HRUS in predicting functional decline in ALS.

2. Material & methods

2.1. Sample

The cohort comprised n = 206 ALS patients who underwent MRI, HRUS or both. As part of an ongoing longitudinal MRI study of the brain, which was initiated in 2011 to investigate the ALS patients’ cognition and their underlying functional and structural brain alterations, patients were recruited at the Departments of Neurology at the Otto-von-Guericke University, Magdeburg; Hannover Medical School; and the University Medical Center, Rostock, Germany. ALS was diagnosed by two out of three experienced neurologists (JP, SP, SV) and patients suffered from definite, probable, possible or suspected ALS according to the El Escorial criteria and its revisions (Brooks et al., 2000; Brooks, 1994; Carvalho et al., 2008; Carvalho and Swash, 2009). Exclusion criteria were concurrent neuromuscular or neurological disorders of the CNS or PNS or motor neuron diseases other than ALS as well as contraindications against MRI. N = 175 (85%) patients underwent 3T MRI of the brain (also mapping the tongue) between 04/2011 and 02/2017. N = 48 (23%) patients underwent HRUS examination of the tongue between 09/2015 and 12/2016, and they were recruited at the Departments of Neurology Magdeburg and Hannover, only (Table 1). There was an overlap of n = 17 (8%) ALS patients who underwent both MRI and HRUS.

Additionally, n = 104 age- and sex-matched healthy control subjects were recruited at the Departments of Neurology at the Otto-von-Guericke University, Magdeburg and the University Medical Center, Rostock, Germany via public announcement. None of the control participants suffered from any neurodegenerative, neuroinflammatory or neuromuscular disorders nor did they display any specific abnormalities on the neurological exam. They were all cognitively normal. N = 87 (84%) of them underwent MRI between 04/2011 and 02/2017 at the Departments of Neurology at the Otto-von-Guericke University, Magdeburg and the University Medical Center, Rostock, Germany; and n = 17 (16%) underwent HRUS during 09/2015 and 12/2016 at the Department of Neurology at the Otto-von-Guericke University, Magdeburg (Table 1).

The discrepancies of sample size between both imaging cohorts originate from the different recruitment periods. The recruitment of HRUS patients or controls started in 2015 and took place over a comparable short period of 15 months; while the recruitment of MRI patients or controls was part of a longer-lasting study that already started in 2011 with a recruitment period of 70 months (please see above). The small overlap between both imaging cohorts not only results from the short overlap period during recruitment, but also from the fact that MRI took place at 2 centers (Magdeburg, Rostock), while HRUS was performed at one center only (Magdeburg).

The local ethics committees approved the study and all subjects gave informed consent (75/11, 16/17).

2.2. Clinical data

The sonographic or MRI examination was considered as a baseline. At baseline disease severity was assessed using the revised ALS functional rating scale (ALSFRS-R) (Cedarbaum et al., 1999). Median (range) disease duration from disease onset to baseline was 17 (0.7–272) months, and from diagnosis to baseline it was 5 (0–122) months (Table 2). N = 125 ALS patients underwent at least one, and n = 78 patients underwent at least two follow-up ALSFRS-R measures. N = 81 patients were lost to longitudinal follow-up due to death (n = 20, 25%), switching to another medical center (n = 44, 54%) or undetermined reasons (n = 17, 21%). Time spans between each follow-up was approximately 3 months. Median (range) time span between baseline and the last available ALSFRS-R was 8.5 (0.2–63) months. Statistical analysis comprised the composite score for bulbar, respiratory, fine and gross motor function (maximum score 48 points) and the bulbar sub-score (maximum score 12 points).

Clinical phenotype distribution at baseline comprised classic ALS (n = 134, 65%), lower motor neuron dominant ALS (LMND ALS, n = 41, 20%) and upper motor neuron dominant ALS (UMND ALS, n = 21, 10%), classified in accordance with previously specified operational definitions (Schreiber et al., 2015). We also included patients with primary lateral sclerosis (PLS, n = 10, 5%), a syndrome of pure upper motor neuron involvement that has been considered part of the motor neuron disease spectrum (Al Chalabi et al., 2016; Finegan et al., 2019) (Table 2). In detail, at the time of tongue imaging, a variable combination of LMN and UMN signs in the bulbar, cervical and lumbo-sacral regions were found in those designated as classic ALS who, in turn, fulfilled the El Escorial criteria of definite or probable ALS. All patients with LMND ALS had clinical and electrophysiological evidence of sporadic progressive pure LMN involvement in one or more regions without clinical signs of UMN dysfunction, which must have been the predominant finding for at least 12 months after symptom onset. UMND ALS patients, in contrast, had either no LMN signs, or, if present: (1) they were restricted to only one neuroaxis level (bulbar, cervical, or lumbo-sacral); and (2) electromyography (EMG) abnormalities were limited to sparse fibrillation potentials/positive sharp waves or minor

### Table 1

Demographics of the sample under consideration.

| Total cohort | MRI cohort | HRUS cohort |
|--------------|------------|-------------|
| ALS n = 206  | CON n = 104 | n = 175     |
| CON n = 87   |            | n = 48      |
| p            |            | n = 17      |
| Age in years* |           | 63 (32;83) | 61 (33;83) |
|              |           | 62 (32;83) | 62 (33;83) |
| Male sex, n (%)*** |   | 129 (63) | 64 (62)     |
| Height in cm*, mean [SD] | | 172 [9] | 173 [10] |
| Weight in kg*, mean [SD] | | 74 [14] | 80 [13] |

Unless otherwise reported median (range) is given. For group comparisons independent-samples t-test*, Mann-Whitney U test** or χ² test*** was conducted. P-values < 0.05 were deemed statistically significant.

ALS, amyotrophic lateral sclerosis; CON, controls; HRUS, high-resolution ultrasound; MRI, magnetic resonance imaging.
enlargement of motor unit potentials in one or at most two muscles for at least 12 months after symptom onset (Schreiber et al., 2015). The enlargement of motor unit potentials in one or at most two muscles for motor neuron dominant.

Concerning the ALSFRS-R total (bulbar sub-score) decline baseline minus the last available score in terms of longitudinal follow-up. Unless otherwise reported median (range) is given. Decline of ALSFRS-R total and ALSFRS-R bulbar* from advanced disease stages in that they could not sit for a longer time or had swallowing disturbances with retention of saliva. All images were imported to ImageJ Version 1.5 (National Institute of Health, USA). The latter was used to perform a visual quality check and consequent manual delineation of the whole tongue area, which was conducted by one investigator (NH), blinded to diagnostic group. The tongue area was defined through its anatomical boundaries that are (a) the hyperechoic tongue root (top), (b) the hyperechoic tongue dorsum (bottom) and (c) the transition from hyperechoic muscular tongue tissue to the hypoechoic tissue of the oral cavity (lateral margin) (Fig. 1). Delineated tongue area was then taken as a region of interest (ROI) to calculate the quantitative features of the tongue. Mean echointensity was defined as the gray level derived from each ROI pixel; coronal area was calculated by multiplying the total number of tongue ROI pixels with the area per pixel; coronal height and width was calculated by multiplying the distance in pixels within the respective plane direction with the corresponding pixel dimension, and the height/width ratio was calculated as well (Fig. 1). The parameters derived from one session's multiple images were consequently averaged.

To assess the measures' reliability, images of n = 10 randomly selected subjects were delineated 3 times by the same investigator (NH) and once by a second investigator (LF).

### 3T MRI of the tongue

At both imaging sites (Magdeburg and Rostock), sagittal cerebral 3D-MPRAGE (magnetization-prepared rapid gradient-echo) scans were acquired at a Siemens MAGNETOM Verio 3 T MRI scanner (Siemens, Erlangen, Germany) using a 32-channel head coil, and applying the same pulse sequence setting (echo time 4.82 ms, repetition time 2500 ms, inversion time 1100 ms, flip angle 7°, voxel size 1 × 1 × 1 mm³, matrix 256 × 256 × 192), respectively. To account for varying head sizes and alignments, the bias-corrected scans were registered to a centered mean template applying an affine (12 degrees of freedom) linear registration using the linear image registration tool (FLIRT) provided as a part of the FMRIB’s Software Library (FSL).
version 6.0 (Jenkinson et al., 2002). Tongue analysis was conducted in a coplanar manner using one section per subject. DICOM images were imported to Mango viewer version 4.0.1 (Research Imaging Institute, University of Texas Health Science Center, USA) to again manually delineate the whole tongue area by the same blinded investigator (NH). In line with the work by Cha and Patten (Cha and Patten, 1989) we concentrated on the midsagittal slice of the MRI volume as this is a distinct slice found with relative ease in clinical practice, irrespective of tongue dimension. Each delineation was performed in the same midsagittal cross-section of the aligned scans to obtain two ROIs, the mid-sagittal tongue ROI (TR) as well as an approximately trapezoidal ROI serving as a reference for the individual samples (Fig. 2). The TR was defined through its anatomical boundaries that are (a) the hyperintense epithelium with its papilla (apex and upper surface), (b) the hyperintense lingual tonsil (back) and (c) the transition from the iso-intense tongue muscle to the hypointense mouth tissue (bottom). The reference ROI (RR) was used to account for the varying median extent of the individuals’ oral cavity. The RR was also defined through its anatomical boundaries starting from the visible tip of the upper front teeth to their inner edge along the palate to the back of the pharynx (Fig. 2). The lower ROI boundary was drawn parallel to the palate interconnecting the tip of the teeth with the back of the pharynx (Fig. 2). To again assess the measures’ reliability images of $n = 10$ randomly chosen subjects were delineated three times by the same investigator (NH) and once by a second investigator (LF). A center line was fitted to the points enclosed in RR, that minimizes the distance to these points in a least square sense (Fig. 2). The points at which the lines perpendicular to the center line intersected with the tooth tip and the back of the pharynx were marked and used to normalize the positions along the center line, with a value of 0 located at the pharynx and 1 at the teeth. In order to assess the tongue size, the mid-sagittal tongue area was evaluated by multiplying the number of pixels contained in TR with the area covered by a single pixel. The median component of the tongue position was assessed by the center point onto the center line (Fig. 2). Applying a solver based upon the Khachiyan algorithm (Todd and Yildirim, 2007), the minimum volume enclosing ellipsoid was estimated for the points contained in TR (Fig. 2). The shape of TR was quantified by the resulting ellipsoids’ numerical eccentricity $a$, and minor axis length $b$ (Fig. 2). The tongue ROI was moreover used to calculate the tongue’s mean unnormalized T1 intensity. All calculations, both for HRUS and MRI measures, were carried out using Mathwork’s MATLAB R2016a, SPM12 and FMRIB Software Library v5 (Jenkinson et al., 2012).

2.5. Statistics

Statistical analyses were conducted using IBM SPSS statistics (Version 24.0). Averages from all HRUS images per patient were taken for each HRUS tongue feature. Intra- and inter-rater reliability was calculated using the intraclass correlation coefficient (ICC).

We first aimed to test for the existence of any relationship between sonographic or MRI data and the cohort’s demographics (age, sex, height, weight), that would have been taken into account for further analysis, as thus performed bivariate correlations, independent-samples t-tests or Mann-Whitney U tests. Bivariate correlations have further been used to assess any relationship between sonographic or MRI data and the ALS patients’ clinical variables (disease duration, baseline ALSFRS-R and its bulbar sub-score). P-values < 0.05 were deemed statistically significant.

Group comparisons were performed using general linear models with each tongue measure considered as dependent variable and group as independent variable. Subgroup comparisons thereby focused on bulbar and respiratory severely affected patients who may present more distinct tongue alterations (Cha and Patten, 1989; Fox and Cohen, 2012; Matsuda et al., 2016; McKee et al., 2013; Nakayama et al., 2017). Group and subgroup comparisons were thus conducted between the following: (i) ALS patients vs. controls, (ii) bulbar-onset patients vs. limb-onset patients vs. controls, (iii) patients requiring assisted ventilation vs. patients not requiring assisted ventilation, (iv) patients with vs. without gastrostomy. Covariate adjustment was conducted where necessary in order to account for any correlations between the tongue measures and the patients’ demographics (see Results). As ALS patients were considered within four different constellations Bonferroni-adjusted $p$-values of $<0.05/4 = 0.013$ were deemed statistically significant.

Furthermore, the potential of imaging-based tongue variables to predict the disease course in ALS patients has so far remained unexplored. As the longitudinal progression of disease was recorded clinically by the ALSFRS-R, random intercept mixed effects linear models with ALSFRS-R total or its bulbar sub-score (dependent variable), various tongue measures (each included as main effect), time in months (main effect) and time × tongue measures interaction effects were conducted. Each tongue measure was median split and the resulting binary dummy-coded variable (performing high (above median) vs. performing low (below median) on each tongue measure, coded as 1 vs. 0) was included in random intercept mixed effects linear models. Estimates $e$ (and the respective confidence intervals) of time interaction effects displayed the measures of interest indicating absolute ALSFRS-R (total or bulbar sub-score) differences between median split levels for each tongue measure over time. Covariate adjustment was performed if necessary (see Results). As there were two outcome measures for each model, Bonferroni-adjusted $p$-values $<0.05/2 = 0.025$ were deemed statistically significant.

In consideration that PLS is a distinct entity, which is part of the
motor neuron disease spectrum, but has a completely different disease course, all analysis were repeated after excluding the PLS cases from the cohort. None of the results changed after exclusion of the PLS patients.

Sample size calculations for a hypothetical, randomized, controlled trial with a 12-month observation period were conducted using a 2-sample t-test, assuming equal group means, a 50% treatment effect on the decline of the ALSFRS-R score, a linear ALSFRS-R decline, a 2-sided significance of 0.05, and a power of 0.8. Analysis was performed using nQuery winter 2019 release (ver 8.5.0) software (Statistical Solutions Ltd, Cork, Ireland).

3. Results

3.1. Demographics and clinical data of the cohort

Demographics of the whole cohort under consideration are given in Table 1. There were no group differences between ALS subjects of the HRUS or MRI cohort and controls with respect to age, sex, height and weight.

Table 2 demonstrates the patients’ clinical data. Mean ALSFRS-R total score was significantly lower and prevalence of gastrostomy, as well as prevalence of patients requiring ventilation was significantly higher in ALS Patients who underwent tongue HRUS compared to those who underwent 3T MRI. There were, however, no differences between the HRUS and MRI ALS sample with respect to their remaining clinical data.

3.2. Tongue features

For each tongue measure intra- and inter-rater reliability was good to excellent with ICC values ranging from 0.8 to 1.0 (for sonography: echointensity 0.9 and 0.9, area 0.8 and 0.9, height 0.8 and 0.8, width 0.8 and 0.9, height/width ratio 0.8 and 0.8; for MRI: T1 intensity 1.0 and 1.0, area 1.0 and 0.7, position 1.0 and 0.8, shape 0.9 and 0.9).

Lower HRUS echointensity was related to greater height ($r = −0.4$, $p = 0.001$) and male sex ($t(63) = 4.0$, $p < 0.001$); larger sonographic area, height and width was related to greater weight ($r = 0.3$, $p = 0.007$; $r = 0.3$, $p = 0.03$; $r = 0.3$, $p = 0.007$). Larger MRI tongue area was correlated with greater weight ($r = 0.3$, $p = 0.002$) and male sex ($Z = −3.3$, $p = 0.001$); sagittal tongue shape was related to age in that older age was associated with a more rounded shaped tongue ($r = 0.2$, $p = 0.009$) (for the detailed results of the correlation analysis see Supplemental Table 1).

Lower T1 intensity ($r = 0.2$, $p = 0.02$) was related to a lower ALSFRS-R bulbar sub-score at baseline (Fig. 3). There was no association between any other HRUS or MRI measures and disease duration, ALSFRS-R total or its bulbar sub-score (baseline measures).

3.3. Group- and subgroup- comparisons

Bulbar- compared to limb-onset patients displayed a significantly lower T1 intensity of the tongue ($p < 0.001$, $F(2) = 8.5$) (Fig. 3). The remaining tongue measures, however, did not differ between bulbar- and limb-onset patients.

Further group comparisons did not reveal any differences between (i) ALS and controls, (ii) patients requiring assisted ventilation vs. patients not requiring assisted ventilation, (iii) patients with vs. without gastrostomy, regarding any tongue measure derived from HRUS or MRI (for the detailed results of the group comparisons see Supplemental Table 2).

Results remained unchanged after excluding patients with assisted ventilation or without gastrostomy or with UMND ALS or PLS phenotypes.

3.4. Longitudinal analysis

Longitudinal analysis using mixed effects linear models revealed significant time interaction effects on ALSFRS-R total and its bulbar sub-score of sonographic tongue area ($e = 0.5$ [CI 0.3; 0.6], $p ≤ 0.001$; $e = 0.2$ [CI 0.1; 0.3], $p ≤ 0.001$), height ($e = 0.3$ [CI 0.1; 0.4], $p = 0.002$; $e = 0.1$ [CI 0.1; 0.2], $p ≤ 0.001$) and width ($e = 0.5$ [CI 0.3; 0.6], $p ≤ 0.001$; $e = 0.2$ [CI 0.1; 0.2], $p = 0.001$) (Supplemental Table 3). In summary, patients with larger baseline sonographic measures (area, height, width) compared to those displaying smaller baseline sonographic measures declined at a monthly rate of 0.3 to 0.5 points (for ALSFRS-R total score) or 0.1 to 0.2 points (for ALSFRS-R bulbar sub-score) slower (Fig. 4). Additionally, we found significant time interaction effects of the MRI-based tongue area on ALSFRS-R bulbar sub-score ($e = 0.05$ [CI 0.1; 0.08], $p = 0.004$) and of the MRI-based tongue shape on the ALSFRS-R total score ($e = 0.1$ [CI 0.02; 0.2], $p = 0.015$). This implies, that patients with larger tongue area or more ellipsoidal tongue shape at baseline MRI declined at a monthly rate of 0.04 (for ALSFRS-R bulbar sub-score) or 0.1 points (for ALSFRS-R total score) slower when compared to those with smaller tongue area or more rounded tongue shape.

In translating these findings to the context of a randomized, controlled clinical trial with a 12-month observation period, a sample size calculation based on the HRUS cohort was carried out to have an 80% chance ($α = 0.05$) of detecting a 50% therapy effect on ALSFRS-R decline, which corresponds to a decreased mean decline from −12.9
per year to −6.5 points per year. Based on the sonographic cohort, the necessary size of each trial arm is estimated as 34 patients. In case possible candidate were also selected to additionally show submedian values for sonographic tongue area, tongue height or tongue width, the necessary size of each trial arm would decrease to 29 (~15%), 32 (~6%) or 26 (~24%), respectively (Supplemental Table 4).

The same evaluation for the MRI cohort with an annualized mean ALSFRS-R decline of −9.0 resulted in a necessary size of each trial arm of 44 patients, which could be reduced to 41 (~7%) or 40 (~9%) by only including patients with submedian MRI-based tongue area or tongue shape, respectively (Supplemental Table 5).

4. Discussion

This large study suggests that lower T1 intensity of the tongue could be a marker of more severe bulbar involvement in ALS. Smaller baseline coronal (HRUS) and sagittal tongue area (MRI), smaller coronal height (HRUS) and width (HRUS) as well as more rounded sagittal tongue shape might have the potential to predicate more rapid functional impairment - not only of bulbar, but also of overall motor function - in ALS. Future studies have to confirm the potential of in-vivo HRUS und MRI tongue features to potentially serve as biomarkers to reflect motor function impairment.

Former tongue studies in ALS were limited by low resolution imaging, insufficient data quantification and small sample sizes. In addition to overcoming all these shortcomings, application of an advanced analysis technique allowed derivation of several novel tongue measures. In particular, qualitative HRUS- and MRI-based tongue measures such as echointensity and T1 intensity, position and shape have been rigorously quantified using a novel approach whereas former studies relied only on visual or semi-quantitative evaluation.

More severe bulbar involvement was related to lower tongue T1 intensity. The skeletal muscle in autopsied ALS patients with long-lasting disease (up to 317 months) is characterized by atrophic fibers, increase of connective tissue and fatty replacement (Ahmadi et al., 2010; Borgia et al., 2017; Dupuis and Echaniz-Laguna, 2010; Trias et al., 2018). These chronic changes resulting from longstanding denervation (May et al., 2000) have been proposed to give rise to increased T1 intensity on MRI (Theodorou et al., 2012), but correlation studies between histopathology and MRI in ALS are absent. There is, indeed, only one study with less than 40 patients with longstanding ALS and one case report, that have specifically taken into account T1 alterations of the tongue muscle. Performing visual analysis, the authors described higher T1 intensities of the tongue (Cha and Patten, 1989; Fox and Cohen, 2012), but replication of their results within larger independent studies is still missing. Low T1 intensity, in contrast, has been proposed to represent edema or inflammatory cellular infiltration indicating subacute denervation (Theodorou et al., 2012; Zhang et al., 2013), and is thought to be apparent at earlier ALS disease stages. Consistent with this interpretation, autopsy studies display cellular (e.g. neutrophils, lymphocytes) skeletal muscle inflammation in ALS patients (Al-Sarraj et al., 2014; Trias et al., 2018), especially in those with comparatively shorter disease duration (up to 55 months). The association between bulbar involvement and lower T1 intensity in our study could presumably mirror the early disease stage of our ALS sample, reflected by a short median baseline disease duration of 17 months and relatively preserved bulbar function (median baseline ALSFRS-R bulbar sub-score of 11 points). However, our data and the provided interpretation of these findings have to be considered with caution. First, effect size of the relationship between bulbar involvement and lower T1 intensity was small. Further, T1 sequences are not well suited to detect edema or inflammation, but have been rather designed to assess atrophy, fibrosis and fatty degeneration. Moreover, considering disease stage to play a role, one would have expected to find any relationship between T1 intensity and disease duration or any group differences between less advanced or more advanced disease stages (as indicated by the need for ventilation and/or gastrostomy). But this was not the case in our ALS sample. Large studies also taking into account the evolution of T1 intensity of the tongue are therefore needed to replicate our findings.

Several baseline tongue features (HRUS: smaller coronal area, height and width; MRI: smaller area and rounded shape) seemed to predict worse long-term motor and bulbar function as assessed using the ALSFRS-R over a median (range) follow-up time of 8.5 (0.23 - 63) months. Concerning ALS patients with the above mentioned tongue features, there was only a slight difference in monthly decline of 0.5 to 0.1 points (ALSFRS-R total score) or 0.2 to 0.1 points (ALSFRS-R bulbar sub-score) compared to ALS patients with the opposite tongue features, which limits the practical relevance of that finding. This is underscored by the variability of the tongue measures, indicated by their relationship to several demographical data (such as age or sex), allowing for the detectability of only extreme changes. The low magnitude of functional decline may not be a problem in a sufficiently large group (like our sample), but it might be not useful at the individual level. Our longitudinal results further indicate that there would be a difference between ALS and controls at some point later in the illness. It is thus no surprise that cohorts with more advanced disease would be more likely to show positive results, but this is arguably of limited use because diagnostic markers are most needed in early disease stages where the clinical profile is typically more ambiguous. However, the results of the power analyses indicate that the presented variables might prove useful if employed as additional inclusion parameters in therapeutic trials, as they would allow focusing on patients with e.g. smaller sonographic tongue area or tongue width which are more likely to show a steep ALSFRS-R decline. This might make it possible to either reduce the size

---

**Fig. 4.** Longitudinal ALSFRS-R total score as a function of baseline sonographic tongue measures. Graphs demonstrate significant time interaction effects of baseline sonographic tongue area (a), width (b) and height (c) on longitudinal ALSFRS-R total score. ALS patients revealing larger baseline sonographic measures (black) compared to patients displaying smaller baseline sonographic measures (red) declined slower when considering the longitudinal ALSFRS-R total score. ALSFRS-R, revised ALS functional rating scale.
of trial arms or increase the power of a trial with the same number patients.

Chronic invasive ventilation (up to 87 months) has been related to tongue enlargement in ALS (Matsuda et al., 2016; McKee et al., 2013; Nakayama et al., 2017), and so could mask tongue atrophy in ALS. None of the patients in our sample required invasive ventilation, however, and there were only 10% receiving non-invasive ventilation. It has been unclear whether non-invasive ventilation results in any structural tongue alterations as well. Our results comparing patients with non-invasive to those without ventilation suggests there is no relationship between non-invasive ventilation and structural tongue features.

Likewise, from a clinical perspective pronounced/sole upper motor neuron involvement with a spastic tongue muscle in UMND ALS and PLS could have also masked tongue atrophy. Furthermore, macroglossia has been reported an early feature of PLS (Chang and Kwak, 2019). On the other hand, tongue muscle atrophy and intensity alterations have been proposed to mainly result from predominant lower motor neuron involvement (Cha and Patten, 1989). Therefore inclusion of UMND ALS and PLS phenotypes may have led to concerns of masked tongue atrophy in ALS, but the data did not support this hypothesis. Overall, incident tongue enlargement possibly resulting from ventilation or spastic phenotypes was thus not masking tongue atrophy and therefore not the reason behind absent tongue group differences between ALS and controls. In this large sample there was a robust relationship between various tongue measures and weight and sex, replicating previous findings (Cha and Patten, 1989; Matsuda et al., 2016; Nakamori et al., 2016). These findings seem to be a better explanation for why the majority of tongue measures were not useful diagnostically - i.e. with so much variability in the normal population, one would require extreme changes (which were not anticipated at early disease stages) to identify a disease-specific marker.

Tongue features did not differ between ALS patients with or without gastrostomy, but with only n = 11 gastrostomy participants, this may simply reflect lack of power. This might be explained by the fact that gastrostomy is considered and placed in any case patients suffering from dysphagia, even at early ALS disease stages without severe bulbar dysfunction (Nunes et al., 2016). Early gastrostomy placement thereby not only reduces the procedure risks but further relates to a better prognosis in ALS (Conde et al., 2018). The participating centers of that study applied an early gastrostomy as recommended, which is reflected by a comparably long mean (SD) duration of having gastrostomy of 8.7 (10.1) months in contrast to the relatively short median disease duration of 17 months.

Both HRUS and MRI are non-invasive investigations. As already stated, HRUS is, however, more feasible in advanced disease stages, in that patients do not need to lie flat as is required during MRI scanning. The MRI-based method, however, has the advantage that it can be acquired as part of the standard cranial MRI that forms part of the routine diagnostic work-up in ALS. Regarding their potential to serve as prognostic markers in ALS, both imaging techniques the longitudinal findings of this study indicate that both HRUS or MRI tongue measures have potential to monitor patients’ long-term motor/bulbar function.

There are several study limitations. As already mentioned, effect sizes were small and associations between clinical and imaging measures were little robust. Moreover, albeit performing several group and subgroup analysis, we have restricted the corrections for multiple testing. Further, despite large, our patient cohort was still quite heterogeneous with small, diverse and unequal clinical subgroups, especially for the distinct imaging modalities. Heterogeneity and variability was also evident in the patients’ longitudinal evaluations with little standardized time intervals between the clinical visits and several cases that have been lost-to-follow up. These are, however, very common issues when performing a clinical study with patients suffering from such a devastating neurodegenerative disease like ALS. Analysis was further restricted to static imaging and offline evaluations, which do not represent sufficient information about the dynamic data reproducibility in the clinical setting. Future studies should further take into account dynamic tongue measures (e.g. fasciculations) as well as bulbar muscle groups beyond the tongue. We further did not consider any possible influence of tongue fasciculations on our static tongue measures. Additionally, MRI analysis took account of T1 images only, and follow-up studies should consider further MRI sequences (e.g. T2 images, diffusion tensor imaging) to bring in other important information regarding tissue properties. Future studies should also focus on 3-dimensional analysis of the tongue in order to detail possibly irregular patterns of tongue atrophy in ALS. One may also consider applying high-frequency probes to provide more details on the sonographic fine muscular structure of the tongue.

In summary, MRI-based T1 tongue intensity appears to be a marker of current bulbar involvement in ALS, whereas larger sonographic tongue area, width and height as well as larger MRI-based tongue area and ellipsoidal tongue shape at baseline could be markers of long-term motor function in ALS. These results highlight the potential role of in-vivo tongue imaging to serve as a tool for current bulbar involvement as well as for motor function evolution over time in ALS, making these imaging approaches potentially useful measures for future clinical or therapeutic trials.

Ethical publication statement

We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Financial disclosure

None of the authors have any disclosures to declare.

CRediT authorship contribution statement

Nathalie Hensiek: Investigation, Conceptualization, Writing - original draft, Visualization, Writing - review & editing. Frank Schreiber: Conceptualization, Methodology, Software, Validation, Formal analysis, Data curation, Writing - review & editing. Thomas Wimmer: Investigation, Validation, Data curation. Jörn Kaufmann: Resources, Data curation, Methodology, Software, Conceptualization. Judith Machts: Data curation, Investigation, Resources. Laura Fahlbusch: Investigation, Validation. Cornelia Garz: Investigation, Resources, Data curation. Susanne Vogt: Data curation, Investigation, Resources. Johannes Prudlo: Investigation, Resources, Data curation, Supervision. Reinhard Dengler: Investigation, Resources, Data curation, Supervision. Susanne Petri: Investigation, Resources, Data curation, Supervision. Peter J. Nestor: Conceptualization, Investigation, Resources, Data curation, Supervision. Stefanie Schreiber: Conceptualization, Resources, Writing - review & editing, Supervision, Project administration.

Declaration of Competing Interests

None of the authors have potential conflicts of interest to be disclosed.

Acknowledgements

We thank Anne-Katrin Baum, Department of Neurology, Otto-von-
Guericke University, Magdeburg, Germany, for technical assistance, and Christa Sobetzkó, Department of Neurology, Otto-von-Guericke University, Magdeburg, Germany, for data collection.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jic.2020.102233.

References

Ahmadi, M., Liu, J.-X., Braintstrom, T., Andersen, P.M., Stål, P., Pedrosa-Domelöff, F., 2010. Human extracellular muscle in ALS. Invest. Ophthalmo. Vis. Sci. 51 (7), 3494–3501. https://doi.org/10.1167/iovs.09-5030.

Al Chalabi, A., Hardiman, O., 2013. The epidemiology of ALS: a conspiracy of genes, environment and time. Nat. Rev. Neurol. 9 (11), 617–628.

Al Chalabi, A., Hardiman, O., Kiernan, M.C., Chio, A., Rix-Brooks, B., van den Berg, L.H., 2016. Amyotrophic lateral sclerosis: moving towards a new classification system. Lancet Neurol. 15 (11), 1182–1194.

Al-Sarraj, S., King, A., Cleveland, M., Swash, M., 2009. Stratifying disease stages with diagnostically relevant markers. Amyotroph. Lateral Scler. Frontotemporal Degener. 10 (3–4), 117–125. https://doi.org/10.3109/18734870902959159.

Al-Sarraj, S., King, A., Cleveland, M., Pradat, P.-F., Corse, A., Rothstein, J.D., Leigh, P.N., Al-Chalabi, A., Hardiman, O., 2013. The epidemiology of ALS: a conspiracy of genes, environment and time. Nat. Rev. Neurol. 9 (11), 617–628.

Ahmadi, M., Liu, J.-X., Braintstrom, T., Andersen, P.M., Stål, P., Pedrosa-Domelöff, F., 2010. Human extracellular muscle in ALS. Invest. Ophthalmo. Vis. Sci. 51 (7), 3494–3501. https://doi.org/10.1167/iovs.09-5030.

Al Chalabi, A., Hardiman, O., 2013. The epidemiology of ALS: a conspiracy of genes, environment and time. Nat. Rev. Neurol. 9 (11), 617–628.

Al Chalabi, A., Hardiman, O., Kiernan, M.C., Chio, A., Rix-Brooks, B., van den Berg, L.H., 2016. Amyotrophic lateral sclerosis: moving towards a new classification system. Lancet Neurol. 15 (11), 1182–1194.

Al-Sarraj, S., King, A., Cleveland, M., Pradat, P.-F., Corse, A., Rothstein, J.D., Leigh, P.N., Al-Sarraj, S., King, A., Cleveland, M., Swash, M., 2009. Stratifying disease stages with diagnostically relevant markers. Amyotroph. Lateral Scler. Frontotemporal Degener. 10 (3–4), 117–125. https://doi.org/10.3109/18734870902959159.

Brooks, R.B., 1994. El Escorial world federation of neurology criteria for the diagnosis of amyotrophic lateral sclerosis. J. Neurol. Sci. 124, 96–107. https://doi.org/10.1016/0022-510X(94)90191-0.

Brooks, R.B., Miller, R.G., Swash, M., Muntnat, T.L., 2000. El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. Amyotroph. Lateral Scler. 1 (5), 293–299. https://doi.org/10.1080/14602940000795591.

Carvalho, M., Dengler, R., Eisen, A., England, J.D., Kaji, R., Kinbara, J., Mills, K., Mitsumoto, H., Nodera, H., Sferrini, J., Swash, M., 2008. Electrodagnostica criteria for diagnosis of ALS. Clin. Neurophysiol. 119 (3), 497–503.

Carvalho, M., Swash, M., 2009. Stratifying disease stages with different progression rates determined by electrophysiological tests in patients with amyotrophic lateral sclerosis. Muscle Nerve 40 (2), 318.

Cedarbaum, J.M., Stambler, N., Malta, E., Fuller, C., Hilt, D., Thurmond, B., Nakanishi, A., Cedarbaum, J.M., Stambler, N., Malta, E., Fuller, C., Hilt, D., Thurmond, B., Nakanishi, A., Tardieu, F., Galetta, S., 2006. The natural history of primary lateral sclerosis. Neurology 66 (5), 647–653.

Chang, M.C., Kwak, S., 2019. Macroglossia in primary lateral sclerosis: a case report. Int J Mol. Sci. 20 (5), 1250. https://doi.org/10.3390/ijms20051250.

Chen, R., Liao, J., Wang, H., Zhang, H., Phillips, R., 2019. Amyotrophic lateral sclerosis: a distinct entity or part of the ALS spectrum? Amyotroph. Lateral Scler. Frontotemporal Degener. 20 (3–4), 133–145. https://doi.org/10.1080/14602947.2018.1550518.

Fox, M.D., Cohen, A.B., 2012. Bright tongue sign. in ALS.. Neurology 79 (14), 1520.