NeuroImage: Clinical 35 (2022) 103094

Contents lists available at ScienceDirect

NeuroImage: Clinical

journal homepage: www.elsevier.com/locate/ynicl

A multivariate Bayesian classification algorithm for cerebral stage prediction by diffusion tensor imaging in amyotrophic lateral sclerosis

Anna Behler [a], Hans-Peter Müller [a], Albert C. Ludolph [a,b], Dorothée Lulé [a], Jan Kassubek [a,b,*]

[a] Department of Neurology, University of Ulm, Germany
[b] German Center for Neurodegenerative Diseases (DZNE), Ulm, Germany

ARTICLE INFO

Keywords:
Diffusion tensor imaging
Magnetic resonance imaging
Amyotrophic lateral sclerosis
Bayesian classifier
Classification uncertainty

ABSTRACT

Background and Objective: Diffusion tensor imaging (DTI) can be used to tract-wise map correlates of the sequential disease progression and, therefore, to assess disease stages of amyotrophic lateral sclerosis (ALS) *in vivo*. According to a threshold-based sequential scheme, a classification of ALS patients into disease stages is possible, however, several patients cannot be staged for methodological reasons. This study aims to implement a multivariate Bayesian classification algorithm for disease stage prediction at an individual ALS patient level based on DTI metrics of involved tract systems to improve disease stage mapping.

Methods: The analysis of fiber tracts involved in each stage of ALS was performed in 325 ALS patients and 130 age- and gender-matched healthy controls. Based on Bayes' theorem and in accordance with the sequential disease progression, a multistage classifier was implemented. Patients were categorized into *in vivo* DTI stages using the threshold-based method and the Bayesian algorithm. By the margin of confidence, the reliability of the Bayesian categorizations was accessible.

Results: Based on the Bayesian multistage classifier, 88% of all ALS patients could be assigned into an ALS stage compared to 77% using the threshold-based staging scheme. Additionally, the confidence of all classifications could be estimated.

Conclusions: By the application of the multi-stage Bayesian classifier, an individualized *in vivo* cerebral staging of ALS patients was possible based on the sequentially involved tract systems and, furthermore, the reliability of the respective classifications could be determined. The Bayesian classification algorithm is an improvement of the threshold-based staging method and could provide a framework for extending the DTI-based *in vivo* cerebral staging in ALS.

1. Introduction

During the disease progression of amyotrophic lateral sclerosis (ALS), four neuropathological stages are progressed based on the distribution of phosphorylated 43 kDa TAR DNA-binding protein (pTDP-43) in the brain as defined in postmortem studies (Braak et al., 2013; Brettschneider et al., 2013; Tan et al., 2015). Using diffusion tensor imaging (DTI), the analysis of white matter neuronal tracts and, in a hypothesis-guided approach, the *in vivo* identification of tracts involved in the progression of ALS is possible (Kassubek et al., 2014). During the course of ALS, the following tracts were demonstrated to become sequentially involved: the corticospinal tract (CST) at stage 1, the corticorubral and corticopontine tracts at stage 2, the corticostriatal pathway at stage 3, and the proximal portion of the perforant pathway at stage 4 (Kassubek et al., 2018, 2014). The analysis of the fractional anisotropy (FA) using a tract of interest (TOI)-based method allowed for a stage classification at an individual level and, moreover, quantitative mapping of disease progression in the brain (Kassubek et al., 2018). This *in vivo* staging method, based on the stereotypical sequential involvement of tract systems, uses a fixed FA threshold for each tract system based on a healthy control group to provide an individual disease stage classification for ALS patients. However, for a substantial percentage of

Abbreviations: ALS, Amyotrophic Lateral Sclerosis; ALS-FRS-R, revised ALS functional rating scale; CST, corticospinal tract; DTI, Diffusion Tensor Imaging; FA, fractional anisotropy; MNI, Montreal Neurological Institute; MRI, Magnetic Resonance Imaging; pTDP-43, phosphorylated 43kDa TAR DNA-binding protein; TOI, tract of interest.

* Corresponding author at: Dept. of Neurology, University of Ulm, Oberer Eselsberg 45, 89081 Ulm, Germany.
E-mail address: jan.kassubek@uni-ulm.de (J. Kassubek).

https://doi.org/10.1016/j.nicl.2022.103094
Received 23 February 2022; Received in revised form 4 June 2022; Accepted 19 June 2022
Available online 21 June 2022

the patients investigated so far, a classification according to this scheme is not possible due to methodological reasons (Kassubek et al., 2018, 2014): the general inter-subject variability of FA values (Vollmar et al., 2010) contributes to an incomplete separation of ALS patients' and healthy controls' FA values in the CST (Kassubek et al., 2014) so that in patients with rather high FA values at the cross-sectional level, the question remains open whether their FA values are not altered or their FA is altered but *per se* is within the range of FA values of controls. While combining the FA of involved tract systems with structural parameters can improve the separation of patients and healthy controls for general disease diagnosis (Kocar et al., 2021), this uncertainty in the assessment of a patient's FA value in the range of overlapping distributions of patients and controls should be incorporated into disease stage imaging.

Bayesian predictions have already shown their usefulness in the medical field in general (Bours, 2021; Webb and Sidebotham, 2020) and in combination with brain imaging in particular (Morales et al., 2013): Bayesian statistics can be used to calculate the probabilities that a measured value belongs to either the patient or the healthy control class based on the distributions of the observed values in those two classes. This is especially helpful in cases where measured values of a patient group cannot be sharply separated from those of a healthy control group as it is the case for the *in vivo* staging approach. Based on a two-class scenario, the classification algorithm can also be extended for the prediction of different multiple stages along an observed variable (Ichikawa et al., 2017). In contrast, established threshold- or cut-off-based classification schemes provide binary decisions, e.g., 'altered' or 'not altered', for observed values.

In this study, we developed a Bayesian classifier with multiple staging alternatives based on multi-dimensional data sets of diffusion metrics of tract systems involved during the progression of ALS. The *a priori* knowledge about sequential involvement of tracts was considered as a boundary condition and thus, individual probability statements for each *in vivo* DTI stage of ALS. Providing a probability for the classification into a certain stage and therefore confidence of classification might have the potential to improve the stratification of quantitative DTI results at the individual level.

2. Methods

2.1. Subjects and clinical characterization

In total, 325 patients (61.5 ± 12.0 years, 198 male/126 female) with clinically definite or probable sporadic ALS according to the revised version of the El Escorial World Federation of Neurology criteria (Brooks et al., 2000) were included in the study. All patients underwent standardized clinical-neurological and routine laboratory examinations. None of the patients had any history of other neurological or psychiatric disorders. Severity of physical symptoms of patients as measured with the revised ALS functional rating scale (ALS-FRS-R) (Cedarbaum et al., 1999) was 40 ± 6 (range 17 – 47). MiToS functional staging was obtained from ALS-FRS-R domains (Chiò et al., 2015). King's clinical staging was estimated based on the ALS-FRS-R domains (Balendra et al., 2014) which provides high agreement with the original staging scheme (Roche et al., 2012). Due to the retrospective nature of this study, the points reached in the single ALS-FRS-R questions were not available for all patients in addition to the total ALS-FRS-R score. Therefore, functional and clinical staging could only be obtained for 206 of the 325 patients. Magnetic Resonance Imaging (MRI) data were acquired either on a 1.5 T scanner (248 patients) or a 3.0 T scanner (77 patients). About half (52%) of these patients with ALS were already included in an earlier DTI-based staging analysis (Kassubek et al., 2018), i.e., the results of this prior report could be confirmed by the application of the staging approach to additional 156 patients who had not received this analysis before. Longitudinal data sets were available for a limited number of patients. In these cases, the baseline data were used for further analysis. In addition, 130 age- and sex-matched healthy controls (57.6 ± 12.1 years, 72 male/58 female, 80 acquired on a 1.5 T scanner and 50 acquired on a 3 T scanner) were used to define statistical comparison. All subjects gave written consent for the MRI protocol according to institutional guidelines. The study was approved by the Ethical Committee of the University of Ulm, Germany (reference # 19/12).

2.2. MRI acquisition

For scanning, a 1.5 T MRI scanner (Magnetom Symphony, Siemens Medical, Erlangen, Germany) and a 3.0 T MRI scanner (Allegra, Siemens Medical, Erlangen, Germany) were used. The 1.5 T protocol consisted of 52 gradient directions including four $b0$ directions ($b = 1000$ s/mm^2, voxel size 2.0 mm × 2.0 mm × 2.8 mm, 128 × 128 × 64 matrix, $TE = 95$ ms, $TR = 8000$ s). The 3.0 T protocol consisted of 49 gradient directions including one $b0$ directions ($b = 1000$ s/mm^2, voxel size 2.2 mm × 2.2 mm × 2.2 mm, 96 × 128 × 52 matrix, $TE = 85$ ms, $TR = 7600$ s).

2.3. DTI analysis

The DTI analysis software '*Tensor Imaging and Fiber Tracking*' (TIFT) (Müller et al., 2007a) was used for data postprocessing. First, DTI data sets were corrected for eddy current distortions and underwent quality control (Müller et al., 2011). In the next step, on all data sets a non-linear spatial normalization to the Montreal Neurological Institute (MNI) stereotaxic standard space (Brett et al., 2002) was performed by using study-specific DTI templates from all patients and controls (Müller et al., 2007b) as described previously (Müller et al., 2012). For each data set, FA maps were calculated and smoothed with a Gaussian filter with 8 mm full width-at-half-maximum.

Based on FA maps of healthy controls (27 data sets each recorded at 1.5 T and 3 T), three-dimensional correction matrices were calculated and used to correct all 3 T data sets as described previously (Rosskopf et al., 2015). An averaged data set from MNI transformed controls' data was used for the identification of the relevant tract systems by a seed-to-target TOI-based approach. The brain structures to be identified were defined according to the ALS-associated staging system (Brettschneider et al., 2013; Kassubek et al., 2018, 2014), i.e., the CST (according to ALS stage 1), the corticorubral and corticopontine tracts (according to ALS stage 2), the corticostriatal pathway (according to ALS stage 3), and the proximal portion of the perforant pathway (according to ALS stage 4). Fiber tracking was a deterministic streamline tracking approach (Müller et al., 2009) at which the FA threshold was set at 0.2 (Kunimatsu et al., 2004), the Eigenvector scalar product threshold was set at 0.9, the seed regions had a radius of 5 mm, and the target regions had a radius of 10 mm. In a final step, the technique of tract-wise FA statistics was applied to select FA values underlying the fiber tracks for arithmetic averaging. Bihemispheric FA values were averaged and tract-wise corrected for age (Behler et al., 2021). The results of corticopontine and corticorubral tracts were averaged since both were involved in ALS stage 2.

2.4. Threshold-based disease staging categorization

The well-established fiber tract-based staging scheme (Kassubek et al., 2014) is based on FA thresholds ($\mu_{HC} - 0.47\sigma_{HC}$) utilizing the mean μ_{HC} and standard deviation σ_{HC} of the healthy control group in the respective tract systems. After applying the sequential staging cascade, a staging categorization was obtained at the individual level.

2.5. Bayesian statistics

In a particular tract system subject to disease-related alterations, based on Bayes' theorem, the posterior probability that the observed FA value x_i of a tract system would be assigned to the patient class, i.e., would be considered as 'altered' (noted as A), is given by:

$$P(A|x_i) = \frac{P(x_i|A)\cdot P(A)}{P(x_i)} \quad (1)$$

where $P(A)$ is the prior probability for class A, $P(x_i|A)$ is the likelihood of observing x_i given A, i.e., in the patient group, and $P(x_i)$ is the overall probability of x_i being observed. For the condition that a value is considered as belonging to the controls class, i.e., 'not altered' (notated as $\overline{A}$), the formula applies analogously:

$$P(\overline{A}|x_i) = \frac{P(x_i|\overline{A})\cdot P(\overline{A})}{P(x_i)} \quad (2)$$

A schematic example of a Bayesian prediction between these two classes A (the patient class) and $\overline{A}$ (the controls class) is schematically shown in Fig. 1. The likelihood for x_i being observed in the ALS patient class, $P(x_i|A)$, and the likelihood for x_i being observed in the healthy control class, $P(x_i|\overline{A})$, overlap for a broad range of x_i. For $P(x_i|A) = P(x_i|\overline{A})$, the confidence of a classification of x_i into one of both classes is reasonably the lowest, and, therefore, the risk for misclassification the highest. Starting from this point, with decreasing or increasing x_i the difference between $P(x_i|A)$ and $P(x_i|\overline{A})$ increases resulting in increasing confidence in classification. The confidence of a classification is highest when there is just evidence for one of both classes.

2.6. Modeling a multistage Bayesian classifier

The binary case just described for classifying an FA value as 'altered' or 'not altered' can be extended to consider all four tract systems simultaneously without *a priori* knowledge of the patients' stages. The multivariate Bayesian algorithm then calculates the posterior probabilities for each stage based on the likelihood of FA values in each of the four tract systems to be 'altered' or 'not altered'. The vector $\boldsymbol{x} = (x_1, x_2, x_3, x_4)$ contains all four FA values observed in the stage-defining tract systems (see Fig. 1). The likelihood $P(\boldsymbol{x}|S)$ of $\boldsymbol{x}$ in each class $S \in \{0, 1, 2, 3, 4\}$ is obtained by multiplication of the likelihoods $P(x_i|A)$ of tract systems involved in the corresponding stage and likelihoods $P(x_i|\overline{A})$ of tract systems not yet involved. It should be noted that the condition of all four x_i to be considered as 'not altered' is referred to as class $S = 0$ here. The posterior probabilities of ALS stages are therefore calculated as.

$$P(S|\boldsymbol{x}) = \frac{1}{P(\boldsymbol{x})}\prod_k P(x_k|A)\cdot\prod_l P(x_l|\overline{A})\cdot P(S) \quad (3)$$

with $k \in \{0, 1, 2, 3, 4\}$ in class S affected tract systems and the number of non-involved tract systems $l = 4 - k$, in order to take the sequential order into account in which the tracts are affected during the disease. An exemplary illustration of the resulting posterior probabilities for each class is given in Fig. 1. The classifier assigns $\boldsymbol{x}$ to ALS stage C according to the following classification rule:

$$C = \begin{cases} \text{'not stageable'} & \text{if } P(0|\boldsymbol{x}) > \sum_{S=1}^{4} P(S|\boldsymbol{x}) \\ \underset{S}{\operatorname{argmax}}\ P(S|\boldsymbol{x}) & \text{otherwise} \end{cases} \quad (4)$$

To use Bayes' theorem in this multi-stage setting, two assumptions were made:

1) likelihoods are normal distributed with mean μ and standard deviation σ
2) standard deviations from the patients and controls group are equal in each tract system: $P(x_k|A) \sim \mathcal{N}(\mu_A, \sigma_{HC})$ and $P(x_k|\overline{A}) \sim \mathcal{N}(\mu_{HC}, \sigma_{HC})$,
3) prior probabilities $P(S)$ are equal.

The assumption of a normal distributed FA values in the tract systems was tested using D'Agostino's K-squared test (D'Agostino, 1971), rejecting the null hypothesis at $p < 0.05$; in the control group, FA values outside $\mu \pm 3\sigma$ were removed. Since the CST is involved in every ALS stage, the variation between the distributions of the respective mean FA of patients and healthy controls were the starting points for defining the mean μ_A of 'altered' FA values in tract systems involved in a later (i.e., higher) stage.

The computational implementation of the Bayesian classifier and the statistical analysis was performed using Python 3.9 with the following machine learning modules: Scikit-learn 1.7 (Pedregosa et al., 2018) and PyTorch 1.9 (Paszke et al., 2019).

$$P(\bar{A}|x_i) = \frac{P(x_i|\bar{A}) \cdot P(\bar{A})}{P(x_i)} \quad (2)$$

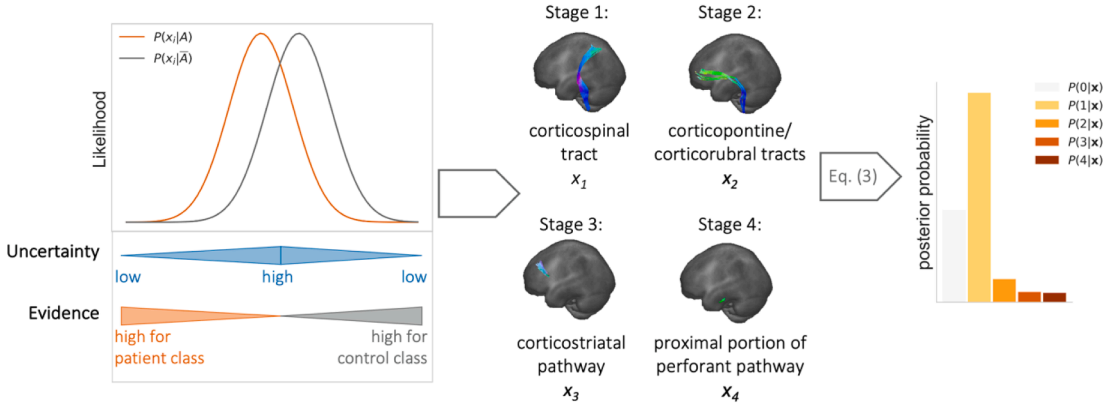

Fig. 1. Schematic illustration of likelihoods and posterior probabilities for the Bayesian prediction of measured FA values x_i belonging either to the patient (orange) or the healthy control class (grey) in a tract system associated with disease progression. The superposition of the two likelihoods leads to a range of greater uncertainty in the classification. Therefore, the evidence for the respective class membership increases with decreasing or increasing x_i. Based on this simple case in one tract system, the likelihoods of FA values of four tract systems are evaluated for the classification of ALS patients and posterior probabilities are calculated for each ALS stage.

2.7. Margin of confidence

Since the decision of the multilevel Bayesian classifier is based on probabilities, the risk of misclassification is accessible (Li and Ma, 2020). The margin of confidence can be used as a confidence variable and is calculated as the difference between the posterior probabilities of the two most probable classes. If those two classes have equal probability, the classification then is 'least confident'.

3. Results

In the healthy control group, the means and standard deviations of FA values in the tract systems considered were obtained as follows: CST, 0.335 ± 0.019: corticopontine and corticorubral tracts, 0.323 ± 0.018; corticostriatal tract, 0.331 ± 0.026; proximal portion of perforant pathway, 0.165 ± 0.020. Among the patient measurements, the mean FA in the CST was 0.319. Based on the differences between patients and healthy controls in the CST, the means of the distribution for the tract systems involved in stages 2 to 4 were calculated to 0.307 for the corticopontine and corticorubral tracts, 0.315 for the corticostriatal tract, and 0.157 for the perforant pathway. Standard deviations of FA in the altered tract systems, according to the stages, were equated with those of healthy controls according to the assumptions made. Based on these values, the likelihoods were estimated. For a comparison with the decision rules of the threshold method, the posterior probabilities were calculated for each tract system according to the Eqs. (1) and (2) (Fig. 2). In the CST (Fig. 2a) and the corticopontine and corticorubral tracts (Fig. 2b), the threshold of a FA considered as altered was equal to the FA with the same posterior probability for class A and $\overline{A}$. In the corticostriatal pathway (Fig. 2c) and the proximal portion of the perforant pathway (Fig. 2d), the threshold was 0.020 and 0.012, respectively, below the FA with equal posterior probabilities.

After implementing the Bayesian multi-stage classifier (Eqs. (3) and (4)), 88% of the ALS patients could be assigned to a DTI stage. In comparison, only 77% of the ALS patients could be assigned using the threshold-based method. The detailed distribution of patients across the stages is shown in Fig. 3 for both staging approaches. 229 patients showed no difference in their stage association between the two classification methods. Out of the 75 patients who could not be classified using the threshold method, Bayesian staging could be used to assign 14 of them to Stage 1, 4 to Stage 2, 7 to Stage 3, and 11 to Stage 4. The margin of confidence for the Bayesian classifications ranged from < 0.001% to 75.7%.

Disease duration and ALS-FRS-R did not differ between patients in different DTI stages. For the ALS-FRS-R, however, a trend for a lower score at higher stages could be identified (Fig. 4). An analysis of MiToS and King's staging of the patients in the different DTI stages also did not show a significant association between the clinical staging and the DTI-

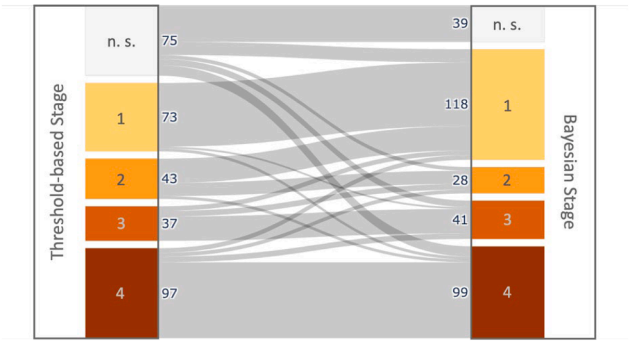

Fig. 3. Number of patients with amyotrophic lateral sclerosis (ALS) in each DTI stage. Staging categorization was performed for 325 patients using the established threshold-based method and the Bayesian classifier. Patients who could not be classified into one of the four DTI stages are categorized as 'not stageable' ('n. s.').

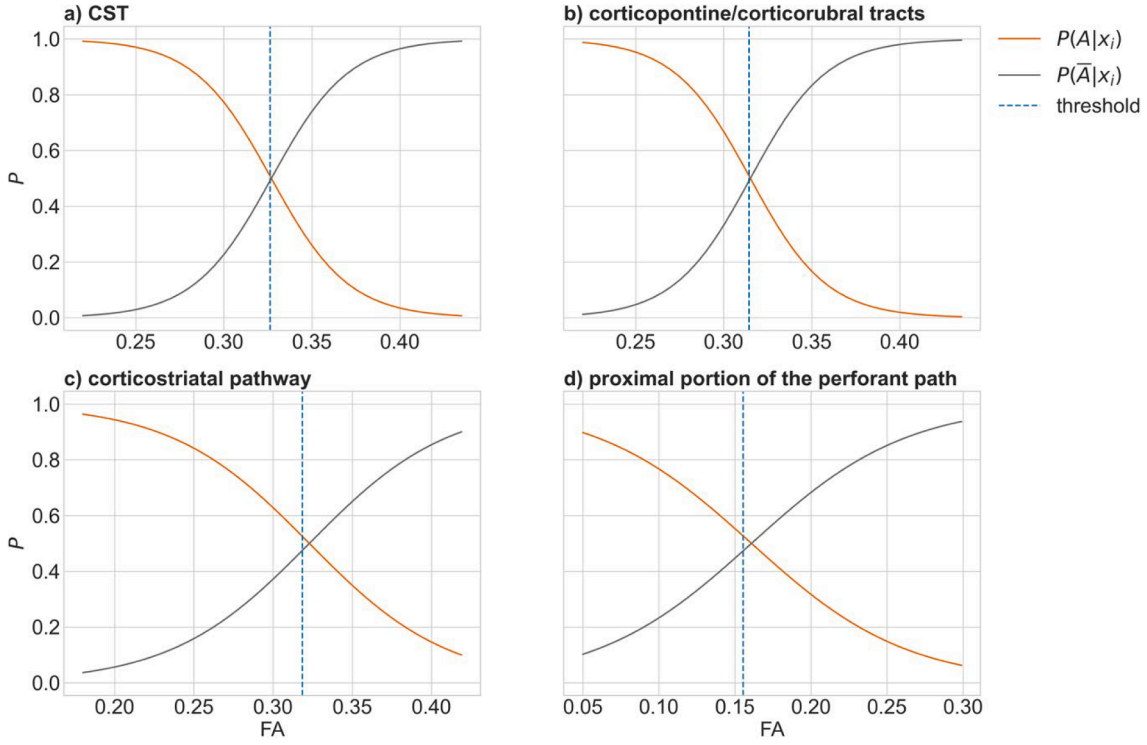

Fig. 2. Posterior probabilities for fractional anisotropy (FA) values in a) the corticospinal tract (CST), b) the corticopontine and corticorubral tracts, c) the corticostriatal pathway, and d) the proximal portion of the perforant pathway to be classified as 'altered' ($P(A|x_i)$, orange) or 'not altered' ($P(\overline{A}|x_i)$, grey). The individual tract systems were analyzed independently and the respective threshold of the threshold-based staging method is shown as dashed blue line.

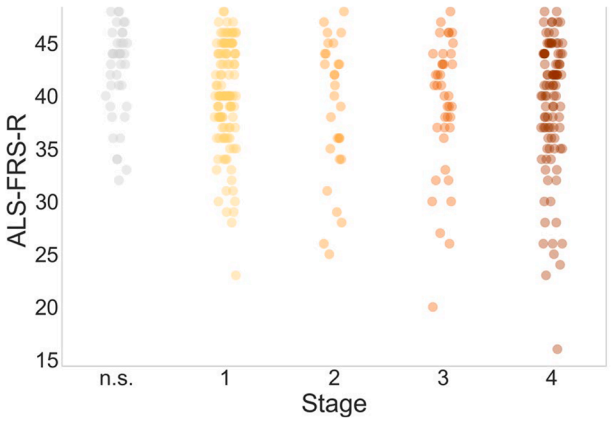

Fig. 4. Revised ALS functional rating score (ALS-FRS-R) of patients with ALS in different stages as classified according to tract-based Bayesian classification. Patients who could not be classified into one of the four DTI stages are categorized as 'not stageable' ('n.s.'). ALS = Amyotrophic Lateral Sclerosis, DTI = Diffusion Tensor Imaging.

based staging (see Supplementary Material Table S1).

4. Discussion

It could be shown how a Bayesian classifier for a multivariate setting with several staging alternatives can be developed and used to classify ALS patients into *in vivo* DTI stages based on their diffusion metrics in the sequentially involved tract systems. The algorithm evaluated the posterior probabilities for each stage based on likelihoods of altered or non-altered FA values in the tract systems involved or not involved according to the respective disease stage. The development of a multivariable Bayesian prediction algorithm has shown that it is possible to obtain probability statements about the impairment of stage-defining tract systems and the staging categorization itself at an individual level for ALS patients in accordance with the sequential involvement of the tract systems during the course of the disease. The transition from fixed 'either-or' decisions to probability statements represents a substantial improvement in the DTI staging approach. Compared to the threshold-based staging method, the new Bayesian multistage classifier reduced the number of not-stageable patients by half. Comparison of the classification of patients using both classification methods (Fig. 3) showed that for most patients, the classification remained the same. Also, the tract-wise comparison of the threshold with the FA values for equal posterior probabilities (Fig. 2) suggests that Bayesian classification should be seen as an extension and improvement of the previous threshold-based approach but does not invalidate it. These advantages of the multivariate Bayesian classification support the use of machine learning approaches to interpret and categorize imaging data at the single-patient level in ALS (Meier et al., 2020; Bede et al., 2022).

Another advantage of the Bayesian approach over a threshold-based classification is the availability of probability statements for each stage, and, therefore, a measure of confidence for each classification. The threshold-based approach cannot provide any information about how certain a specific classification is, although due to the uncertainty of an observed value, misclassification could occur, especially for values in whose measurement uncertainty the threshold lies. Due to the higher information content compared to the threshold method, the Bayesian classification of a given patient into a stage can be evaluated differently. For example, if the certainty is low, the DTI measurement could be repeated to obtain more confidence in the staging classification or unnecessary tests could be avoided if the classification is certain. As a quantitative measure, FA values are subject to a certain degree of uncertainty (Tanno et al., 2021) and while the inclusion of such measurement uncertainty at the individual level is not possible with the threshold-based method, the Bayesian classifier could be extended accordingly (Qin et al., 2011).

The prior probabilities within the Bayesian staging algorithm correspond to the probability for the stages prior to DTI and provide the opportunity to incorporate existing knowledge about a patient, e.g., prevalence or previous test results. Additional specifications of the prior probabilities might further reduce the number of not stageable patients; other measures that are associated with disease progression, such as performance in cognitive tests (Lulé et al., 2018), could be used to estimate prior probabilities and to assess disease progression in a multimodal approach. Such a combination of factors would be consistent with the natural decision-making process of physicians (Gill et al., 2005) and might help to enhance the potential role of DTI-based MRI as a biomarker in ALS (Baek et al., 2020; Kalra et al., 2020).

Given the disease's clinical heterogeneity, it is not surprising that a higher DTI stage was not associated with higher disease duration. The DTI staging approach in general assesses only the ALS-associated neuropathological changes in the brain according to the proposed propagation pattern (Brettschneider et al., 2013), which can only be reflected to a limited extent by assessing functionality with the ALS-FRS-S and clinical MiToS and King's staging, given that many clinical deficits and dysfunctions are determined by the degeneration of the second motor neuron.

This study is not without limitations. The likelihoods were assumed to be normally distributed with equal standard deviations for altered and normal FA values in each tract system. In tract systems affected during the disease, the standard deviation of FA values could be higher since the FA values decrease continuously during the course. The patients' ALS stages were not confirmed by post-mortem examination; calculation of the likelihoods based on autopsy-confirmed data could improve the confidence of the classifications.

In summary, we demonstrated how Bayesian statistics can be used to design an algorithm that improves the *in vivo* classification of ALS patients into a neuropathological stage according to the sequential scheme of disease progression. In addition to classification, confidence statements can be made about the stage categorization itself at the individual level. This approach could be a further step ahead to form a basis for combining DTI with other modalities to stratify MRI data according to the neuropathological stages *in vivo* and thus could improve the accuracy of DTI-based imaging biomarkers in ALS.

CRediT authorship contribution statement

Anna Behler: Formal analysis, Investigation, Visualization, Writing – original draft. **Hans-Peter Müller:** Formal analysis, Investigation, Software, Supervision, Writing – original draft. **Albert C. Ludolph:** Investigation, Writing – review & editing. **Dorothée Lulé:** Data curation, Formal analysis, Writing – review & editing. **Jan Kassubek:** Project administration, Conceptualization, Supervision, Investigation, Writing – original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We would like to thank Sonja Fuchs for her help in the acquisition of MRI data and Kornelia Günther for her help in assembling the clinical data. Further thanks to the Ulm University Center for Translational Imaging MoMAN for its support.

Funding

None.

Author contributions

AB: study concept and design, data acquisition, data analysis and interpretation, drafting of manuscript.

H-PM: study concept and design, data analysis and interpretation of data, critical revision of the manuscript for important intellectual content.

ACL: critical revision of manuscript for important intellectual content.

DL: data analysis and interpretation of data, critical revision of the manuscript for important intellectual content.

JK: study concept and design, analysis and interpretation, critical revision of manuscript for important intellectual content, study supervision.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.nicl.2022.103094.

References

Baek, S.-H., Park, J., Kim, Y.H., Seok, H.Y., Oh, K.-W., Kim, H.-J., Kwon, Y.-J., Sim, Y., Tae, W.-S., Kim, S.H., Kim, B.-J., 2020. Usefulness of diffusion tensor imaging findings as biomarkers for amyotrophic lateral sclerosis. Sci. Rep. 10, 5199. https://doi.org/10.1038/s41598-020-62049-0.

Balendra, R., Jones, A., Jivraj, N., Knights, C., Ellis, C.M., Burman, R., Turner, M.R., Leigh, P.N., Shaw, C.E., Al-Chalabi, A., 2014. Estimating clinical stage of amyotrophic lateral sclerosis from the ALS Functional Rating Scale. Amyotroph. Lateral. Scler. Frontotemporal. Degener. 15, 279–284. https://doi.org/10.3109/21678421.2014.897357.

Bede, P., Murad, A., Lope, J., Li Hi Shing, S., Finegan, E., Chipika, R.H., Hardiman, O., Chang, K.M., 2022. Phenotypic categorisation of individual subjects with motor neuron disease based on radiological disease burden patterns: A machine-learning approach. J. Neurol. Sci. 432, 120079. https://doi.org/10.1016/j.jns.2021.120079.

Behler, A., Kassubek, J., Müller, H.-P., 2021. Age-related alterations in DTI metrics in the human brain—consequences for age correction. Front. Aging Neurosci. 13, 682109 https://doi.org/10.3389/fnagi.2021.682109.

Bours, M.J.L., 2021. Bayes' rule in diagnosis. J. Clin. Epidemiol. 131, 158–160. https://doi.org/10.1016/j.jclinepi.2020.12.021.

Braak, H., Brettschneider, J., Ludolph, A.C., Lee, V.M., Trojanowski, J.Q., Tredici, K.D., 2013. Amyotrophic lateral sclerosis—a model of corticofugal axonal spread. Nat. Rev. Neurol. 9, 708–714. https://doi.org/10.1038/nrneurol.2013.221.

Brett, M., Johnsrude, I.S., Owen, A.M., 2002. The problem of functional localization in the human brain. Nat. Rev. Neurosci. 3, 243–249. https://doi.org/10.1038/nrn756.

Brettschneider, J., Del Tredici, K., Toledo, J.B., Robinson, J.L., Irwin, D.J., Grossman, M., Suh, E., Van Deerlin, V.M., Wood, E.M., Baek, Y., Kwong, L., Lee, E.B., Elman, L., McCluskey, L., Fang, L., Feldengut, S., Ludolph, A.C., Lee, V.-M.-Y., Braak, H., Trojanowski, J.Q., 2013. Stages of pTDP-43 pathology in amyotrophic lateral sclerosis: ALS stages. Ann. Neurol. 74, 20–38. https://doi.org/10.1002/ana.23937.

Brooks, B.R., Miller, R.G., Swash, M., Munsat, T.L., 2000. El Escorial revisited: Revised criteria for the diagnosis of amyotrophic lateral sclerosis. Amyotroph. Lateral Scler. Other Motor Neuron Disord. 1, 293–299. https://doi.org/10.1080/146608200300079536.

Cedarbaum, J.M., Stambler, N., Malta, E., Fuller, C., Hilt, D., Thurmond, B., Nakanishi, A., 1999. The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. J. Neurol. Sci. 169, 13–21. https://doi.org/10.1016/S0022-510X(99)00210-5.

Chiò, A., Hammond, E.R., Mora, G., Bonito, V., Filippini, G., 2015. Development and evaluation of a clinical staging system for amyotrophic lateral sclerosis. J. Neurol. Neurosurg. Psychiatry 86, 38–44. https://doi.org/10.1136/jnnp-2013-306589.

D'Agostino, R.B., 1971. An omnibus test of normality for moderate and large size samples. Biometrika 58, 341–348. https://doi.org/10.1093/biomet/58.2.341.

Gill, C.J., Sabin, L., Schmid, C.H., 2005. Why clinicians are natural bayesians. BMJ 330, 1080–1083. https://doi.org/10.1136/bmj.330.7499.1080.

Ichikawa, S., Motosugi, U., Enomoto, N., Matsuda, M., Onishi, H., 2017. Noninvasive hepatic fibrosis staging using mr elastography: The usefulness of the bayesian prediction method: MR Elastography With Bayesian Method. J. Magn. Reson. Imaging 46, 375–382. https://doi.org/10.1002/jmri.25551.

Kalra, S., Müller, H.-P., Ishaque, A., Zinman, L., Korngut, L., Genge, A., Beaulieu, C., Frayne, R., Graham, S.J., Kassubek, J., 2020. A prospective harmonized multicenter DTI study of cerebral white matter degeneration in ALS. Neurology 95, e943–e952. https://doi.org/10.1212/WNL.0000000000010235.

Kassubek, J., Müller, H.-P., Del Tredici, K., Brettschneider, J., Pinkhardt, E.H., Lule, D., Bohm, S., Braak, H., Ludolph, A.C., 2014. Diffusion tensor imaging analysis of sequential spreading of disease in amyotrophic lateral sclerosis confirms patterns of TDP-43 pathology. Brain 137, 1733–1740. https://doi.org/10.1093/brain/awu090.

Kassubek, J., Müller, H.-P., Del Tredici, K., Lulé, D., Gorges, M., Braak, H., Ludolph, A.C., 2018. Imaging the pathoanatomy of amyotrophic lateral sclerosis in vivo: targeting a propagation-based biological marker. J. Neurol. Neurosurg. Psychiatry 89, 374–381. https://doi.org/10.1136/jnnp-2017-316365.

Kocar, T.D., Behler, A., Ludolph, A.C., Müller, H.-P., Kassubek, J., 2021. Multiparametric microstructural MRI and machine learning classification yields high diagnostic accuracy in amyotrophic lateral sclerosis: proof of concept. Front. Neurol. 12, 745475 https://doi.org/10.3389/fneur.2021.745475.

Kunimatsu, A., Aoki, S., Masutani, Y., Abe, O., Hayashi, N., Mori, H., Masumoto, T., Ohtomo, K., 2004. The optimal trackability threshold of fractional anisotropy for diffusion tensor tractography of the corticospinal tract. Magn. Reson. Med. Sci. 3, 11–17. https://doi.org/10.2463/mrms.3.11.

Li, H.-H., Ma, W.J., 2020. Confidence reports in decision-making with multiple alternatives violate the Bayesian confidence hypothesis. Nat. Commun. 11, 2004. https://doi.org/10.1038/s41467-020-15581-6.

Lulé, D., Böhm, S., Müller, H.-P., Aho-Özhan, H., Keller, J., Gorges, M., Loose, M., Weishaupt, J.H., Uttner, I., Pinkhardt, E., Kassubek, J., Del Tredici, K., Braak, H., Abrahams, S., Ludolph, A.C., 2018. Cognitive phenotypes of sequential staging in amyotrophic lateral sclerosis. Cortex 101, 163–171. https://doi.org/10.1016/j.cortex.2018.01.004.

Meier, J.M., Burgh, H.K., Nitert, A.D., Bede, P., Lange, S.C., Hardiman, O., Berg, L.H., Heuvel, M.P., 2020. Connectome-based propagation model in amyotrophic lateral sclerosis. Ann. Neurol. 87, 725–738. https://doi.org/10.1002/ana.25706.

Morales, D.A., Vives-Gilabert, Y., Gómez-Ansón, B., Bengoetxea, E., Larrañaga, P., Bielza, C., Pagonabarraga, J., Kulisevsky, J., Corcuera-Solano, I., Delfino, M., 2013. Predicting dementia development in Parkinson's disease using Bayesian network classifiers. Psychiatry Res. Neuroimaging 213, 92–98. https://doi.org/10.1016/j.pscychresns.2012.06.001.

Müller, H.-P., Süssmuth, S.D., Landwehrmeyer, G.B., Ludolph, A., Tabrizi, S.J., Kloppel, S., Kassubek, J., 2011. Stability effects on results of diffusion tensor imaging analysis by reduction of the number of gradient directions due to motion artifacts: an application to presymptomatic Huntington's disease. PLoS Curr. 3, RRN1292. https://doi.org/10.1371/currents.RRN1292.

Müller, H.-P., Unrath, A., Huppertz, H.-J., Ludolph, A.C., Kassubek, J., 2012. Neuroanatomical patterns of cerebral white matter involvement in different motor neuron diseases as studied by diffusion tensor imaging analysis. Amyotroph. Lateral. Scler. 13, 254–264. https://doi.org/10.3109/17482968.2011.653571.

Müller, H.-P., Unrath, A., Ludolph, A.C., Kassubek, J., 2007a. Preservation of diffusion tensor properties during spatial normalization by use of tensor imaging and fibre tracking on a normal brain database. Phys. Med. Biol. 52, N99–N109. https://doi.org/10.1088/0031-9155/52/6/N01.

Müller, H.-P., Unrath, A., Riecker, A., Pinkhardt, E.H., Ludolph, A.C., Kassubek, J., 2009. Intersubject variability in the analysis of diffusion tensor images at the group level: fractional anisotropy mapping and fiber tracking techniques. Magn. Reson. Imaging 27, 324–334. https://doi.org/10.1016/j.mri.2008.07.003.

Müller, H.-P., Unrath, A., Sperfeld, A.D., Ludolph, A.C., Riecker, A., Kassubek, J., 2007b. Diffusion tensor imaging and tractwise fractional anisotropy statistics: quantitative analysis in white matter pathology. Biomed. Eng. Online 6, 42. https://doi.org/10.1186/1475-925X-6-42.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., Chintala, S., 2019. PyTorch: An Imperative Style, High-Performance Deep Learning Library, in: Wallach, H., Larochelle, H., Beygelzimer, A., Alché-Buc, F. d', Fox, E., Garnett, R. (Eds.), Advances in Neural Information Processing Systems 32. Curran Associates, Inc., pp. 8024–8035.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Müller, A., Nothman, J., Louppe, G., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2018. Scikit-learn: Machine Learning in Python. arXiv:1201.0490 [cs].

Qin, B., Xia, Y., Wang, S., Du, X., 2011. A novel Bayesian classification for uncertain data. Knowl.-Based Syst. 24, 1151–1158. https://doi.org/10.1016/j.knosys.2011.04.011.

Roche, J.C., Rojas-Garcia, R., Scott, K.M., Scotton, W., Ellis, C.E., Burman, R., Wijesekera, L., Turner, M.R., Leigh, P.N., Shaw, C.E., Al-Chalabi, A., 2012. A proposed staging system for amyotrophic lateral sclerosis. Brain 135, 847–852. https://doi.org/10.1093/brain/awr351.

Rosskopf, J., Müller, H.-P., Dreyhaupt, J., Gorges, M., Ludolph, A.C., Kassubek, J., 2015. Ex post facto assessment of diffusion tensor imaging metrics from different MRI protocols: Preparing for multicentre studies in ALS. Amyotroph. Lateral. Scler. Frontotemporal Degener. 16, 92–101. https://doi.org/10.3109/21678421.2014.977297.

Tan, R.H., Kril, J.J., Fatima, M., McGeachie, A., McCann, H., Shepherd, C., Forrest, S.L., Affleck, A., Kwok, J.B.J., Hodges, J.R., Kiernan, M.C., Halliday, G.M., 2015. TDP-43 proteinopathies: pathological identification of brain regions differentiating clinical phenotypes. Brain 138, 3110–3122. https://doi.org/10.1093/brain/awv220.

Tanno, R., Worrall, D.E., Kaden, E., Ghosh, A., Grussu, F., Bizzi, A., Sotiropoulos, S.N., Criminisi, A., Alexander, D.C., 2021. Uncertainty modelling in deep learning for

safer neuroimage enhancement: demonstration in diffusion MRI. Neuroimage 225, 117366. https://doi.org/10.1016/j.neuroimage.2020.117366.

Vollmar, C., O'Muircheartaigh, J., Barker, G.J., Symms, M.R., Thompson, P., Kumari, V., Duncan, J.S., Richardson, M.P., Koepp, M.J., 2010. Identical, but not the same: Intra-site and inter-site reproducibility of fractional anisotropy measures on two 3.0T scanners. Neuroimage 51, 1384–1394. https://doi.org/10.1016/j.neuroimage.2010.03.046.

Webb, M.P.K., Sidebotham, D., 2020. Bayes' formula: a powerful but counterintuitive tool for medical decision-making. BJA Educ. 20, 208–213. https://doi.org/10.1016/j.bjae.2020.03.002.