



Preoperative language tract integrity is a limiting factor in recovery from aphasia after glioma surgery

Gordian Prasse^{a,*}, Hans-Jonas Meyer^b, Cordula Scherlach^a, Jens Maybaum^a,
Anastasia Hoffmann^a, Johannes Kasper^c, Michael Karl Fehrenbach^c, Florian Wilhelmy^c,
Jürgen Meixensberger^c, Karl-Titus Hoffmann^a, Tim Wende^c

^a Institute of Neuroradiology, University Hospital Leipzig, 04103 Leipzig, Germany

^b Department of Radiology, University Hospital Leipzig, 04103 Leipzig, Germany

^c Department of Neurosurgery, University Hospital Leipzig, 04103 Leipzig, Germany

ARTICLE INFO

Keywords:

MRI
DTI
Glioma
Neurosurgery
White matter integrity
Aphasia
Rehabilitation

ABSTRACT

Aphasia can occur in a broad range of pathological conditions that affect cortical or subcortical structures. Here we test the hypothesis that white matter integrity of language pathways assessed by preoperative diffusion tensor imaging (DTI) is associated with language performance and its recovery after glioma resection. 27 patients with preoperative DTI were included. Segmentation of the arcuate fascicle (AF), the inferior fronto-occipital fascicle (IFOF), the inferior longitudinal fascicle (ILF), the superior longitudinal fascicle (SLF), and the uncinate fascicle (UF) was performed with a fully-connected neural network (FCNN, TractSeg). Median fractional anisotropy (FA) was extracted from the resulting volumes as surrogate marker for white matter integrity and tested for correlation with clinical parameters. After correction for demographic data and multiple testing, preoperative white matter integrity of the IFOF, the ILF, and the UF in the left hemisphere were independently and significantly associated with aphasia three months after surgery. Comparison between patients with and without aphasia three months after surgery revealed significant differences in preoperative white matter integrity of the left AF ($p = 0.021$), left IFOF ($p = 0.015$), left ILF ($p = 0.003$), left SLF ($p = 0.001$, $p = 0.021$, $p = 0.043$ for respective sub-bundles 1–3), left UF ($p = 0.041$) and the right AF ($p = 0.027$). Preoperative assessment of white matter integrity of the language network by time-efficient MRI protocols and FCNN-driven segmentation may assist in the evaluation of postoperative rehabilitation potential in glioma patients.

1. Introduction

Aphasia is a frequent neurological symptom that can occur in the context of numerous pathological conditions of the central nervous system, including vascular incidents, neurodegenerative disease, hemorrhage as well as primary and secondary neoplasms (Geranmayeh et al., 2014). The association of anatomical and functional brain structures with language-related abilities has been widely explored using a range of methodologies. Diffusion-tensor (DTI) magnetic resonance imaging (MRI) is an important tool in the assessment of white matter microstructure (Basser and Pierpaoli, 1996; Le Bihan et al., 2001). The collinear orientation of axons within fiber bundles facilitates water diffusion along their trajectory and limits diffusion in directions diverging from it. Measuring water diffusivity along a number of spatial

axes and fitting a descriptive model, e.g., the diffusion tensor model, enables the estimation of the main trajectory of anatomical white matter tracts within a volume element (voxel), allowing the assessment of white matter integrity in a non-invasive, *in vivo* way.

According to the contemporary dual-stream model, language pathways can be subdivided into a ventral and a dorsal stream, allowing for parallel signal processing (Chang et al., 2015; Dick et al., 2014; Hickok and Poeppel, 2007; Tuncer et al., 2021). The dorsal part consists of the *Arcuate Fascicle* (AF) and the *Superior Longitudinal Fascicle* (SLF) (Catani and Mesulam, 2008; Saur et al., 2008). The ventral part includes the *Inferior Fronto-Occipital Fascicle* (IFOF), the *Inferior Longitudinal Fascicle* (ILF), and the *Uncinate Fascicle* (UF) (Almairac et al., 2015; Duffau et al., 2009; Mandonnet et al., 2007). The SLF anatomy is subject to current discussion; it can be subdivided into two reproducible parts (SLF2 and

* Corresponding author.

E-mail address: gordian.prasse@medizin.uni-leipzig.de (G. Prasse).

SLF3) and one that could be shown to have a close to non-distinguishable relationship to the cingulum (SLF1) (Wang et al., 2016).

Objective assessment of white matter integrity by means of DTI faces several obstacles in the clinical setting. First and most importantly, perifocal edema leads to an increase of extracellular water, thus decreasing the fractional anisotropy (FA), complicating the assessment of white matter integrity and further computational processing of the image data, e. g., tractography. Clinical routine can limit imaging scan time for more dense acquisition of imaging data as well as resources in post-processing, both elements that contribute to the compensation of the more difficult assessment setting compared to research in healthy subjects. However, processing time and reproducibility can be optimized by application of automatic segmentation (Wasserthal et al., 2018).

Symptomatic brain mass lesions frequently undergo surgical resection as part of the therapeutic concept. The gold standard for the identification of language-eloquent areas is direct electrical stimulation (DES) *in situ*. For speech monitoring, this technique can only be applied during awake craniotomy and is therefore limited in its applicability (Sanai et al., 2008). The combination of preoperative DTI and intraoperative DES could be shown to enhance surgical performance and safety while retaining a high rate of functional preservation (Bello et al., 2008). While advanced MRI techniques such as DTI and functional MRI have been shown to be able to contribute to the planning of the surgical procedure to achieve a safe resection (Essayed et al., 2017), there is an ongoing debate on the limitations of advanced neuroimaging, e.g., DTI and functional MRI (fMRI) and the associated drawbacks in terms of patient selection and possible undertreatment (Azad and Duffau, 2020). While there is a significant correlation between white matter integrity of the language network and aphasia symptoms, this correlation seems to be independent of the etiology of aphasia (Zhang et al., 2021).

Here we test the hypothesis that preoperatively assessed white matter integrity is associated with short-term language outcome in glioma surgery and potentially able to identify patients at risk for lack of rehabilitation potential. To explore potential influence of neuronal plasticity on recovery, both lesions of the left and the right hemisphere were included in this study, and both hemispheres were included in the analysis. To investigate benefits of DTI in treatment planning and to demonstrate potential usability in the clinical setting, a minimalistic DTI protocol and a novel automated post-processing pipeline are used, requiring minimal expert interaction.

2. Materials and methods

2.1. Participants

This study was approved by the ethics committee of the University of Leipzig (297/21-ek). The hospital database was searched for glioma patients with preoperative DTI scans between January 2020 and January 2021. Inclusion criteria for this prospective cohort design were: (1) Completed gross total resection (GTR) or subtotal resection (STR) of a previously therapy-naive glioma, (2) histopathological confirmation of the diagnosis, (3) right-handedness, (4) MRI scan of sufficient quality for analysis and successful preprocessing, as checked visually, (5) absence of severe perioperative complication that might affect speech outcome, e. g. stroke, hemorrhage, or abscess, (6) age > 18 years.

2.2. Clinical data

To objectively quantify aphasia by communication ability, similar to the aphasia scoring in the National Institute of Neurological Disorders Stroke Scale (NIHSS, Brott et al., 1989), the aphasia score (AS) can be used, providing a semi-quantitative scale (Ille et al., 2016; Krieg et al., 2014; Sollmann et al., 2015): (0) no deficit, (1) mild deficit (normal speech comprehension and/or conversational speech with slight amnesic aphasia, adequate communication ability), (2) medium deficit

(minor disruption of speech comprehension and/or conversational speech, adequate communication ability), and (3) severe deficit (major disruption of speech comprehension and/or conversational speech, clear impairment of communication ability). AS was assessed by a neurosurgeon according to in-house documentation of the neurological status at three points in time during the clinical course: preoperative, post-operative in the intensive care setting, and at three-month follow-up. The scoring investigator was blinded to the medical imaging.

2.3. Imaging

MRI was performed as part of the standard preoperative workup on a 3 T system (Ingenia, Philips, Eindhoven, Netherlands). The DTI acquisition was optimized for short scan time and consisted of single-shot *epi*-planar imaging (*EPI*) sequences using the following parameters: number of diffusion directions = 32, repetition time = 7010 ms, echo time = 102 ms, b-value = 1000 s/m², number of b = 0 volumes: 1, field of view = 222 × 222 mm², voxel size = 1.98 × 1.98 × 2.7 mm³, number of excitations = 1, gap between slices = 0, and acquisition time = 330 s.

Preprocessing included bias field correction using the N4 method (Tustison et al., 2013), removal of Gibbs ring artifacts (Neto Henriques, 2018), and denoising using the Patch2Self method (Fadnavis et al., 2020) based on self-supervised learning as part of the Dipy framework (Garyfallidis et al., 2014). Intra-individual registration was performed rigorously using elastix software framework (Klein et al., 2010; Shamonin et al., 2013) in four sequential steps: translation, rigid, and twice affine. In order to present imaging data as clean as possible as input to the automated tractography pipeline, emphasis was laid on preprocessing steps that do not require extra imaging time. The intra-individual registration step was optimized for rigorously clean results. Imaging and preprocessing results were checked visually for quality.

Automated segmentation of white matter tracts was performed by TractSeg (Wasserthal et al., 2018) after spatial reorientation of the DTI volumes with FLIRT (Jenkinson et al., 2002; Jenkinson and Smith, 2001). Tract segmentations underwent visual quality control to identify implausible results in the presence of altered brain anatomy, e.g., false-positive segmentation of edema, tumor or adjacent fiber bundles, and false-negative segmentation of anatomically distinct parts of the respective tracts.

FA maps were generated from the preprocessed DTI volumes using MRtrix 3 (Tournier et al., 2019). Median FA values were extracted using the automatically generated segmentation maps for each tract.

A graphical summary of the automated tractometry pipeline is provided in Fig. 1.

The datasets generated during and analyzed in the current study are not publicly available due to the privacy of research participants but are available from the corresponding author upon reasonable request.

2.4. Statistical analysis

The statistical evaluation utilized established Python (version 3.8) statistics modules (McKinney, 2010; Seabold and Perktold, 2010; Vallat, 2018; Waskom, 2021). FA values of all tracts were tested for normal distribution, as determined by the Shapiro-Wilk test, yielding non-normal distribution in all tracts. Correlation analyses were conducted using non-parametric Spearman's ρ , controlling for demographic variables by applying partial correlation and correcting for multiple tests using the Benjamini-Hochberg method (Benjamini and Hochberg, 1995) (denoted as p and p_{corr} , respectively). Group comparison was conducted using analysis of covariance (ANCOVA). All significance tests between FA and clinical variables were corrected for age, sex, affected hemisphere, and grading as confounding demographic covariates.

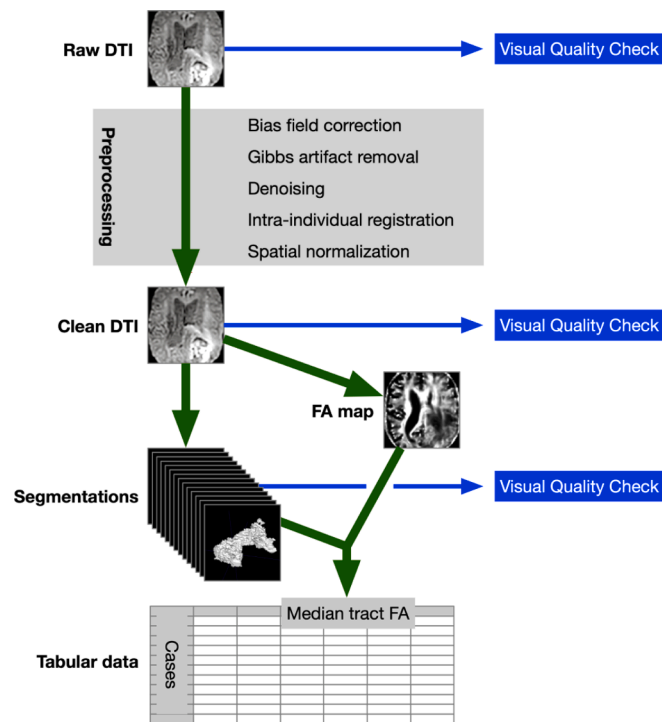


Fig. 1. Schematic illustration of the automated tractometry process. For details on the individual steps and tools used, the reader is referred to the methods section.

3. Results

3.1. Imaging and patient cohort

The initial screening yielded 71 cases, of which 28 were histopathologically diagnosed with glioma. One case underwent stereotactic biopsy and was excluded. In total, 27 patients were included and analyzed. One patient was lost to follow-up. Demographic data are summarized in Table 1. Preoperative imaging took place median one day prior to the date of surgery (range 1–5 days).

3.2. Analysis of clinical features and DTI metrics

Correlation between age and the aphasia scores revealed no association for $AS_{initial}$ ($p = 0.323$) but significance for AS_{postop} ($p = 0.017$) and AS_{3mo} ($p = 0.016$). $AS_{initial}$ showed association with AS_{postop} ($p = 0.003$) and AS_{3mo} ($p < 0.001$). Group comparisons between affected hemispheres yielded significantly higher AS in left-hemispheric tumors initially ($p = 0.024$), postoperatively ($p = 0.01$), and at three months follow-up ($p = 0.007$); there was no difference in age distribution ($p = 0.3$).

Median tract FA values are provided in Table 2. Spearman's ρ and median FA were used for correlation analysis.

Correlation between AS and median tract FA is summarized in Table 3. When corrected for demographic data and multiple testing, no median tract FA was significantly associated with preoperative or postoperative aphasia. However, preoperative median FA of the left IFOF, the left ILF, and the left UF were significantly associated with aphasia score three months after surgery. Fig. 2 shows the graphical analysis of the results. Two illustrative example cases are presented in Fig. 3.

Prediction of language impairment by median tract FA was evaluated by receiver operating characteristic (ROC) using the Monte Carlo method; results are given in Table 5. The highest AUC of 0.92 (CI: 0.75–1.0) was obtained by the left ILF for a FA cutoff of 0.29, yielding

Table 1

Dataset demographic and clinical characteristics.

| Patients [n] | 27 |
|----------------------------------|--------------|
| Age [years] (range) | 57.1 (31–83) |
| Male [%] | 59.3 |
| Left hemisphere affected [n] (%) | 14 (51.9 %) |
| Glioma grade | |
| - IV | 18 |
| - III | 5 |
| - II | 3 |
| - I | 1 |
| Tumor type | |
| - HGG | 23 |
| - LGG | 4 |
| Tumor location | |
| Left hemisphere | |
| -L temporal | 5 |
| -L parietal | 3 |
| -L frontal | 2 |
| -L basal ganglia | 1 |
| -L parietotemporal | 1 |
| -L frontoparietal | 1 |
| -L frontotemporal | 1 |
| Right hemisphere | |
| -R parietotemporal | 4 |
| -R frontal | 4 |
| -R frontoparietal | 2 |
| -R temporal | 2 |
| -R frontotemporal | 1 |
| Extent of Resection (EOR) | |
| - GTR [n] (%) | 14 (52 %) |
| - STR [n] (%) | 13 (48 %) |
| $AS_{initial}$ [mean] (SD) | 0.44 (0.64) |
| AS_{postop} [mean] (SD) | 0.74 (1.06) |
| AS_{3mo} [mean] (SD) | 0.31 (0.62) |
| $AS_{initial} \geq 1$ [n] | 17 / 10 |
| $AS_{postop} \geq 1$ [n] | 17 / 10 |
| $AS_{3mo} \geq 1$ [n] | 20 / 6 |

Table 2

FA measures for language tracts. Lat.: laterality.

| Tract | Lat. | Median FA (range) |
|-------|------|-------------------|
| AF | L | 0.36 (0.17–0.55) |
| AF | R | 0.36 (0.17–0.61) |
| IFOF | L | 0.39 (0.22–0.62) |
| IFOF | R | 0.37 (0.22–0.61) |
| ILF | L | 0.38 (0.19–0.61) |
| ILF | R | 0.40 (0.16–0.81) |
| SLF1 | L | 0.41 (0.14–0.66) |
| SLF1 | R | 0.39 (0.19–0.62) |
| SLF2 | L | 0.40 (0.16–0.63) |
| SLF2 | R | 0.37 (0.17–0.63) |
| SLF3 | L | 0.40 (0.12–0.65) |
| SLF3 | R | 0.38 (0.17–0.69) |
| UF | L | 0.42 (0.26–0.76) |
| UF | R | 0.39 (0.22–0.72) |

1.0 and 0.84 for sensitivity and specificity and 0.95 and 1.0 for positive and negative predictive value, respectively.

For analysis of individual tract contributions to language impairment, the applicability of linear regression for multivariate modeling was tested by calculating the tract variance inflation factors (VIF, Maroulides and Raykov, 2019), yielding very strong collinearity (VIF median: 340, range 57–1000). Accordingly, ANCOVA was used to explore the contribution of individual tracts to language function while controlling for the influence of other tracts and demographic data. Remarkably, intergroup comparison revealed no statistically significant differences in language tract integrity between patients with and without aphasia before or after surgery. In contrast, all left hemisphere language tracts, as well as the right AF, were independently and significantly associated with aphasia three months after surgery. Table 4

Table 3

Correlation analysis of AS and preoperative median FA using Spearman's ρ . Given are p values controlled for effects of demographic data (p) and additionally corrected for multiple tests (p_{corr}). Lat.: laterality. r: Spearman correlation coefficient. Significant correlations are in bold typeface and marked with "*" ($p < 0.05$).

| Tract | Lat. | AS_{initial} | | | AS_{postop} | | | $AS_{3\text{mo}}$ | | |
|-------|------|-----------------------|-------|-------------------|----------------------|---------------|-------------------|-------------------|---------------|-------------------|
| | | r | p | P_{corr} | r | p | P_{corr} | r | p | P_{corr} |
| AF | L | -0.31 | 0.414 | 0.725 | -0.52 | 0.136 | 0.475 | -0.60 | 0.024* | 0.065 |
| AF | R | -0.06 | 0.214 | 0.577 | -0.15 | 0.293 | 0.495 | -0.11 | 0.224 | 0.246 |
| IFOF | L | -0.31 | 0.196 | 0.577 | -0.53 | 0.039* | 0.270 | -0.57 | 0.007* | 0.026* |
| IFOF | R | -0.15 | 0.643 | 0.900 | -0.30 | 0.342 | 0.495 | -0.33 | 0.162 | 0.246 |
| ILF | L | -0.43 | 0.093 | 0.577 | -0.57 | 0.027* | 0.270 | -0.66 | 0.001* | 0.011* |
| ILF | R | -0.11 | 0.931 | 0.931 | -0.34 | 0.297 | 0.495 | -0.40 | 0.097 | 0.206 |
| SLF1 | L | -0.18 | 0.848 | 0.931 | -0.41 | 0.385 | 0.495 | -0.44 | 0.203 | 0.246 |
| SLF1 | R | 0.01 | 0.528 | 0.822 | -0.14 | 0.389 | 0.495 | -0.11 | 0.303 | 0.278 |
| SLF2 | L | -0.17 | 0.902 | 0.931 | -0.37 | 0.615 | 0.615 | -0.42 | 0.340 | 0.287 |
| SLF2 | R | -0.06 | 0.318 | 0.636 | -0.19 | 0.310 | 0.495 | -0.17 | 0.200 | 0.246 |
| SLF3 | L | -0.08 | 0.931 | 0.931 | -0.34 | 0.469 | 0.505 | -0.34 | 0.395 | 0.311 |
| SLF3 | R | -0.16 | 0.247 | 0.577 | -0.20 | 0.434 | 0.505 | -0.20 | 0.289 | 0.278 |
| UF | L | -0.28 | 0.209 | 0.577 | -0.47 | 0.134 | 0.475 | -0.58 | 0.004* | 0.023* |
| UF | R | -0.26 | 0.218 | 0.577 | -0.40 | 0.263 | 0.495 | -0.41 | 0.112 | 0.206 |

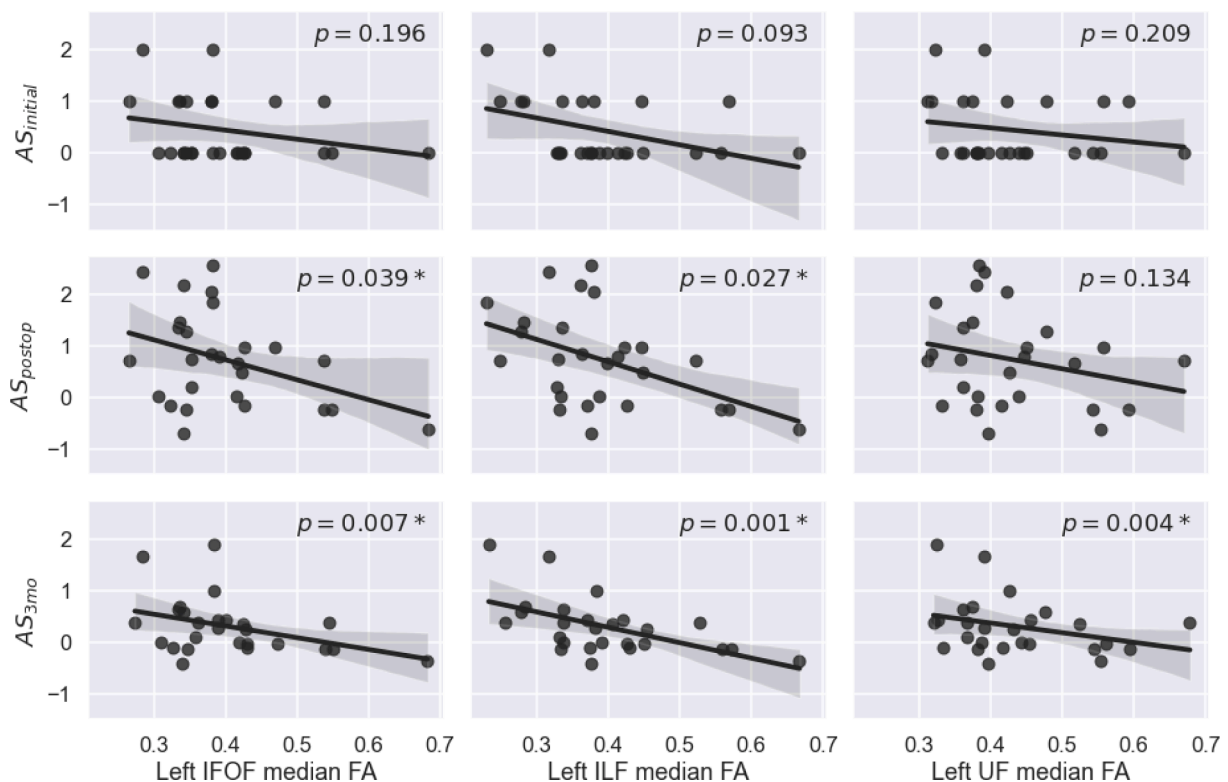


Fig. 2. Regression plots for median FA and aphasia score (AS). According to correlation with age, for AS_{postop} and $AS_{3\text{mo}}$ both axes and for AS_{initial} FA are corrected for age as covariate. Given p values are corrected for demographic data.

provides an overview of the evaluation of individual tract contributions.

4. Discussion

4.1. Subcortical language network

This study investigated associations between preoperative white matter integrity of the subcortical language system by measuring median FA and aphasia at three points in time during the clinical course of glioma resection in right-handed patients: preoperatively, postoperatively, and at three months follow-up.

The major finding is that lower white matter integrity of each tract of the left hemisphere subcortical language system and the right AF is independently associated with worse language outcome after three

months past resection, irrespective of demographic data. This result suggests that damage to each individual subcortical component of the subcortical language system in the left hemisphere and the right AF at some point can pose a limitation to the short-term language outcome. While there was no case of functional deterioration between initial presentation and three-month follow-up in this cohort, four individuals achieved language improvement. There is broad evidence that the extent of resection (Duffau, 2016; Rossi et al., 2019; Zigiotta et al., 2020) and the functional disability (Dietterle et al., 2020) are of major influence on progression-free survival and overall survival in glioma. The data collected here suggest that DTI may help to distinguish individuals who benefit from surgical resection from those who do not. Of clinical relevance is the identification of cases where gross total resection or even supratotal resection of glioma can be considered despite their

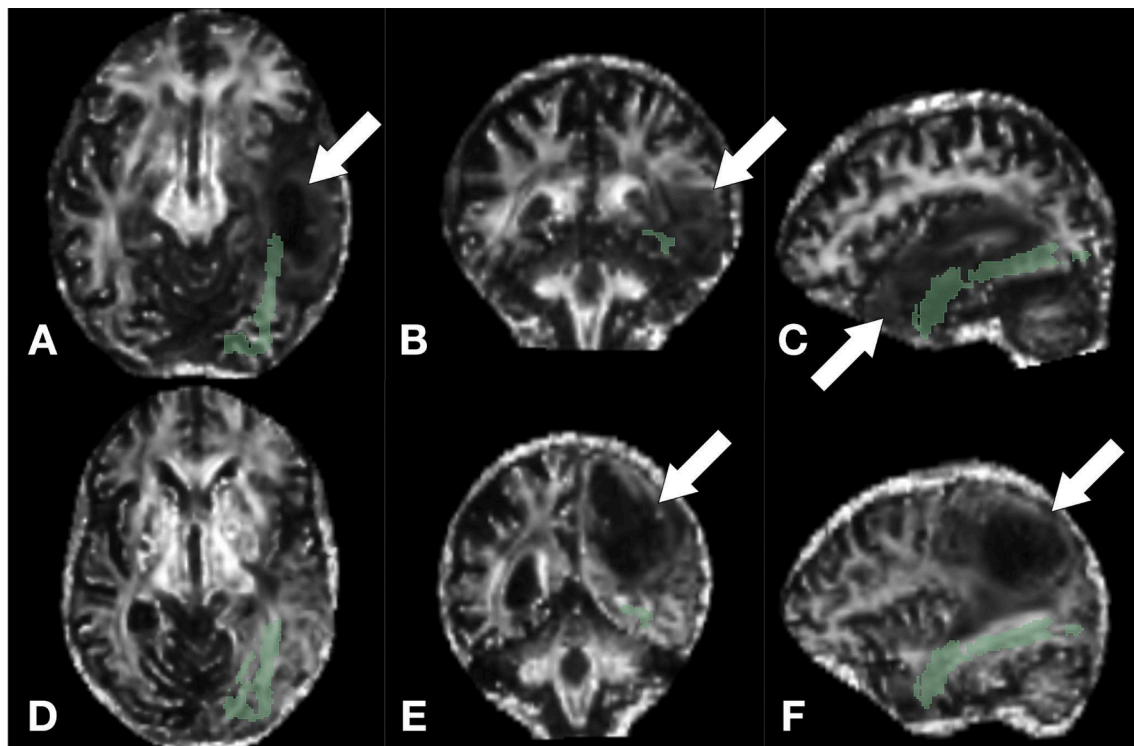


Fig. 3. FA maps of two example cases illustrating the association between ILF integrity and speech impairment. Top row (A-C): Left temporal glioblastoma (WHO IV). $AS_{\text{initial}} = 2$, $AS_{\text{postop1}} = 2$, $AS_{3\text{mo}} = 2$, median FA = 0.21, age = 62 years. Bottom row (D-F): Left parietal glioblastoma (WHO IV). $AS_{\text{initial}} = 1$, $AS_{\text{postop}} = 0$, $AS_{3\text{mo}} = 0$, median FA = 0.54, age = 64 years. Tumor locations are marked with arrows.

Table 4

Results for group comparison after dichotomization by language impairment. Given p values are corrected for the influence of demographic data and median FA values of all other tracts of the subcortical language system on a per-tract basis using ANCOVA. Lat.: Laterality. Significant correlations are in bold type-face and marked with “*” ($p < 0.05$).

| Tract | Lat. | $AS_{\text{initial}} 0 \text{ vs } \geq 1$ | | $AS_{\text{postop}} 0 \text{ vs } \geq 1$ | | $AS_{3\text{mo}} 0 \text{ vs } \geq 1$ | |
|-------|------|--|-------|---|-------|--|---------------|
| | | Median FA | p | Median FA | p | Median FA | p |
| AF | L | 0.41 0.31 | 0.750 | 0.42 0.26 | 0.773 | 0.42 0.25 | 0.021* |
| AF | R | 0.37 0.34 | 0.931 | 0.39 0.34 | 0.872 | 0.38 0.34 | 0.027* |
| IFOF | L | 0.42 0.35 | 0.659 | 0.43 0.32 | 0.760 | 0.44 0.27 | 0.015* |
| IFOF | R | 0.40 0.36 | 0.537 | 0.41 0.35 | 0.960 | 0.42 0.32 | 0.998 |
| ILF | L | 0.42 0.30 | 0.992 | 0.42 0.26 | 0.934 | 0.43 0.23 | 0.003* |
| ILF | R | 0.42 0.36 | 0.326 | 0.42 0.36 | 0.922 | 0.43 0.33 | 0.668 |
| SLF1 | L | 0.43 0.42 | 0.360 | 0.43 0.37 | 0.783 | 0.45 0.32 | 0.001* |
| SLF1 | R | 0.40 0.36 | 0.810 | 0.44 0.36 | 0.807 | 0.42 0.36 | 0.100 |
| SLF2 | L | 0.44 0.40 | 0.101 | 0.44 0.37 | 0.299 | 0.45 0.29 | 0.021* |
| SLF2 | R | 0.38 0.34 | 0.988 | 0.38 0.34 | 0.953 | 0.39 0.34 | 0.211 |
| SLF3 | L | 0.45 0.43 | 0.664 | 0.45 0.41 | 0.523 | 0.47 0.33 | 0.043* |
| SLF3 | R | 0.41 0.33 | 0.667 | 0.41 0.33 | 0.779 | 0.41 0.33 | 0.831 |
| UF | L | 0.45 0.39 | 0.497 | 0.44 0.35 | 0.625 | 0.45 0.31 | 0.041* |
| UF | R | 0.40 0.34 | 0.774 | 0.41 0.34 | 0.901 | 0.41 0.31 | 0.440 |

eloquent location when the functional language outcome can be predicted not to improve clinically relevant.

At the time of presentation, no tract contributed to language dysfunction independently from the others. This may be interpreted as the result of the language network compensating for the loss of integrity of its components. The occurrence of the clinical symptom of aphasia may therefore be the result of a network dysfunction as a whole and may be less due to disruption of individual network elements. This result reflects the common observation of low-grade glioma patients usually presenting with no or mild neurological deficits at the time of diagnosis (Desmurget et al., 2007; Duffau, 2005).

The subcortical tracts analyzed here differ regarding sensitivity and specificity for prediction of 3-month language impairment (see Table 5). Results for the left AF, for instance, suggest high sensitivity and specificity (0.90 and 0.83, respectively), while the right AF's results suggest low sensitivity and high specificity (0.35 and 1.00, respectively). Aside from the main results of this study, the quantification of the predictive power of individual tracts should be interpreted with great care because it is not only limited by the small sample size but also by the individual distribution of the lesions in this sample. Nonetheless, right-hemisphere tracts show consistently lower sensitivity and higher or equal specificity and FA cutoff values than their left-hemisphere counterparts. This pattern is most likely due to the inclusion of only right-handed individuals, leading to the language networks being lateralized probably exclusively to the left side. Accordingly, the higher FA cutoff values of right-hemisphere tracts may reflect the process of recruitment after loss of tissue integrity: While network compensation enables language function tolerating injured components, preexisting injury to contralateral homologs limits language function at the recruitment stage at a low threshold. This result is in line with studies in stroke patients highlighting the importance of the right hemisphere for language recovery (Forkel and Catani, 2018; Kiran and Thompson, 2019). In this context, the presented results are compatible with the two-level model of interindividual variability proposed by Duffau, which suggests high

Table 5

ROC-AUC analysis for prediction of language impairment at three months ($AS_{3mo} = 0$ vs $AS_{3mo} \geq 1$) by median tract FA. Lower FA values indicate higher risk for impairment. Confidence intervals, as determined by the Monte Carlo method, are given in brackets. Lat.: Laterality. Cutoff: Youden cutoff FA value. AUC: area under the receiver operating characteristic curve. PPV: positive predictive value. NPV: negative predictive value.

| Tract | Lat. | Cutoff | AUC | Sensitivity | Specificity | PPV | NPV |
|-------|------|--------|------------------|------------------|------------------|------------------|------------------|
| AF | L | 0.29 | 0.87 [0.65–1.00] | 0.90 [0.75–1.00] | 0.83 [0.43–1.00] | 0.95 [0.84–1.00] | 0.71 [0.33–1.00] |
| AF | R | 0.44 | 0.67 [0.57–0.78] | 0.35 [0.14–0.56] | 1.00 [1.00–1.00] | 1.00 [1.00–1.00] | 0.32 [0.12–0.53] |
| IFOF | L | 0.41 | 0.82 [0.72–0.92] | 0.65 [0.43–0.84] | 1.00 [1.00–1.00] | 1.00 [1.00–1.00] | 0.46 [0.19–0.73] |
| IFOF | R | 0.41 | 0.77 [0.67–0.88] | 0.55 [0.33–0.76] | 1.00 [1.00–1.00] | 1.00 [1.00–1.00] | 0.41 [0.17–0.67] |
| ILF | L | 0.29 | 0.92 [0.75–1.00] | 1.00 [1.00–1.00] | 0.84 [0.50–1.00] | 0.95 [0.85–1.00] | 1.00 [1.00–1.00] |
| ILF | R | 0.39 | 0.74 [0.52–0.90] | 0.65 [0.43–0.84] | 0.84 [0.50–1.00] | 0.93 [0.75–1.00] | 0.42 [0.15–0.73] |
| SLF1 | L | 0.34 | 0.73 [0.51–0.93] | 0.80 [0.60–0.95] | 0.67 [0.25–1.00] | 0.89 [0.73–1.00] | 0.50 [0.14–0.88] |
| SLF1 | R | 0.44 | 0.75 [0.64–0.86] | 0.50 [0.27–0.71] | 1.00 [1.00–1.00] | 1.00 [1.00–1.00] | 0.38 [0.17–0.60] |
| SLF2 | L | 0.33 | 0.76 [0.50–0.95] | 0.85 [0.68–1.00] | 0.66 [0.20–1.00] | 0.89 [0.74–1.00] | 0.57 [0.17–1.00] |
| SLF2 | R | 0.45 | 0.70 [0.59–0.81] | 0.40 [0.18–0.62] | 1.00 [1.00–1.00] | 1.00 [1.00–1.00] | 0.33 [0.12–0.57] |
| SLF3 | L | 0.38 | 0.71 [0.46–0.92] | 0.75 [0.55–0.91] | 0.66 [0.20–1.00] | 0.88 [0.71–1.00] | 0.44 [0.11–0.78] |
| SLF3 | R | 0.46 | 0.70 [0.60–0.81] | 0.40 [0.20–0.62] | 1.00 [1.00–1.00] | 1.00 [1.00–1.00] | 0.33 [0.12–0.55] |
| UF | L | 0.39 | 0.80 [0.60–0.94] | 0.75 [0.55–0.94] | 0.84 [0.50–1.00] | 0.94 [0.79–1.00] | 0.51 [0.20–0.83] |
| UF | R | 0.39 | 0.80 [0.70–0.90] | 0.60 [0.39–0.81] | 1.00 [1.00–1.00] | 1.00 [1.00–1.00] | 0.44 [0.17–0.69] |

cortical variation and low subcortical variation, thus, shifting focus from traditionally considered indispensable ‘eloquent’ cortical areas to preservation of an ‘invariant common core’ of subcortical anatomical structures in neurosurgical care (Duffau, 2017).

The extent of resection is of potential influence on the language outcome. In this cohort, there were no differences between GTR and STR regarding $AS_{initial}$ ($p = 0.21$), AS_{postop} ($p = 0.91$), AS_{3mo} ($p = 0.15$), and age ($p = 0.20$).

4.2. Components of the subcortical language network

The strongest associations, as determined by corrected correlation analysis, were found for AS_{3mo} and the perioperative FA values of the left-hemispheric IFOF, ILF, and UF. Notably, the results do not show independent associations of AF integrity with aphasia at the time of the MRI acquisition, although there are significant differences in AF integrity between patients with and without aphasia after three months (Bernal and Ardila, 2009; Glasser and Rilling, 2008). Damage to the AF, as measured by FA, has been shown to be associated with language impairment in subcortical stroke (Noh et al., 2021). Regarding patients undergoing glioma surgery, in a study including 50 patients, the preoperative infiltration of the AF was associated with worsening of language function at three-month follow-up (Tuncer et al., 2021). In contrast to the current analysis, not tract FA was used as a biomarker, but tumor-to-tract distance, which necessitates more expert post-processing steps, leading to a relevant observer bias.

A recent meta-analysis on diffusion metrics of the dual pathway tracts that pooled 1353 aphasia patients from 46 studies found that the left IFOF integrity correlated with the most language domains, including comprehension, naming, reading, and global aphasia severity (Zhang et al., 2021). FA of the left ILF and UF could also be related to the comprehension subdomains single word level and semantic association, FA of the left ILF with sentence level. In accordance with these pooled results, in the data presented here, preoperative left ILF integrity shows the strongest correlation to overall language function ($r = 0.63$), followed by left IFOF and left UF ($r = 0.62$ and $r = 0.53$, respectively). The preoperative FA of the left ILF was able to predict the occurrence of aphasia at three months with an AUC of 0.92 (FA cutoff = 0.39).

Notably, the only right-hemisphere tract independently associated with language outcome is the right AF. In a study including 16 S patients, Forkel et al. could show that adding the volume of the right AF’s long segment to a predictive language outcome model improves the predictive power (Forkel et al., 2014).

Furthermore, this analysis acknowledges the influence of age on both language outcome and measured FA. FA of all measured tracts was significantly correlated to age, as previously reported (Correia et al., 2008; Pfefferbaum et al., 2000; Pfefferbaum and Sullivan, 2003).

4.3. Technical considerations

TractSeg is an approach to DTI data assessment that consists of a pre-trained fully-connected neural network (FCNN) based on the popular encoder-decoder U-Net architecture (Wasserthal et al., 2018). It could be shown to deliver superior quality in the segmentation of white matter tracts compared with other DTI pipeline solutions, both in the presence of altered brain anatomy due to mass lesions and reduced-quality clinical scans. Visual inspection of the segmentation results confirmed the system’s ability to identify white matter tracts in areas of markedly reduced FA values due to tumor invasion or perifocal edema. The inclusion of disturbed parts of fiber bundles is crucial in order to assess their functional integrity. Streamline-based tractography may fail in areas of white matter damage, potentially excluding voxels of reduced FA in the segmentation, thus leading to underestimation of damage and introducing bias. The high number of associations between aphasia scores and left hemispheric language network elements is expected in the right-handed study population and underpins this study design’s ability to investigate language dysfunction.

The TractSeg method used in this study to acquire the tract segmentation offers the advantage of *a priori* knowledge extracted from the dataset of the Human Connectome Project (Van Essen et al., 2013) that the FCNN was primarily trained on to identify anatomical structures. This approach could be shown to be superior to other automated, streamline-based tractography methods (Wasserthal et al., 2018). Manual tractography using standardized ROIs (Fekonja et al., 2019) can be considered an alternative but requires expert knowledge and experience to achieve reliable and reproducible results, which is more time-consuming while the same general limitations apply. In contrast to the presented, purely anatomical approach, functional MRI (fMRI) may provide valuable insight into the relevance of cortical areas in the context of aphasia in neurosurgical patients and neuronal plasticity (Cargnelutti et al., 2020). While using ‘eloquent’ cortical areas identified by fMRI as seeds for tractography seems to offer the possibility of gaining more individualized anatomical tractograms, some authors state concern about the potential of both underselection of patients for surgery and increased likelihood of partial or subtotal resection due to false positives (Azad and Duffau, 2020). fMRI relies methodologically on indirect assessment of neuronal function by BOLD imaging and is heavily dependent on patient compliance during acquisition, which may be impaired in clinical use, especially in aphasic patients. Navigated transcranial magnetic stimulation (nTMS) as well offers ability to acquire tractography seeds based on individual cortical functionality by evoking potentials in a non-invasive transcranial manner and could be shown to achieve high correlation to direct cortical stimulation (Picht et al., 2013). This method was able to identify interhemispheric connectivity as a risk factor for surgery-related aphasia (Sollmann et al.,

2017).

4.4. Future directions

Further investigation of methodological aspects can be part of future work. Variation of parameters of automated tractography pipelines warrant further investigation. While leaving out steps in preprocessing as registration or noise reduction is suggestive of lower quality data for further processing, a comparison of different pipelines in a dataset of clinical MRI scans can effectively contribute to further refinement of tractography automation and explore its limits. Additional functional MRI can help create a homogenous dataset regarding hemispheric dominance in language-related white matter tracts. In order to acquire a dataset that forms a solid base for clinical decision-making and patient counseling, a larger case number and follow-up times beyond three months are of interest. Future studies should not neglect the influence of the non-dominant hemisphere on plasticity and neurological outcome.

At the immediate postoperative stage, the temporal relation to the surgical trauma along with brain edema are the most likely contributors to functional deficits. Both of these confounders, along with individual variations in the clinical course, cannot be taken into account by the current statistical analysis of preoperative data. Postoperative measurement of white matter integrity of the language network is of future interest to the investigation of brain plasticity and predictive modeling of individual rehabilitation potential.

As previously reported, FA-values were not normally distributed in the analyzed tracts (Wende et al., 2021), and median FA was the metric of choice in this study. While mean is the most frequently used metric, there is a lack of evidence on the optimal metric to summarize FA values for tract-based analysis.

4.5. Limitations

Several aspects must be considered when weighing this study's results. The present study exclusively included glioma patients, which might limit the application of its results. However, the etiology of aphasia is not a critical moderator influencing the correlations with tract FA (Zhang et al., 2021). In the surgical context, though, due to the rather infiltrative than compressive growth pattern in glioma, mass resection might more likely lead to permanent damage to integral subcortical structures than secondary malignancies, e.g., metastases. Both the small sample size and the heterogeneity of the patient cohort that results from inclusion of different glioma gradings and age groups are major limitations to the generalizability of the results. To reduce the likelihood of false positives, correction for confounder variables and multiple testing were part of the analysis.

Aphasia severity was scored in retrospective according to medical documentation of the neurological status, which opens opportunity for investigator bias. However, to cope with investigator bias, the scoring investigator was blinded to medical imaging. The semiquantitative AS scale reflects communication ability. It does not distinguish between fluent and nonfluent aphasia and is unsuitable for differential diagnosis of impaired speech. AS is prone to inaccuracies regarding disabilities unrelated to white matter integrity. A high percentage of patients suffering from brain tumors show cognitive disturbances (Taphoorn and Klein, 2004). The detection of more subtle cognitive symptoms, which may include language processing, is highly dependent on the extent of neuropsychological testing. AS, as used in this study, is likely to underestimate the frequency of language disturbance at the preoperative time point but presents as a feasible tool for quantification of the individual development and for interindividual comparison in this context.

This study only included right-handed patients, reflecting about 90 % of the general population (Papadatou-Pastou et al., 2020). The relation of handedness to language lateralization is under active research and discussion. Lateralization of subcortical pathways is, as of yet, not fully attributable to right- and left-handers (Budisavljevic et al., 2021).

DTI data acquisition in the presented data was optimized for short acquisition time, potentially conceding image quality. White matter damage due to glioma infiltration, brain edema, or otherwise leads to a decrease in local FA values, which hinders the identification of voxels that belong to an anatomical tract. While this is the data of interest in our investigation, it is also an inherent limitation of the tractography methodology based on diffusion imaging.

5. Conclusion

This study was designed to explore the relationship of preoperative white matter integrity in the subcortical language system to clinically apparent aphasia preoperatively, postoperatively, and three months after glioma resection in right-handed patients. The left AF, IFOF, ILF, and UF, as well as right AF, showed significantly reduced preoperative white matter integrity in patients with aphasia three months after surgery compared to patients with patent language function. The data presented here are in line with current literature, underlining the importance of integrity of critical subcortical structures for patent network function and preservation of plasticity in rehabilitation of language function. Time-efficient 32-directions DTI protocols, together with reproducible FCNN-driven tract segmentation, can provide a solid base for future studies investigating relations of language network features with the clinical outcome.

Funding

This research received no external funding.

CRediT authorship contribution statement

Gordian Prasse: Conceptualization, Formal analysis, Methodology, Software, Visualization, Writing – original draft, Investigation, Data curation, Project administration. **Hans-Jonas Meyer:** Writing – review & editing. **Cordula Scherlach:** Resources, Supervision, Writing – review & editing. **Jens Maybaum:** Project administration. **Anastasia Hoffmann:** Investigation, Data curation, Writing – review & editing. **Johannes Kasper:** Writing – review & editing. **Michael Karl Fehrenbach:** Writing – review & editing. **Florian Wilhelmy:** Resources, Writing – review & editing. **Jürgen Meixensberger:** Supervision, Writing – review & editing. **Karl-Titus Hoffmann:** Project administration, Supervision, Writing – review & editing. **Tim Wende:** Conceptualization, Formal analysis, Methodology, Software, Investigation, Data curation, Validation, Project administration, Supervision, Writing – review & editing.

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.

Data availability

The datasets generated during and analyzed in this study are not publicly available due to the privacy of research participants but are available from the corresponding author on reasonable request.

References

- Almairac, F., Herbet, G., Moritz-Gasser, S., de Champfleury, N.M., Duffau, H., 2015. The left inferior fronto-occipital fasciculus subserves language semantics: a multilevel lesion study. *Brain Struct. Funct.* 220, 1983–1995. <https://doi.org/10.1007/s00429-014-0773-1>.
- Azad, T.D., Duffau, H., 2020. Limitations of functional neuroimaging for patient selection and surgical planning in glioma surgery. *Neurosurg. Focus* 48, E12. <https://doi.org/10.3171/2019.11.FOCUS19769>.

- Basser, P.J., Pierpaoli, C., 1996. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *J. Magn. Reson. B* 111, 209–219. <https://doi.org/10.1006/jmrb.1996.0086>.
- Bello, L., Gambini, A., Castellano, A., Carrabba, G., Acerbi, F., Fava, E., Giussani, C., Cadioli, M., Blasi, V., Casarotti, A., Papagno, C., Gupta, A.K., Gaini, S., Scotti, G., Falini, A., 2008. Motor and language DTI Fiber Tracking combined with intraoperative subcortical mapping for surgical removal of gliomas. *Neuroimage* 39, 369–382. <https://doi.org/10.1016/j.neuroimage.2007.08.031>.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Series B Stat. Methodol.* 57, 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>.
- Bernal, B., Ardila, A., 2009. The role of the arcuate fasciculus in conduction aphasia. *Brain* 132, 2309–2316. <https://doi.org/10.1093/brain/awp206>.
- Brott, T., Adams Jr, H.P., Olinger, C.P., Marler, J.R., Barsan, W.G., Biller, J., Spilker, J., Holleran, R., Eberle, R., Hertzberg, V., 1989. Measurements of acute cerebral infarction: a clinical examination scale. *Stroke* 20, 864–870. <https://doi.org/10.1161/01.str.20.7.864>.
- Budisavljevic, S., Castiello, U., Begliomini, C., 2021. Handedness and White Matter Networks. *Neuroscientist* 27, 88–103. <https://doi.org/10.1177/1073858420937657>.
- Cargnelutti, E., Ius, T., Skrap, M., Tomasino, B., 2020. What do we know about pre- and postoperative plasticity in patients with glioma? A review of neuroimaging and intraoperative mapping studies. *Neuroimage Clin* 28, 102435. <https://doi.org/10.1016/j.nicl.2020.102435>.
- Catani, M., Mesulam, M., 2008. The arcuate fasciculus and the disconnection theme in language and aphasia: history and current state. *Cortex* 44, 953–961. <https://doi.org/10.1016/j.cortex.2008.04.002>.
- Chang, E.F., Raygor, K.P., Berger, M.S., 2015. Contemporary model of language organization: an overview for neurosurgeons. *J. Neurosurg.* 122, 250–261. <https://doi.org/10.3171/2014.10.JNS132647>.
- Correia, S., Lee, S.Y., Voorn, T., Tate, D.F., Paul, R.H., Zhang, S., Salloway, S.P., Malloy, P.F., Laidlaw, D.H., 2008. Quantitative tractography metrics of white matter integrity in diffusion-tensor MRI. *Neuroimage* 42, 568–581. <https://doi.org/10.1016/j.neuroimage.2008.05.02>.
- Desmurget, M., Bonnetblanc, F., Duffau, H., 2007. Contrasting acute and slow-growing lesions: a new door to brain plasticity. *Brain* 130, 898–914. <https://doi.org/10.1093/brain/awl300>.
- Dick, A.S., Bernal, B., Tremblay, P., 2014. The language connectome: new pathways, new concepts. *Neuroscientist* 20, 453–467. <https://doi.org/10.1177/1073858413513502>.
- Dietterle, J., Wende, T., Wilhelm, F., Eisenlöffel, C., Jähne, K., Taubenheim, S., Arlt, F., Meixensberger, J., 2020. The prognostic value of peri-operative neurological performance in glioblastoma patients. *Acta Neurochir.* 162, 417–425. <https://doi.org/10.1007/s00701-019-04136-4>.
- Duffau, H., 2005. Lessons from brain mapping in surgery for low-grade glioma: insights into associations between tumour and brain plasticity. *The Lancet Neurology* 4 (8), 476–486.
- Duffau, H., 2016. Long-term outcomes after supratotal resection of diffuse low-grade gliomas: a consecutive series with 11-year follow-up. *Acta Neurochir.* 158, 51–58. <https://doi.org/10.1007/s00701-015-2621-3>.
- Duffau, H., 2017. A two-level model of interindividual anatomo-functional variability of the brain and its implications for neurosurgery. *Cortex* 86, 303–313. <https://doi.org/10.1016/j.cortex.2015.12.009>.
- Duffau, H., Gatignol, P., Moritz-Gasser, S., Mandonnet, E., 2009. Is the left uncinate fasciculus essential for language? A cerebral stimulation study. *J. Neurol.* 256, 382–389. <https://doi.org/10.1007/s00415-009-0053-9>.
- Essayed, W.I., Zhang, F., Unadkat, P., Cosgrove, G.R., Golby, A.J., O'Donnell, L.J., 2017. White matter tractography for neurosurgical planning: A topography-based review of the current state of the art. *Neuroimage Clin* 15, 659–672. <https://doi.org/10.1016/j.nicl.2017.06.011>.
- Fadnavis, S., Batson, J., Garyfallidis, E., 2020. Patch2Self: Denoising Diffusion MRI with Self-Supervised Learning. *arXiv [cs.LG]*.
- Fekonja, L., Wang, Z., Bährend, I., Rosenstock, T., Rösler, J., Wallmeroth, L., Vajkoczy, P., Picht, T., 2019. Manual for clinical language tractography. *Acta Neurochir.* 161, 1125–1137. <https://doi.org/10.1007/s00701-019-03899-0>.
- Forkel, S.J., Catani, M., 2018. Lesion mapping in acute stroke aphasia and its implications for recovery. *Neuropsychologia* 115, 88–100. <https://doi.org/10.1016/j.neuropsychologia.2018.03.036>.
- Forkel, S.J., Thiebaut de Schotten, M., Dell'Acqua, F., Kalra, L., Murphy, D.G.M., Williams, S.C.R., Catani, M., 2014. Anatomical predictors of aphasia recovery: a tractography study of bilateral perisylvian language networks. *Brain* 137, 2027–2039. <https://doi.org/10.1093/brain/awu113>.
- Garyfallidis, E., Brett, M., Amirbekian, B., Rokem, A., Van Der Walt, S., Descoteaux, M., Nimmo-Smith, I., 2014. Dipy, a library for the analysis of diffusion MRI data. *Front. Neuroinform.* 8, 8. <https://doi.org/10.3389/fninf.2014.00008>.
- Geranmayeh, F., Brownset, S.L.E., Wise, R.J.S., 2014. Task-induced brain activity in aphasic stroke patients: what is driving recovery? *Brain* 137, 2632–2648. <https://doi.org/10.1093/brain/awu163>.
- Glasser, M.F., Rilling, J.K., 2008. DTI tractography of the human brain's language pathways. *Cereb. Cortex* 18, 2471–2482. <https://doi.org/10.1093/cercor/bhn011>.
- Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402. <https://doi.org/10.1038/nrn2113>.
- Ille, S., Kulchyska, N., Sollmann, N., Wittig, R., Beurskens, E., Butenschoen, V.M., Ringel, F., Vajkoczy, P., Meyer, B., Picht, T., Krieg, S.M., 2016. Hemispheric language dominance measured by repetitive navigated transcranial magnetic stimulation and postoperative course of language function in brain tumor patients. *Neuropsychologia* 91, 50–60. <https://doi.org/10.1016/j.neuropsychologia.2016.07.025>.
- Jenkinson, M., Smith, S., 2001. A global optimisation method for robust affine registration of brain images. *Med. Image Anal.* 5, 143–156. [https://doi.org/10.1016/s1361-8415\(01\)00036-6](https://doi.org/10.1016/s1361-8415(01)00036-6).
- Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17, 825–841. [https://doi.org/10.1016/s1053-8119\(02\)91132-8](https://doi.org/10.1016/s1053-8119(02)91132-8).
- Kiran, S., Thompson, C.K., 2019. Neuroplasticity of Language Networks in Aphasia: Advances, Updates, and Future Challenges. *Front. Neurol.* 10, 295. <https://doi.org/10.3389/fneur.2019.00295>.
- Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P.W., 2010. elastix: a toolbox for intensity-based medical image registration. *IEEE Trans. Med. Imaging* 29, 196–205. <https://doi.org/10.1109/TMI.2009.2035616>.
- Krieg, S.M., Sollmann, N., Hauck, T., Ille, S., Meyer, B., Ringel, F., 2014. Repeated mapping of cortical language sites by preoperative navigated transcranial magnetic stimulation compared to repeated intraoperative DCS mapping in awake craniotomy. *BMC Neurosci.* 15 (1) <https://doi.org/10.1186/1471-2202-15-20>.
- Le Bihan, D., Mangin, J.F., Poupon, C., Clark, C.A., Pappata, S., Molko, N., Chabriat, H., 2001. Diffusion tensor imaging: concepts and applications. *J. Magn. Reson. Imaging* 13, 534–546. <https://doi.org/10.1002/jmri.1076>.
- Mandonnet, E., Nouet, A., Gatignol, P., Capelle, L., Duffau, H., 2007. Does the left inferior longitudinal fasciculus play a role in language? A brain stimulation study. *Brain* 130, 623–629. <https://doi.org/10.1093/brain/awl361>.
- Marcoulides, K.M., Raykov, T., 2019. Evaluation of variance inflation factors in regression models using latent variable modeling methods. *Educ. Psychol. Meas.* 79, 874–882. <https://doi.org/10.1177/0013164418817803>.
- McKinney, W., 2010. Data Structures for Statistical Computing in Python, in: Proceedings of the 9th Python in Science Conference. Presented at the Python in Science Conference, SciPy. <https://doi.org/10.25080/majora-92bf1922-00a>.
- Neto Henriques, R., 2018. Advanced methods for diffusion MRI data analysis and their application to the healthy ageing brain. <https://doi.org/10.17863/CAM.29356>.
- Noh, J.S., Lee, S., Na, Y., Cho, M., Hwang, Y.M., Tae, W.-S., Pyun, S.-B., 2021. Integrity of arcuate fasciculus is a good predictor of language impairment after subcortical stroke. *J. Neurolinguistics* 58, 100968. <https://doi.org/10.1016/j.jneuroling.2020.100968>.
- Papadatou-Pastou, M., Ntolka, E., Schmitz, J., Martin, M., Munafò, M.R., Ocklenburg, S., Paracchini, S., 2020. Human handedness: A meta-analysis. *Psychol. Bull.* 146, 481–524. <https://doi.org/10.1037/bul0000229>.
- Pfefferbaum, A., Sullivan, E.V., 2003. Increased brain white matter diffusivity in normal adult aging: relationship to anisotropy and partial voluming. *Magn. Reson. Med.* 49, 953–961. <https://doi.org/10.1002/mrm.10452>.
- Pfefferbaum, A., Sullivan, E.V., Hedehus, M., Lim, K.O., Adalsteinsson, E., Moseley, M., 2000. Age-related decline in brain white matter anisotropy measured with spatially corrected echo-planar diffusion tensor imaging. *Magn. Reson. Med.* 44, 259–268. [https://doi.org/10.1002/1522-2594\(200008\)44:2<259::aid-mrm13>3.0.co;2-6](https://doi.org/10.1002/1522-2594(200008)44:2<259::aid-mrm13>3.0.co;2-6).
- Picht, T., Krieg, S.M., Sollmann, N., Rösler, J., Niraula, B., Neuvonen, T., Savolainen, P., Lioumis, P., Mäkelä, J.P., Deletis, V., Meyer, B., Vajkoczy, P., Ringel, F., 2013. A comparison of language mapping by preoperative navigated transcranial magnetic stimulation and direct cortical stimulation during awake surgery. *Neurosurgery* 72, 808–819. <https://doi.org/10.1227/NEU.0b013e3182889e01>.
- Rossi, M., Ambrogio, F., Gay, L., Gallucci, M., Conti Nibali, M., Leonetti, A., Puglisi, G., Sciorino, T., Howells, H., Riva, M., Pessina, F., Navarra, P., Franzese, C., Simonelli, M., Rudà, R., Bello, L., 2019. Is supratotal resection achievable in low-grade gliomas? Feasibility, putative factors, safety, and functional outcome. *J. Neurosurg.* 132, 1692–1705. <https://doi.org/10.3171/2019.2.JNS183408>.
- Sanai, N., Mirzadeh, Z., Berger, M.S., 2008. Functional outcome after language mapping for glioma resection. *N. Engl. J. Med.* 358, 18–27. <https://doi.org/10.1056/NEJMoa067819>.
- Saur, D., Kreher, B.W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.-S., Umarova, R., Musso, M., Glauche, V., Abel, S., Huber, W., Rijntjes, M., Hennig, J., Weiller, C., 2008. Ventral and dorsal pathways for language. *Proc. Natl. Acad. Sci. U. S. A.* 105, 18035–18040. <https://doi.org/10.1073/pnas.0805234105>.
- Seabold, S., Perktold, J., 2010. Statsmodels: Econometric and statistical modeling with python, in: Proceedings of the 9th Python in Science Conference. Presented at the Python in Science Conference, SciPy. <https://doi.org/10.25080/majora-92bf1922-011>.
- Shamonin, D.P., Bron, E.E., Lelieveldt, B.P.F., Smits, M., Klein, S., Staring, M., Alzheimer's Disease Neuroimaging Initiative., 2013. Fast parallel image registration on CPU and GPU for diagnostic classification of Alzheimer's disease. *Front. Neuroinform.* 7, 50. <https://doi.org/10.3389/fninf.2013.00050>.
- Sollmann, N., Ille, S., Hauck, T., Maurer, S., Negwer, C., Ringel, F., Meyer, B., Krieg, S.M., 2015. The impact of preoperative language mapping by repetitive navigated transcranial magnetic stimulation on the clinical course of brain tumor patients. *BMC Cancer* 15, 261. <https://doi.org/10.1186/s12885-015-1299-5>.
- Sollmann, N., Negwer, C., Tussis, L., Hauck, T., Ille, S., Maurer, S., Giglhuber, K., Bauer, J.S., Ringel, F., Meyer, B., Others., 2017. Interhemispheric connectivity revealed by diffusion tensor imaging fiber tracking derived from navigated transcranial magnetic stimulation maps as a sign of language function at risk in patients with brain tumors. *J. Neurosurg.* 126, 222–233.
- Taphoorn, M.J.B., Klein, M., 2004. Cognitive deficits in adult patients with brain tumours. *Lancet Neurol.* 3, 159–168. [https://doi.org/10.1016/S1474-4422\(04\)00680-5](https://doi.org/10.1016/S1474-4422(04)00680-5).
- Tournier, J.-D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Christiaens, D., Jeurissen, B., Yeh, C.-H., Connelly, A., 2019. MRtrix3: A fast, flexible

- and open software framework for medical image processing and visualisation. bioRxiv. <https://doi.org/10.1101/551739>.
- Tuncer, M.S., Salvati, L.F., Grittner, U., Hardt, J., Schilling, R., Bährend, I., Silva, L.L., Fekonja, L.S., Faust, K., Vajkoczy, P., Rosenstock, T., Picht, T., 2021. Towards a tractography-based risk stratification model for language area associated gliomas. *Neuroimage Clin* 29, 102541. <https://doi.org/10.1016/j.nicl.2020.102541>.
- Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C., 2013. N4ITK: Improved N3 Bias Correction. *IEEE Trans. Med. Imaging* 29, 1310–1320. <https://doi.org/10.1109/TMI.2010.2046908>.
- Vallat, R., 2018. Pingouin: statistics in Python. *Journal of Open Source Software* 3, 1026. <https://doi.org/10.21105/joss.01026>.
- Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E.J., Yacoub, E., Ugurbil, K., 2013. The WU-Minn Human Connectome Project: an overview. *Neuroimage* 80, 62–79.
- Wang, X., Pathak, S., Stefanescu, L., Yeh, F.-C., Li, S., Fernandez-Miranda, J.C., 2016. Subcomponents and connectivity of the superior longitudinal fasciculus in the human brain. *Brain Struct. Funct.* 221, 2075–2092. <https://doi.org/10.1007/s00429-015-1028-5>.
- Waskom, M., 2021. seaborn: statistical data visualization. *J. Open Source Softw.* 6, 3021. <https://doi.org/10.21105/joss.03021>.
- Wasserthal, J., Neher, P., Maier-Hein, K.H., 2018. TractSeg - Fast and accurate white matter tract segmentation. *Neuroimage* 183, 239–253. <https://doi.org/10.1016/j.neuroimage.2018.07.070>.
- Wende, T., Kasper, J., Wilhelmy, F., Dietel, E., Hamerla, G., Scherlach, C., Meixensberger, J., Fehrenbach, M.K., 2021. Assessment of a Reliable Fractional Anisotropy Cutoff in Tractography of the Corticospinal Tract for Neurosurgical Patients. *Brain Sci* 11 (5), 650.
- Zhang, J., Zhong, S., Zhou, L., Yu, Y., Tan, X., Wu, M., Sun, P., Zhang, W., Li, J., Cheng, R., Wu, Y., Yu, Y., Ye, X., Luo, B., 2021. Correlations between Dual-Pathway White Matter Alterations and Language Impairment in Patients with Aphasia: A Systematic Review and Meta-analysis. *Neuropsychol. Rev.* 31, 402–418. <https://doi.org/10.1007/s11065-021-09482-8>.
- Zigiotto, L., Annicchiarico, L., Corsini, F., Vitali, L., Falchi, R., Dalpiaz, C., Rozzanigo, U., Barbareschi, M., Avesani, P., Papagno, C., Duffau, H., Chioffi, F., Sarubbo, S., 2020. Effects of supra-total resection in neurocognitive and oncological outcome of high-grade gliomas comparing asleep and awake surgery. *J. Neurooncol.* 148, 97–108. <https://doi.org/10.1007/s11060-020-03494-9>.