



OPEN

## Multi-scale image analysis and prediction of visual field defects after selective amygdalohippocampectomy

Bastian David<sup>1</sup>, Jasmine Eberle<sup>1,2</sup>, Daniel Delev<sup>3</sup>, Jennifer Gaubatz<sup>1</sup>, Conrad C. Prillwitz<sup>1</sup>, Jan Wagner<sup>4</sup>, Jan-Christoph Schoene-Bake<sup>5</sup>, Guido Luechters<sup>6</sup>, Alexander Radbruch<sup>7</sup>, Bettina Wabbels<sup>8</sup>, Johannes Schramm<sup>9</sup>, Bernd Weber<sup>10</sup>, Rainer Surges<sup>1</sup>, Christian E. Elger<sup>1</sup> & Theodor Rüber<sup>1,11,12</sup>✉

Selective amygdalohippocampectomy is an effective treatment for patients with therapy-refractory temporal lobe epilepsy but may cause visual field defect (VFD). Here, we aimed to describe tissue-specific pre- and postoperative imaging correlates of the VFD severity using whole-brain analyses from voxel- to network-level. Twenty-eight patients with temporal lobe epilepsy underwent pre- and postoperative MRI (T1-MPRAGE and Diffusion Tensor Imaging) as well as kinetic perimetry according to Goldmann standard. We probed for whole-brain gray matter (GM) and white matter (WM) correlates of VFD using voxel-based morphometry and tract-based spatial statistics, respectively. We furthermore reconstructed individual structural connectomes and conducted local and global network analyses. Two clusters in the bihemispheric middle temporal gyri indicated a postsurgical GM volume decrease with increasing VFD severity (FWE-corrected  $p < 0.05$ ). A single WM cluster showed a fractional anisotropy decrease with increasing severity of VFD in the ipsilesional optic radiation (FWE-corrected  $p < 0.05$ ). Furthermore, patients with (vs. without) VFD showed a higher number of postoperative local connectivity changes. Neither in the GM, WM, nor in network metrics we found preoperative correlates of VFD severity. Still, in an explorative analysis, an artificial neural network meta-classifier could predict the occurrence of VFD based on presurgical connectomes above chance level.

Temporal lobe epilepsy (TLE) is the most common focal epilepsy affecting 25% to 40% of all epilepsy patients<sup>1,2</sup>. Approximately 40% of TLE patients are pharmaco-resistant and the superiority of epilepsy surgery as compared to pharmacotherapy has been repetitively shown<sup>3–5</sup>. The two most commonly applied procedures are the anterior temporal lobectomy and selective amygdalohippocampectomy (sAH)<sup>6</sup>. While several studies have reported no difference in patients becoming postoperatively seizure-free<sup>6</sup>, the theoretical advantage of sAH lies in the postoperatively lower cognitive loss due to preservation of the temporal cortex and the underlying white matter. sAH may be performed using a transylvian<sup>7</sup> or a subtemporal approach<sup>8</sup>. The advantage of the subtemporal as compared to the transylvian approach is that a partial disconnection of the temporal stem can be avoided. Here, a portion of the fusiform gyrus is removed to gain access to the mesial temporal structures. Between 60 and 80% of patients are rendered seizure-free after sAH<sup>4,5,9,10</sup>. However, visual field deficits (VFD)

<sup>1</sup>Department of Epileptology, University Hospital Bonn, Sigmund-Freud-Str. 25, 53127 Bonn, Germany. <sup>2</sup>Clinic for Neurology and Palliative Medicine, Municipal Hospital Köln-Merheim, Cologne, Germany. <sup>3</sup>Department of Neurosurgery, RWTH University Aachen, Aachen, Germany. <sup>4</sup>Department of Neurology, University of Ulm and Universitäts- and Rehabilitationskliniken Ulm, Ulm, Germany. <sup>5</sup>Department of Pediatric Kidney, Liver and Metabolic Diseases, Hannover Medical School, Hannover, Germany. <sup>6</sup>Center for Development Research, University of Bonn, Bonn, Germany. <sup>7</sup>Department of Neuroradiology, University Hospital Bonn, Bonn, Germany. <sup>8</sup>Department of Ophthalmology, University Hospital Bonn, Bonn, Germany. <sup>9</sup>Medical Faculty, University Hospital Bonn, Bonn, Germany. <sup>10</sup>Institute of Experimental Epileptology and Cognition Research, University Hospital Bonn, Bonn, Germany. <sup>11</sup>Department of Neurology, Epilepsy Center Frankfurt Rhine-Main, Goethe University Frankfurt, Frankfurt am Main, Germany. <sup>12</sup>Center for Personalized Translational Epilepsy Research (CePTER), Goethe-University Frankfurt, Frankfurt am Main, Germany. ✉email: theodor.rueber@ukbonn.de

have been reported to postoperatively occur in 15% to 100% of patients undergoing resective surgery in the temporal lobe<sup>11–13</sup>, precluding the ability to drive a car even in patients who are permanently seizure-free. These VFD, which usually manifest as contralateral homonymous upper quadrant anopia, often referred to as ‘pie in the sky’<sup>14</sup>, result from the spatial proximity of the Meyer’s loop (ML) to the resection cavity in the temporal lobe. Patients who have undergone sAH using a subtemporal approach have been reported to show a lower rate of postoperative VFD as compared to those who have undergone sAH using a transylvian approach<sup>15</sup>. With an estimated number of 100,000 TLE patients being potentially amenable to epilepsy surgery every year only in the United States<sup>4</sup>, an understanding of how the damage occurs and on how to prevent it during surgery is of vast clinical relevance. Beside its relevance on clinical grounds, the observation of structural changes following neurosurgery is of interest in the context of neurodegeneration and neuroplasticity<sup>16</sup> and has been aspired by the authors of several studies: while it has been generally asserted, that VFDs are more likely after epilepsy surgery in the left temporal lobe due to hemispherical asymmetry of the ML<sup>12</sup>, various studies have employed Diffusion Tensor Imaging and tractography to delineate the ML and anatomically relate it to either anatomical landmarks or to the resection cavity<sup>17–22</sup>. Resection size and the Euclidean distance between the temporal pole and the most anterior part of the ML are the most promising correlates/predictors of postoperative VFD<sup>17,22</sup>. However, a multi-modal, data-driven approach to identify structural underpinnings of perioperative VFD in different tissue types and on multiple scales has not yet been taken. Here, we apply several whole-brain analyses to imaging and perimetry data of patients undergoing sAH to probe for presurgical gray (GM) and white matter (WM) predictors of postoperative VFD. Furthermore, we aim to describe direct as well as indirect effects of the surgical intervention on both the level of voxels as well as structural connectomes and investigate how they relate to VFD. In an explorative analysis, we further aim to preoperatively predict postoperative VFD using a combination of structural connectomics and supervised machine learning algorithms.

## Results

**Clinical group differences.** Of the 28 patients included in the study, 21 showed postsurgical VFD (11 incomplete homonymous quadrantanopia, 6 complete homonymous quadrantanopia, 4 incomplete homonymous hemianopia) while the other 7 showed no VFD in the automated Goldmann perimetry. VFD and no VFD patient groups did not differ significantly in age, gender, duration of epilepsy or surgery-scan interval (all  $p > 0.05$ ). Subtemporal and transylvian surgery procedure groups did not differ in the mentioned demographic variables ( $p > 0.05$ ). However, patients who underwent sAH using a subtemporal access showed less severe VFDs ( $p < 0.05$ ). Postoperative resection size and the preoperative Euclidean distance between the temporal pole and the most anterior part of Meyer’s loop were tested in a regression analyses with the presence of VFD as dependent variable yielding non-significant results (both  $p > 0.45$ ).

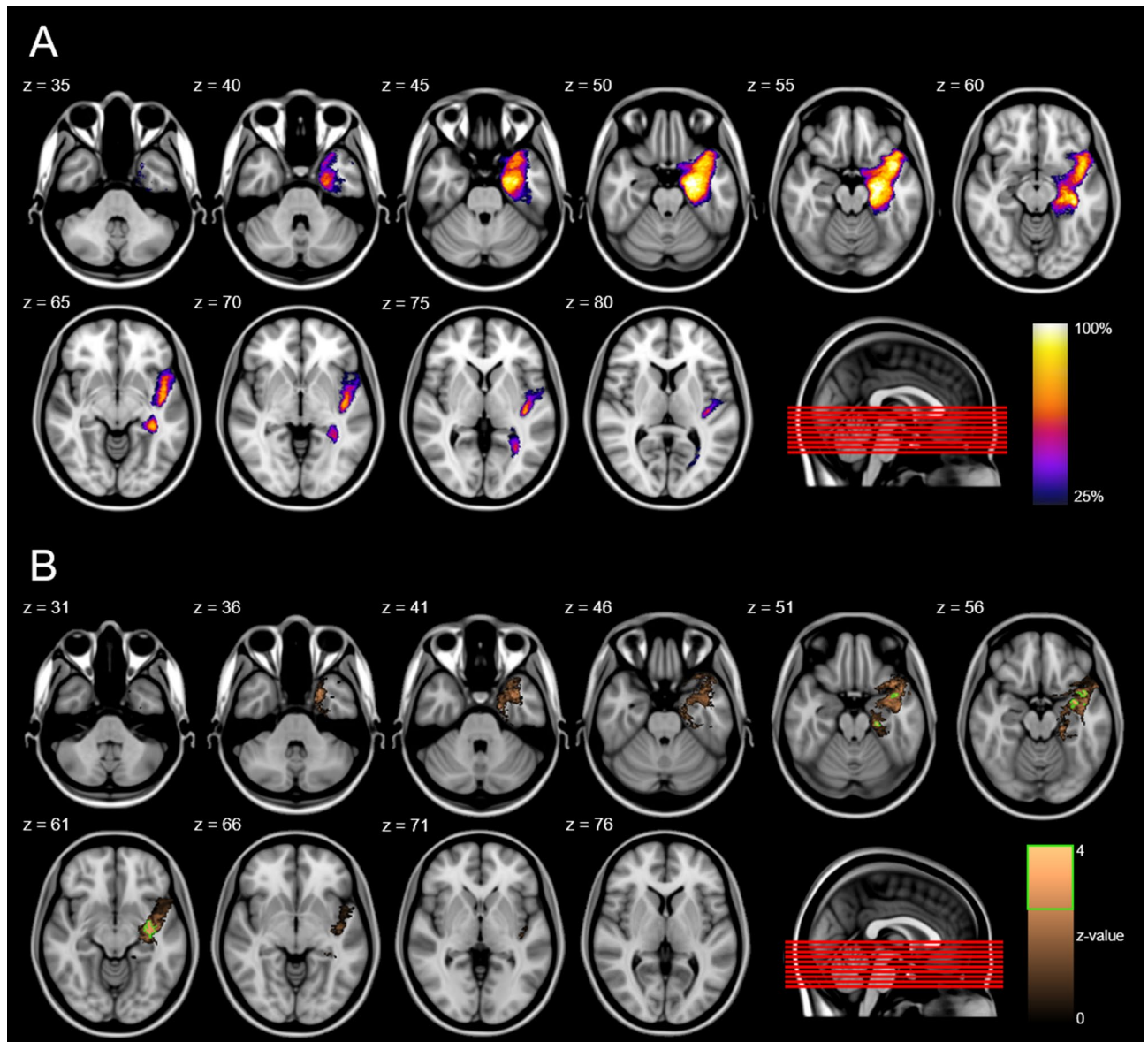
**VLSM results.** Our VLSM analysis including all 28 manual lesion masks showed a significant association between lesioned voxels and postsurgical VFD severity in the ipsilesional external capsule and the uncinate fasciculus (FWE-corrected  $p < 0.05$ ; Fig. 1B; volume = 423 mm<sup>3</sup>). Two smaller significant clusters were found in the GM of the ipsilesional temporal pole (FWE-corrected  $p < 0.05$ ; volume = 41 mm<sup>3</sup>) and parahippocampal gyrus (FWE-corrected  $p < 0.05$ ; volume = 25 mm<sup>3</sup>).

**VBM results.** Using a permutation-based paired *t*-test comparing pre- and postsurgical T1-weighted scans for the subgroup of patients who underwent a transylvian surgical procedure ( $n = 18$ ), we found a significant decrease in ipsilesional GM volume in our patient cohort (Fig. 2A). The largest cluster extended over large parts of subcortical structures, namely the ipsilesional caudate, putamen, pallidum, and thalamus. Apart from subcortical structures, the cluster of postsurgical GM decrease furthermore covered parts of the insular cortex as well as the inferior temporal and middle temporal gyrus (all FWE-corrected  $p < 0.001$ ). The opposite contrast of a postsurgical GM volume increase resulted in a cluster covering the ipsilesional inferior frontal gyrus which, however, did not survive FWE-correction (uncorrected  $p < 0.001$ ). Clusters remained significant after exclusion of patients with a surgery-scan-interval larger than 12 months (see Suppl. Fig. S2A).

In a second analysis, probing for a linear relationship between the degree of VFD and postsurgical GM volume, we found two significant clusters in the posterior division of both the ipsi- and the contralesional middle temporal gyrus showing a GM volume decrease with increasing VFD degree (FWE-corrected  $p < 0.05$ ; Fig. 2B). This linear relationship can be described for both the transylvian and subtemporal patient cohort. The opposite contrast as well as the same contrasts applied to the presurgical T1 scans did not yield any significant results.

**TBSS results.** Parallel to the VBM analysis, we conducted a permutation-based paired *t*-test comparing pre- and postsurgical FA of the transylvian subgroup. Similar to the GM changes described above, we found significantly decreased FA-values in large parts of the ipsilesional temporal and inferior frontal lobe (FWE-corrected  $p < 0.05$ ; Fig. 3A). Clusters extended over the inferior and superior longitudinal and fronto-occipital fasciculus, as well as the anterior thalamic radiation and the uncinate fasciculus. The opposite contrast, however, yielded a significant cluster of postsurgically increased FA in the ipsilesional corona radiata, including especially the corticospinal tract. This cluster, however, did not survive FWE-correction (uncorrected  $p < 0.001$ ). All clusters remained significant after exclusion of patients with a surgery-scan-interval larger than 12 months (see Suppl. Fig. S2B).

Testing for a linear relationship between FA and the degree of VFD, we found a single cluster showing an FA decrease with increasing extent of the VFD (FWE-corrected  $p < 0.05$ ; Fig. 3B). The cluster coincided with the location of the sagittal stratum within the trajectory of the ipsilesional optic radiation as determined by probabilistic tractography. The linear relationship can be appreciated in both the transylvian and subtemporal

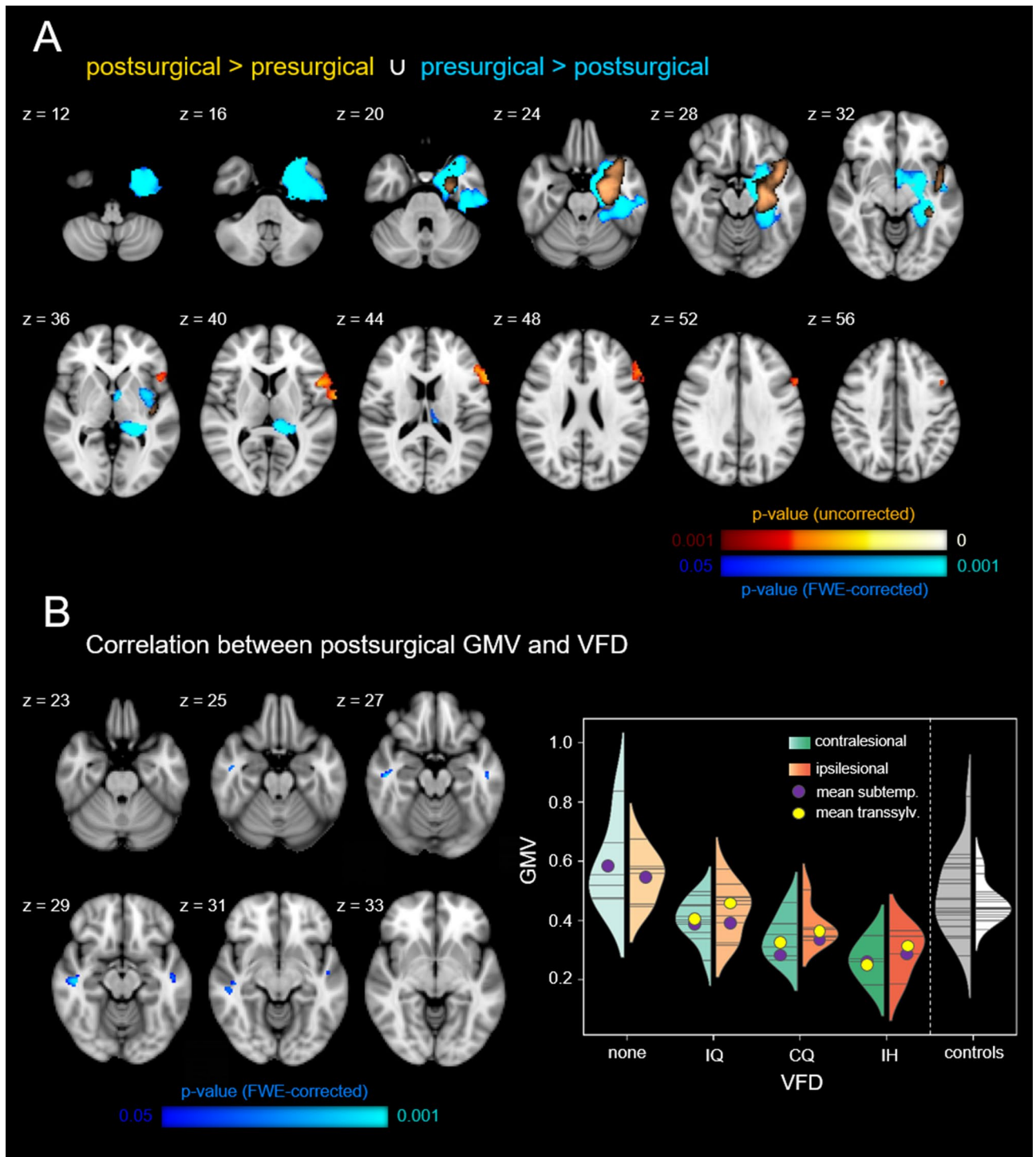


**Figure 1.** Canonical lesion mask and voxel-based lesion symptom mapping. **(A)** Canonical lesion mask after non-linear registration of all patients showing the overall lesion extent. **(B)** VLSM analysis showing a significant association between lesioned voxels and postsurgical VFDs (FWE-corr.  $p < 0.05$ ). Green outlines indicate clusters reaching significance.  $z$  indicates axial level in MNI space.

patient group. The opposite contrast as well as the correlation analysis on presurgical DTI scans did not yield any significant results.

**Connectivity differences between groups.** Comparing pre- and postsurgical mean connectivity matrices, sAH can be appreciated in the zeroed postsurgical connections including amygdala and hippocampus (see Supplementary Fig. S1). Apart from this obvious observation, a slight overall drop in streamline count of connections within the ipsilesional hemisphere (upper left quadrant of connectivity matrices) in both the VFD and no VFD patient group can be seen. However, sole visualization of the connectivity matrices does not yield obvious differences between the two patient groups.

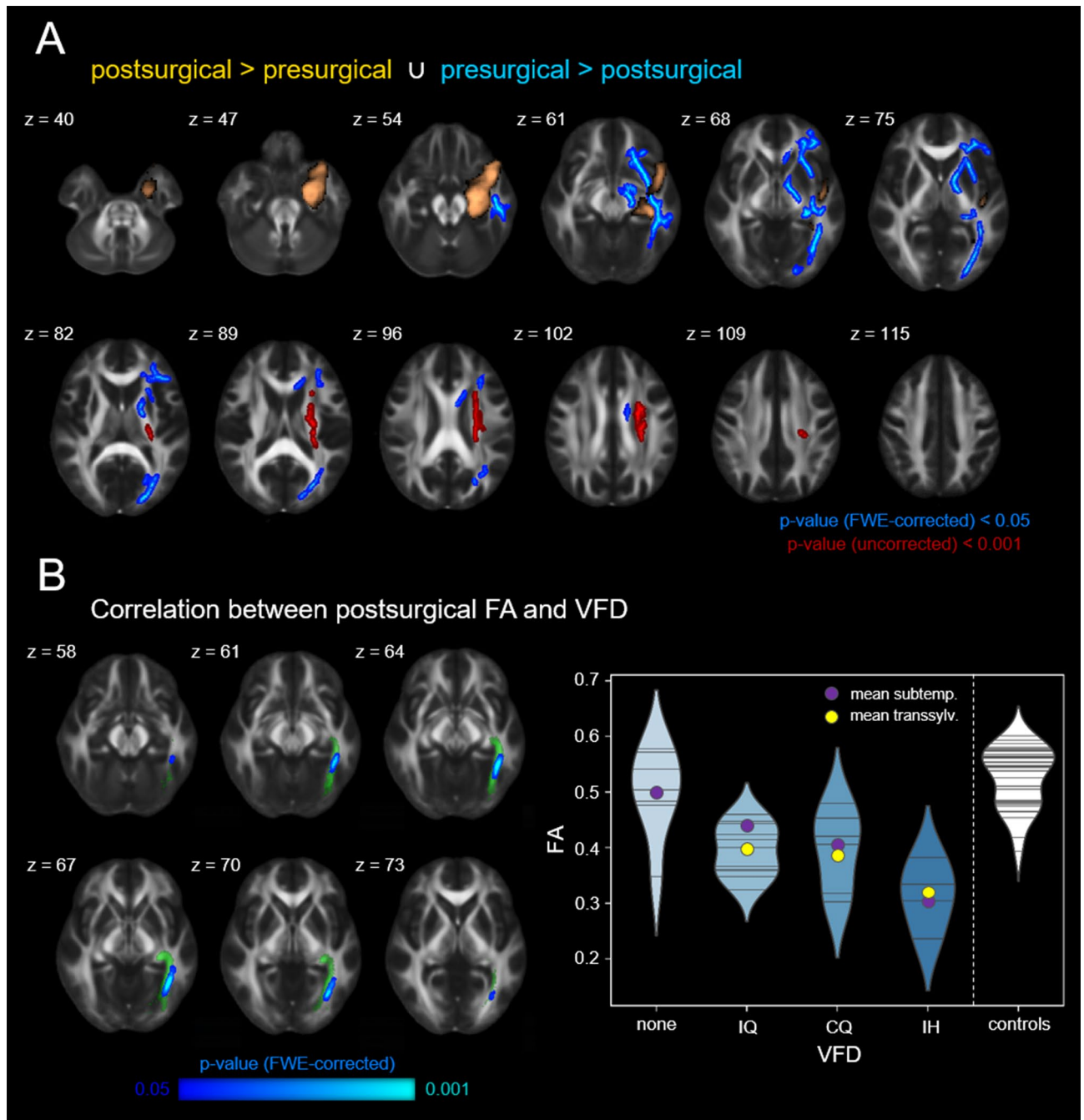
Using permutation-based paired  $t$ -tests between pre- and postsurgical scans, a decrease in streamline count of four edges including six nodes within the ipsilesional hemisphere was found in patients showing no VFD after sAH (FWE-corr.  $p < 0.05$ , see Fig. 4A). In contrast to that, patients with postsurgical VFD showed an extensive loss of connectivity in a total of 73 edges involving 28 different brain regions (FWE-corr.  $p < 0.05$ ; see Tables S1 and S2 for a list of all affected edges). Affected edges covered most of the ipsilesional temporal lobe, subcortical and prefrontal areas as well as temporo-occipital connections. Additionally, three brain regions from the contralateral hemisphere were included, namely the superior temporal gyrus, superior frontal gyrus, and pericalcarine cortex. No significantly increased streamline counts were found in the opposite contrast.



**Figure 2.** Voxel-based morphometry results. **(A)** Gray matter volume comparison of pre- and postsurgical T1-weighted scans of the transylvian patient subgroup (FWE-corr.  $p < 0.05$ ). **(B)** Correlation analysis between postsurgical gray matter volume and visual field deficit: the lower the gray matter volume, the higher the deficit (FWE-corr.  $p < 0.05$ ). The canonical lesion mask is shown in bronze. Violin plots show the distributions of mean gray matter volumes within the contra- and ipsilesional cluster with each horizontal line representing the readout of one patient/control.  $z$  indicates axial level in MNI space. *GMV* gray matter volume, *IQ* incomplete homonymous quadrantopia, *CQ* complete homonymous quadrantopia, *IH* incomplete homonymous hemianopia.

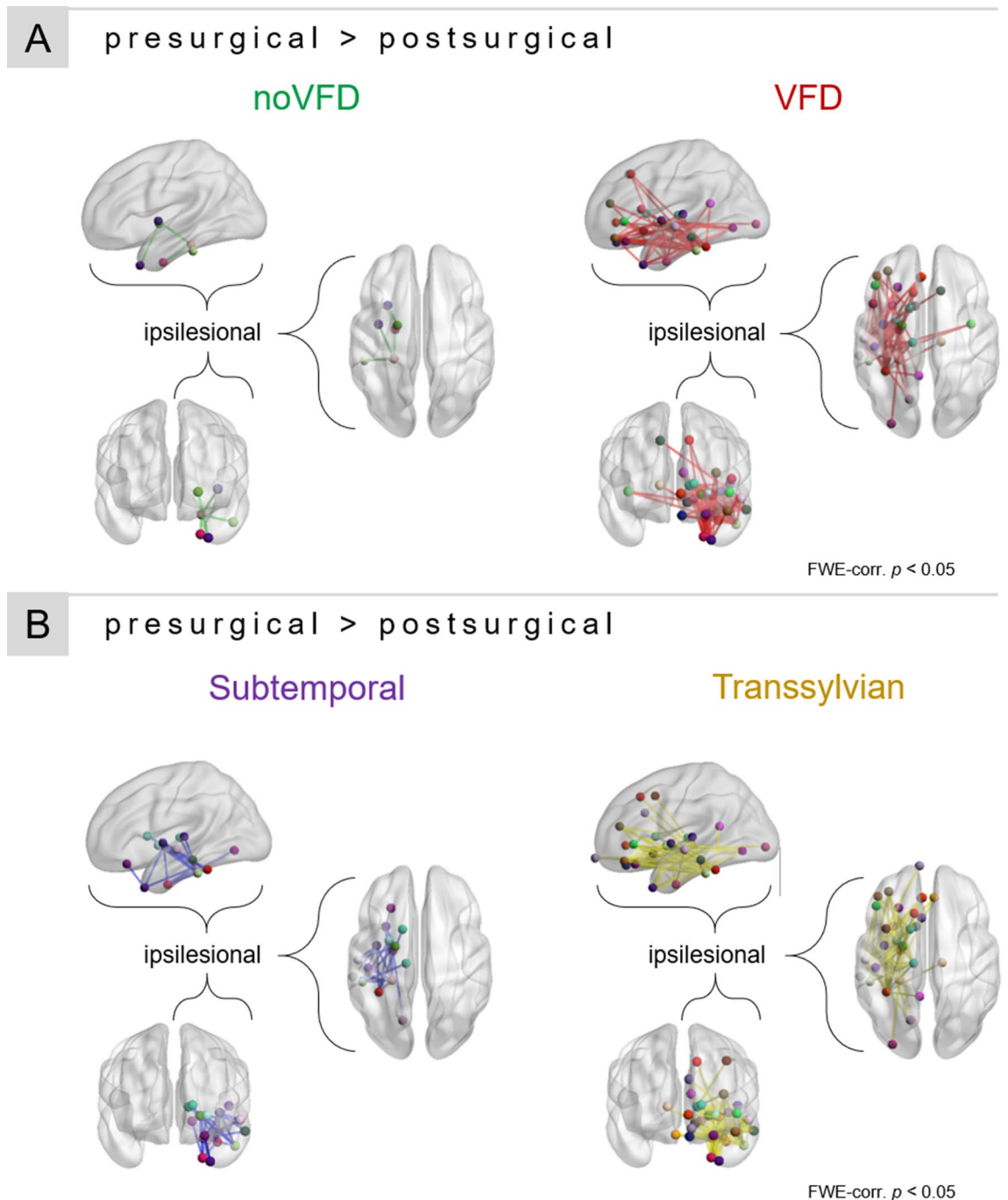
Splitting the sample according to surgical procedures, a similar pattern of connectivity differences in the pre- and postsurgical connectome comparisons can be appreciated: after subtemporal SAH, a significant decrease





**Figure 3.** Tract-based spatial statistics results. (A) Comparison of fractional anisotropy in pre- and postsurgical FA maps of the transylvian patient subgroup (FWE-corr.  $p < 0.05$ ). (B) Correlation analysis between postsurgical fractional anisotropy and visual field deficit: the lower the fractional anisotropy, the larger the visual field deficit (FWE-corr.  $p < 0.05$ ). The trajectory of the ipsilesional optic radiation as determined by probabilistic tractography is depicted in green. The canonical lesion mask is shown in bronze. Violin plots show the distributions of mean FA values within the significant cluster with each horizontal line representing the readout of one patient/control.  $z$  indicates axial level in MNI space. FA fractional anisotropy, IQ incomplete ipsilesional quadrantanopia, CQ complete homonymous quadrantanopia, IH incomplete homonymous hemianopia.

in streamline count was seen in 24 strictly ipsilesional edges spanning 15 nodes mostly involving the temporal lobe and subcortical brain regions (FWE-corr.  $p < 0.05$ , see Fig. 4B). In comparison to that, a more widespread loss in in connectivity was found after transylvian sAH, showing decreased streamline counts in 70 mostly ipsilesional edges involving 29 brain regions, two of them on the contralateral hemisphere (fusiform gyrus, superior temporal gyrus) (FWE-corr.  $p < 0.05$ , see Tables S3 and S4 for a list of all affected edges). The opposite contrast did not yield any significant results.



**Figure 4.** Connectivity-based statistics results. (A) Patients with postsurgical VFD showed a more widespread decrease in streamline count in the longitudinal comparison of pre- and postsurgical connectomes than patients without VFD. (B) Patients who underwent a transylvian sAH showed more widespread connectivity decreases than patients who underwent a subtemporal approach.

**Graph-theoretic differences between VFD and noVFD group.** Global graph-theoretic measures were read out for a cross-sectional comparison of patients with and without postsurgical VFD in both pre- and postsurgical networks. Despite the large deviation in the longitudinal comparison, no significant differences were found between the two groups, neither in presurgical nor postsurgical networks, at a Bonferroni-corrected  $p$ -value threshold of 0.005. Also, a more lenient threshold of uncorrected  $p < 0.05$  did not yield any significant differences.

**Classifier performances on VFD prediction.** Off-the-shelf classifier performances on presurgical prediction of postsurgical VFD differed vastly: due to the unbalanced dataset, most classifiers showed a high sen-

Model	TP	FP	TN	FN	Specificity	Sensitivity	PPV	NPV	Accuracy	F1 score	$p$
k-nearest neighbors	17	7	0	4	0.00	80.95	70.83	0.00	60.71	0.76	0.94
Gaussian Naive Bayes	20	6	1	1	14.29	95.24	76.92	50.00	75.00	0.85	0.14
Logistic regression	16	6	1	5	14.29	76.19	72.73	16.66	60.71	0.74	0.7
Decision tree	18	4	3	3	42.86	85.71	81.82	50.00	75.00	0.84	0.21
SVM	14	6	1	7	14.29	66.66	70.00	12.50	53.57	0.68	0.86
Random forest	19	7	0	2	0.00	90.41	73.08	0.00	68.86	0.81	0.95
Extremely randomized trees	17	6	1	4	14.29	80.95	73.91	20.00	64.29	0.77	0.76
Gradient tree boosting	19	3	4	2	57.14	90.41	86.36	66.67	82.14	0.88	0.13
AdaBoost	19	3	4	2	57.14	90.41	86.36	66.67	82.14	0.88	0.13
ANN meta-classifier	20	1	6	1	<b>85.71</b>	<b>95.24</b>	<b>95.24</b>	<b>85.71</b>	<b>92.86</b>	<b>0.95</b>	<b>*0.03</b>

**Table 1.** Classifier performances. Bold values indicate best performing classifier. *TP* = true positives, *FP* false positives, *TN* true negatives, *FN* false negatives, *PPV* positive predictive value, *NPV* negative predictive value.

sitivity in positively predicting VFD ranging from 66.67 to 90.41%, only few classifiers were able to distinguish patients without postsurgical VFD above chance level, yielding specificity scores of maximally 57.14%. Gradient Tree Boosting and AdaBoost algorithms showed the strongest performance of individual classifiers, both with a total accuracy of 82.14% (sensitivity = 90.41%, specificity = 57.14%). However, none of the individual classifier reached significance in the permutation test (see Table 1 for a full performance report).

In contrast to the individual classifiers, the ANN meta-classifier based on two boosting algorithms showed an increase in performance with a sensitivity of 95.24% and a specificity of 85.71%, yielding a total accuracy of 92.86%. Permutation testing confirmed a significant deviation from the performance on null distributions ( $p < 0.05$ ). The post-hoc analysis of feature importances of the two boosting classifiers revealed, despite having the same individual classification accuracy, a different underlying weighting of connections. Specifically, the AdaBoost algorithm focused mainly on the connections inferior temporal gyrus to precuneus (Gini = 0.5), caudal anterior cingulate gyrus to nucleus accumbens (Gini = 0.28), and caudal anterior cingulate to rostral middle frontal gyrus (Gini = 0.2). In contrast to that, Gradient Tree Boosting showed a more widespread weighting of connections and a strong bias towards the connection between paracentral and postcentral gyrus (Gini = 0.41; see Fig. 5A). In exemplary connectomes, a lower streamline count in the connection from e.g. the inferior temporal gyrus to precuneus but a stronger overall connectivity in prefrontal areas can be observed in the patient without VFD compared to the exemplary patient with postsurgical VFD (see Fig. 5B).

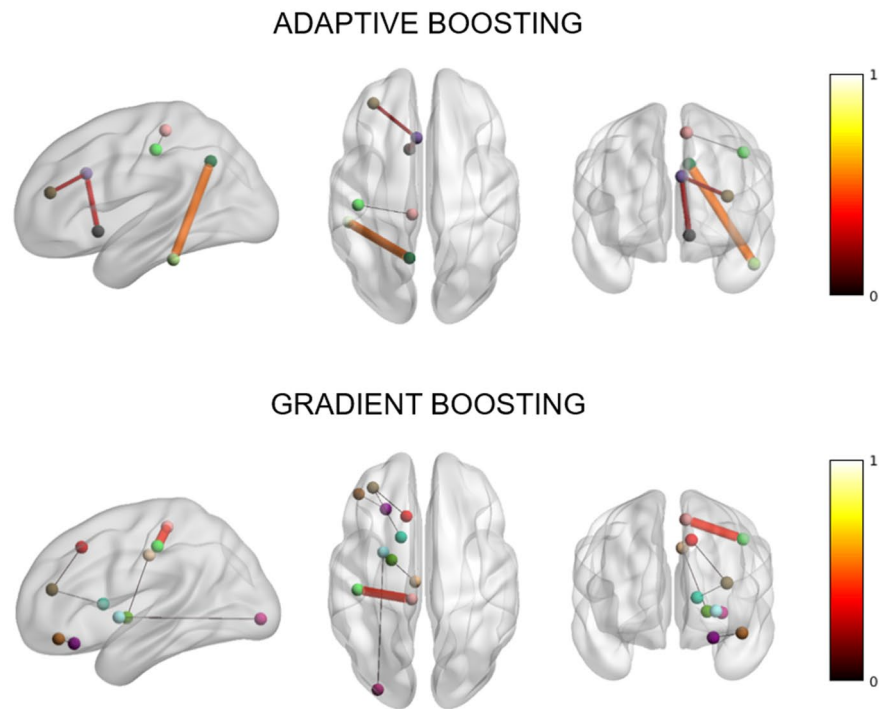
In an additional explorative analysis, we added the information about the surgical procedure to the ANN meta-classifier. Using this additional information about the surgical planning further boosted the performance of the classifier to 96.43% accuracy. Other metrics can be found in the Supplementary Table S5.

## Discussion

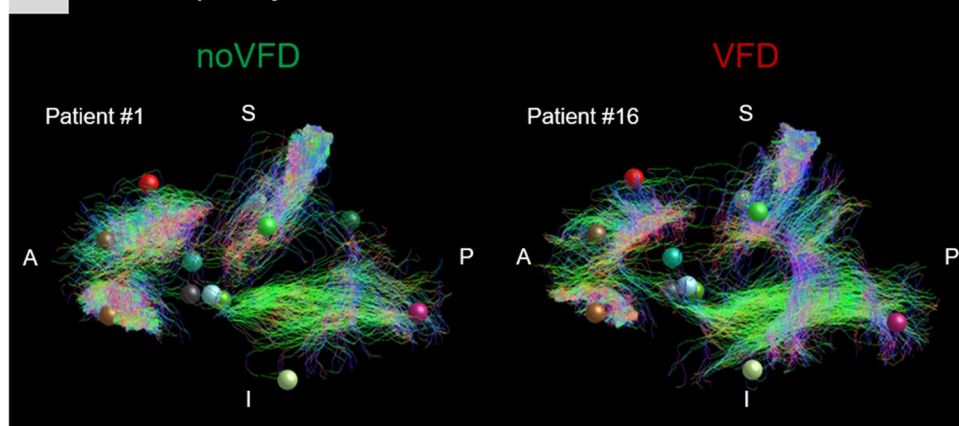
In this study, we set out to find presurgical correlates of postsurgical VFD. While we here describe abundant postsurgical differences between patients with and without VFD in both gray and white matter structures, we could not find presurgical differences, neither on the level of voxels nor on the level of the respective structural networks. Despite the missing statistical significance, we could utilize supervised machine learning algorithms to extract patterns that seemingly distinguish these two patient groups purely based on presurgical structural connectomes with above-chance accuracies.

Imaging analyses of patients undergoing resective surgery in the temporal lobe have been repetitively performed<sup>23–27</sup>. This study is the first to globally relate gray and white matter sequelae of surgery to VFD. The structural changes observed are generally in line with results of these previous studies: both degeneration and neuroplastic reorganization following epilepsy surgery can occur and are mirrored by decrease or increase in gray matter volume or fractional anisotropy<sup>28,29</sup>. The voxel-based lesion-symptom mapping and correlation analyses allow to identify white matter changes associated with VFD. It should be added that these results may be correlates or causal links of the VFD. Most importantly the differentiation between causal links and correlates can only be undertaken on the basis of common knowledge about anatomy and physiology of the visual system. While the changes observed in the ML replicate the findings of previous studies and are somewhat expected, the bilateral nature of the VBM-cluster in the posterior division of both the ipsi- and the contralesional middle temporal gyri is somewhat surprising. This may be explained by *diaschisis*/secondary degeneration of so-called homotopic connectivity: *diaschisis* means the post-lesional change of brain structures remote from and connected with the anatomical site of damage<sup>30</sup>. *Homotopic connectivity* describes the special connectedness of mirror areas of the brain hemispheres<sup>31</sup>. This may also explain the white matter changes found in the external capsule where cortico-cortical association fibers are known to pass, which connect the two hemispheres. Accordingly, bihemispheric white matter changes have been identified previously as sequelae of temporal lobe surgery<sup>24</sup>. The possible reorganization, however, was not reflected in global network metrics based on the structural connectomes. The other surprising result pertains the location and form of the TBSS-cluster correlating with the degree of the VFD. The present study is the first to show changes in the course of optic radiation after sAH with an objective, i.e. ROI-independent, approach. However, it remains open to speculation why this cluster was found in the sagittal stratum and not more anterior in the temporal lobe. Possible reasons include the strong interindividual anatomical variability of the anterior part of the ML. The anatomical distance between the temporal pole and ML

## A Feature importances



## B Exemplary connectomes



**Figure 5.** Features included in the ANN meta-classifier. **(A)** Specific post-hoc feature importances of the two boosting classifiers are visualized within the connectome. Both colour value and edge width indicate the respective Gini importance. **(B)** Exemplary streamline connectomes of the edges used by the ANN meta-classifier of a patient without (patient #1) and a patient with (patient #16) postsurgical VFD. Connectomes were downsampled to a total of 5000 streamlines for visualization purposes only.

varies from 22 to 37 mm<sup>32</sup>. As the posterior part of the ML is not known to show the same degree of anatomical variability, only here diffusivity parameters could be related to the severity of VFD. Another possible explanation is the convergence of the anterior, central, and the posterior bundle of the optic radiation in the sagittal stratum leading to maximal fiber coherence, which makes the analysis of diffusivity parameters more sensitive. Interestingly, postoperative resection size and preoperative Euclidean distance between the temporal pole and the most anterior part of the ML did, in contrast to previous studies<sup>17,22</sup>, not serve as a predictor/correlate of VFD in the current study. Despite the relatively clear causal chain with the lesioning of ML, a clear consensus on which correlates or predictors are most valid has not been reached by previous studies, let alone has one marker been implemented in the clinical routine.



An important limitation of our study is the variability in the interval between timepoint of surgery and postoperative scan, ranging from 2 to 21 months. While we did not find any linear relationships between the surgery-scan-interval in neither VBM nor TBSS analysis and controlled for this variable in all pre- vs. postsurgical analyses, the possible non-linear effects of this variability are still unaccounted for. However, reducing this variability by excluding patients with a surgery-scan-interval greater than 12 months did not alter the overall results from our analyses (see Suppl. Fig. S1).

The discrepancy between the absence of presurgical structural correlates of VFDs and the seemingly successful prediction of VFDs using automated classifiers may seem contradictory, however, the difference between the outcome of statistical inference and statistical learning methods (the latter being classically evaluated using cross-validation) has been discussed multiple times<sup>33,34</sup>. Nevertheless, our explorative classification analysis, testing a multitude of *off-the-shelf* algorithms, should be treated with great caution. Specifically, the use of a *leave-one-out* cross-validation scheme in small datasets is thoroughly discussed in the current neuroimaging literature<sup>35</sup>. It was shown that already minimal changes in the design of a supervised classifier can lead to so-called *vibration effects* which manifest in large differences in the standard performance evaluation metrics. Therefore, the real generalization performance of a classifier cannot be reliably approximated in small datasets as the one reported in this study. To give some information about the reliability of our classifiers, we utilized a permutation test randomly swapping the group labels and evaluating their performance on each permutation. The outcomes of this test already show that especially the simpler classifiers fail to extract real group differences for classification indicating instable generalization performances. On the same note, another big limitation of our classification approach is the strong class imbalance of our dataset. Only seven of the 28 patients in our study showed no postsurgical VFD. A trivial classifier, always predicting postsurgical VFD irrespective of the input, would thereby already result in an accuracy of 75%. Several of the individual classifiers, such as k-nearest neighbours, Naïve Bayes or Random Forest, show exactly this kind of vulnerability for the dominance of the VFD class. However, especially the adaptive and gradient boosting algorithms, developed to overcome this kind of dataset bias, seem to find a more balanced decision boundary between the two classes and reach a higher discriminatory power, also reflected in the different distribution of feature importances over individual connections. The stacking of these by a simple ANN as meta-classifier seemingly combines the best of both worlds. Lastly, feeding this classifier with information about the surgical procedure further boosts its performance. This near-perfect performance, however, can be deceptive and should be, additionally to the already mentioned limitations, seen in context with the unbalanced number of patients undergoing transsylvian and subtemporal amygdalohippocampectomy. While this is unfortunate in the context of our analyses, it results from the challenge of prospectively acquiring data in a clinical setting. Therefore, taking together both the size and the class-imbalance of the dataset, the generalizability of even the best performing classifier in this study cannot be guaranteed and should therefore be understood as mere explorative analysis that may a hint towards a possible predictive approach in the future, that needs to be developed in a larger, balanced sample and validated on yet another, external dataset.

## Methods

**Study design and patient cohort.** Twenty-eight patients with pharmacoresistant mesial temporal lobe epilepsy were prospectively included in our study (mean age  $\pm$  SD: 38.82  $\pm$  12.80; 15 females). All patients underwent presurgical assessment at the Department of Epileptology followed by a sAH at the Department of Neurosurgery of the University Hospital Bonn between 2009 and 2012. Ten patients were operated using a subtemporal access and 18 patients using a transsylvian access for clinical reasons. T1 and DTI scans were acquired one day before and several months after (mean interval  $\pm$  SD: 6.6  $\pm$  5.3 months) the surgical procedure. At the same days of the pre- and postoperative scans, the presence and severity of VFD was assessed using a Twinfield perimeter (Oculus Inc., Wetzlar, Germany), which is able to perform automated static and kinetic perimetry according to the Goldmann standard at the Department of Ophthalmology of the University Hospital Bonn. See Table 2 for details. Additional to the patient cohort, T1 and DTI scans of 32 healthy controls (mean age  $\pm$  SD: 32.52  $\pm$  6.70; 16 females) were acquired. This study and all its experimental protocols were approved by the Institutional Review Board of the medical faculty of the University of Bonn. All methods were performed in accordance with the guidelines and regulations of this ethics board and in accordance with the Declaration of Helsinki. Informed consent was obtained from all participants and/or their legal guardians. The data that support the findings of this study are available on request from the corresponding author. Datasets are not publicly available as they contain information that could comprise the privacy of research participants.

**Image acquisition and preprocessing.** For all patients pre- and postsurgical T1 and DTI scans were acquired using a Siemens Magnetom Trio (3 T) MRI-scanner. Scans were acquired with an 8-channel head receive coil. We ran a 3D T1-weighted anatomical sequence (resolution = 1.0  $\times$  1.0  $\times$  1.0 mm, TR = 1570 ms, TE = 3.42 ms, flip angle = 15°) and a diffusion-weighted single shot dual echo spin-echo echo planar imaging sequence (resolution = 1.72  $\times$  1.72  $\times$  1.7 mm<sup>3</sup>, TR = 12,000 ms, TE = 100 ms, flip angle = 90°) with 60 directions and a b-value of 1000 s/mm<sup>2</sup> as well as six baseline scans with a b-value of 0 s/mm<sup>2</sup>. Preprocessing of T1 and DTI datasets was realized using FMRIB's Software Library 5.0 (FSL)<sup>36</sup> and the Tolerably Obsessive Registration and Tensor Optimization Indolent Software Ensemble (TORTOISE)<sup>37</sup>. T1 scans were skull-stripped and corrected for b0 field inhomogeneities. DTI scans were corrected for susceptibility-induced geometric distortions using a constraint registration approach<sup>38</sup>. In the same registration step, DTI scans were corrected for within-subject motion and eddy-currents, keeping interpolation effects minimal. Finally, a diffusion tensor model was fitted voxel-wise and fractional anisotropy (FA) values were calculated. For all following analyses, ipsilesional hemispheres were flipped to the same side.

ID	Age at onset of seizures (years)	Age at surgery (years)	Surgery–post-scan–interval (months)	Neurosurgical approach	Side	Preoperative VFD	Postoperative VFD	Postoperative driving ability	ILAE outcome (last follow-up)
1	1–5	16–20	21	Transsylv	R	0	1	No	5
2	11–15	16–20	4	Transsylv	R	0	1	Yes	1
3	16–20	41–45	19	Subtemp	L	0	0	Yes	4
4	21–25	46–50	12	Transsylv	L	1	1	No	1
5	1–5	31–35	11	Subtemp	R	0	0	Yes	1
6	11–15	51–55	10	Subtemp	L	0	2	No	3
7	1–5	46–50	6	Subtemp	L	0	0	Yes	1
8	11–15	31–35	16	Transsylv	L	0	2	No	1
9	1–5	41–45	5	Transsylv	L	0	2	No	1
10	1–5	46–50	5	Transsylv	R	0	1	Yes	1
11	26–30	46–50	11	Transsylv	L	0	2	No	1
12	26–30	31–35	3	Transsylv	L	0	2	No	5
13	1–5	21–25	4	Transsylv	L	0	1	Yes	3
14	1–5	41–45	3	Transsylv	R	0	1	Yes	1
14	16–20	46–50	3	Subtemp	L	2	0	Yes	4
16	31–35	71–75	3	Transsylv	L	0	2	No	1
17	31–35	36–40	3	Transsylv	R	1	3	No	4
18	11–15	51–55	3	Subtemp	L	0	1	Yes	1
19	16–20	26–30	3	Transsylv	L	0	1	No	1
20	21–25	26–30	6	Transsylv	R	0	3	No	3
21	36–40	36–40	3	Transsylv	L	0	1	No	1
22	26–30	31–35	2	Transsylv	L	0	1	Yes	1
23	11–15	21–25	5	Transsylv	R	1	1	No	1
24	36–40	36–40	11	Subtemp	L	0	0	Yes	1
25	6–10	36–40	3	Transsylv	R	0	3	No	1
26	21–25	61–65	4	Subtemp	L	2	3	No	3
27	11–15	26–30	3	Subtemp	L	0	0	Yes	3
28	16–20	26–30	2	Subtemp	R	0	0	Yes	1

**Table 2.** Demographical and clinical characteristics of the patient group. Classification of visual field deficit: 0 = no anopia, 1 = incomplete homonymous quadrantanopia, 2 = complete homonymous quadrantanopia, 3 = incomplete homonymous hemianopia. (L = left, R = Right, subtemp = subtemporal access, transsylv = transylvian access).

**Voxel-based lesion-symptom mapping.** To analyze the tissue-independent influence of the specific lesion extent and location on the degree of VFD, we conducted a voxel-based lesion-symptom mapping (VLSM) analysis using the nonparametric mapping toolbox (NPM) implemented in MRICron version 8<sup>39</sup>. For all patients, lesion masks of the resection cavity were manually demarcated on the T1. Based on the individual lesion masks a colour-coded overlay map of all affected voxels across patients was calculated to provide an overview of the lesion extents (Fig. 1A). This was achieved by normalizing each patient's presurgical T1 scan to the MNI152 template using non-linear registrations<sup>36</sup>. These non-linear warp fields were in turn combined with matrices resulting of the linear registration of the postsurgical to the presurgical T1 scan and applied to the individual lesion masks. The normalized lesion masks were averaged and thresholded at 20%. Based on this normalization, the lesion masks were included in the voxel-based lesion-symptom mapping analysis. Here, we used a permutation-based Brunner–Munzel rank test to analyze the statistical contribution of each lesion voxel on the postsurgical binocular results of the automated perimetry. For valid statistical inference, only voxels affected in at least 10 patients were considered<sup>40</sup>. Results were corrected for multiple comparisons using a permutation-based family-wise error (FWE) correction. Clusters were considered significant at FWE-corrected  $p < 0.05$ .

**Voxel-based morphometry.** To analyze possible GM correlates of VFD the pre- and postsurgical T1 scans were entered into a voxel-based morphometry (VBM) analysis using the optimized FSL-VBM protocol<sup>41</sup>. First, the skull-stripped structural scans were segmented into GM partial volume estimates which were in turn non-linearly registered to the MNI152 standard space. For the computation of the postsurgical normalization warp fields the individual lesions were masked out, leading to a registration solely based on the intact brain tissue. From the resulting pre- and postsurgical normalized images of all patients, a study-specific and left–right symmetric brain template was calculated. All scans in native space were then registered to this template using again non-linear transformations and lesion masking. This step included the multiplication of the normalized partial volume estimates by the Jacobian of their respective warp field, to correct for local expansions and contractions

due to the non-linear component of the registration. These modulated GM partial volume estimates were then smoothed using an isotropic three-dimensional Gaussian kernel with a sigma of 3 mm. Permutation-based threshold-free cluster enhancement was applied to statistically evaluate GM differences between pre- and postsurgical scans (paired *t*-test) and correct for multiple comparisons<sup>42</sup>. All statistical models were adjusted for *between-scan interval*, *age at scan*, *surgery-scan interval*, and *TLE-laterality*.

**Tract-based spatial statistics.** We conducted a tract-based spatial statistics analysis (TBSS) of patients' pre- and postsurgical FA maps to identify possible WM correlates of VFDs<sup>43</sup>. All FA maps were registered to the MNI152 standard space using non-linear transformations. Precise normalization of postsurgical maps was achieved by masking lesional voxels in the computation of the warp fields. A mean FA image was created, masked with the canonical lesion mask (see Fig. 1A) and thinned, resulting in a mean FA skeleton representing the centers of all WM pathways common to the patient group. Each patient's aligned FA maps were then projected on to the mean skeleton mask and in turn used in a permutation-based statistical analysis. Parallel to the above described VBM analyses, statistical models comparing pre- and postsurgical scans as well as linear relationships between WM differences and VFD degree were implemented using two-dimensional threshold-free cluster enhancement and adjusted for *between-scan interval*, *age at scan*, *surgery-scan interval*, and *TLE-laterality*. To inform anatomical localization of the results, the optic radiation was reconstructed by means of probabilistic fiber tractography in every patient and a canonical optic radiation was generated (see Supplementary Material for a detailed description).

**Node delineation.** To estimate a structural connectome, we needed to delineate respective network nodes. For this purpose, we parcellated the presurgical T1-scans of all patients into 84 regions according to the Desikan-Killiany atlas using the standard FreeSurfer (version 5.0) processing stream<sup>44</sup>. Individual parcellations were visually controlled and, if needed, manually corrected. To delineate the same nodes in the postsurgical scans, manual lesion masks were constructed and presurgical brain parcellations were linearly registered to the respective postsurgical T1 volume. To account for possible registration biases due to postsurgical brain distortions, we utilized a robust registration method for longitudinal datasets to transform both pre- and postsurgical images to their midspace<sup>45</sup>. Thereby, the registration is not biased towards either the pre- or postsurgical scan. Registered parcellations were in turn multiplied by the inverted lesion mask, leaving us with postsurgical network nodes excluding the individual lesion. In a second step, both pre- and postsurgical nodes had to be linearly registered to the diffusion space. To ensure an accurate anatomical correspondence between T1 and DTI volumes, we calculated a mean b0 image, inverted its contrast and, finally, used boundary-based linear registration to translate our anatomical nodes to diffusion space<sup>46,47</sup>. All registration steps were controlled visually.

**Structural connectome processing.** It is known that the classical diffusion tensor model fails to adequately model the abundant regions of multiple fiber orientations in the brain, which in turn negatively affects the structural connectome construction<sup>48–50</sup>. For this reason, we constructed the connectivity-based structural connectome using a higher order diffusion model as implemented by MRtrix3<sup>51</sup>. First, a regular diffusion tensor model was fitted voxelwise for the whole DTI volume and fractional anisotropy (FA) values were calculated. Voxels showing the highest FA (>0.7, i.e. being the closest to a 'single fiber' voxel) determined the response function which in turn was used for the constrained spherical deconvolution of the diffusion-weighted signal estimating the specific fiber-orientation distribution (FOD) for every voxel<sup>52,53</sup>. Based on the voxelwise FODs a whole-brain probabilistic streamline tractography could be conducted. To mitigate the effect of spurious fiber trajectories in GM regions, we used "anatomically-constrained tractography" by segmenting the GM-WM interface as seed- and termination mask in diffusion space from the previously described parcellations of the T1 scans<sup>54</sup>. From this mask, 10,000,000 streamlines were dynamically seeded and progressed using a second-order integration over the FODs<sup>55</sup>. Further tracking parameters included a step size of 0.85 mm, minimum streamline length of 4 mm, maximum length of 250 mm and a FOD amplitude cut-off value of 0.05. Ensuring the robustness of streamline estimates, whole-brain tractograms were selectively filtered by a factor of 0.5, concluding tractograms of 5,000,000 streamlines each<sup>56</sup>. Finally, seed and termination points were mapped onto the parcellated network nodes and streamline counts between nodes were taken as edge weights of the connectivity graphs.

**Connectivity-based statistics.** Global differences in connectivity-graphs were assessed using the Network-based statistics toolbox<sup>57</sup>. Longitudinal comparisons of pre- and postsurgical networks were tested using paired *t*-tests. To avoid the report of obvious differences between pre- and postsurgical connections to and from amygdala and hippocampus, all connections of these regions were zeroed in presurgical connectomes as well. Significance was evaluated using network-based statistics including non-parametric permutation testing with 10,000 permutations to correct for the family-wise error rate. Results were considered significant at a FWE-corrected *p*-value below 0.05.

**Graph-theoretic measures.** To quantify and compare global network characteristics between patients with and without postsurgical VFD, we calculated graph-theoretic network measures using the Brain Connectivity Toolbox. The following global network-metrics were considered: (1) Average degree (describing the average number of links connected to a node), (2) average clustering coefficient (average number of triangles around nodes), (3) transitivity (normalized clustering coefficient for the individual degree of nodes), (4) density (fraction of present connections to possible connections), (5) routing efficiency (average inverse shortest path length), (6) assortativity (tendency of nodes to link to nodes with a similar degree), (7) coreness (amount of subgraphs comprising of nodes with the same degree), (8) average node betweenness centrality (average fraction

of all shortest paths containing a node), and (9) small-worldness (transitivity over the average shortest path length). Network measures were chosen to describe global features between the networks under consideration, namely the functional integration and segregation of the networks as well as the network resilience to lesions. For mathematical formulations and interpretations the interested reader is referred to the review by Rubinov and Sporns<sup>58</sup>. Group comparisons of network measures were conducted using two-tailed unpaired *t*-tests. Statistical inference was confirmed using bootstrap analyses with 10,000 times resampling.

**Connectome-based classification.** As an additional explorative analysis, we trained five of the most common supervised classification algorithms on the connectomes of the lesioned hemisphere using their default implementation in scikit-learn<sup>59</sup>: (1) k-nearest neighbors, (2) Gaussian Naïve Bayes, (3) logistic regression, (4) Decision Tree, and (5) Support Vector Machine. As our sample is unbalanced and small compared to typical statistical learning problems, we furthermore included four tree-based ensemble and boosting classifiers: (6) Random Forests, (7) Extremely Randomized Trees, (8) Gradient Tree Boosting, and (9) AdaBoost. Ensemble methods and boosting can be favorable for this type of datasets, as here weak (Decision Tree) learners are iteratively trained on the dataset and finally combined to a strong learner. After each added weak learner, the weight of misclassified samples gets increased and forcing the next learner to focus more on those samples, thereby counteracting the dominance of one class in the dataset<sup>60,61</sup>. The generalizability of all classifiers was evaluated using a *Leave-One-Out* cross-validation scheme. We statistically compared the individual classification performances to the classifiers' performance on 10,000 null distributions using permutation testing. Furthermore, to ensure that classifiers do not use spurious structural connections, all classifiers were trained on a sparse connectome including only connections with a minimum streamline count of 10 in all patients.

**Meta classifier construction.** To increase overall capacity of our presurgical classification procedure, we combined the two boosting classifiers in an additional ensemble learner by means of stacking. Here, the two base-classifiers were trained on our original presurgical sparse lesioned connectomes. A supervised meta-classifier in turn was added to integrate the outputs of the two classifiers as meta-features during training. Our meta-classifier in this case was a backpropagating artificial neural network (ANN) comprising of two hidden layers with 100 and 50 neurons using a rectified linear unit activation function and two sigmoid output neurons. Like the other supervised classifiers, our experimental stacking classifier was trained on the data using *Leave-One-Out* cross-validation and performance scores were statistically compared to 10,000 null distributions constructed using permutation testing. Furthermore, a post-hoc analysis of Gini importance according to the two boosting classifiers was conducted to visualize the specific connections utilized in the classification.

Received: 26 September 2020; Accepted: 24 December 2020

Published online: 14 January 2021

## References

1. Semah, F. *et al.* Is the underlying cause of epilepsy a major prognostic factor for recurrence?. *Neurology* **51**, 1256–1262 (1998).
2. Thijs, R. D., Surges, R., O'Brien, T. J. & Sander, J. W. Epilepsy in adults. *Lancet* **393**, 689–701 (2019).
3. Semah, F. & Ryvlin, P. Can we predict refractory epilepsy at the time of diagnosis?. *Epileptic Disord.* **7**(Suppl 1), S10–13 (2005).
4. Engel, J. *et al.* Early surgical therapy for drug-resistant temporal lobe epilepsy: A randomized trial. *JAMA* **307**, 922–930 (2012).
5. Wiebe, S., Blume, W. T., Girvin, J. P., Eliasziw, M. & Effectiveness and Efficiency of Surgery for Temporal Lobe Epilepsy Study Group. A randomized, controlled trial of surgery for temporal-lobe epilepsy. *N. Engl. J. Med.* **345**, 311–318 (2001).
6. Jain, P., Tomlinson, G., Snead, C., Sander, B. & Widjaja, E. Systematic review and network meta-analysis of resective surgery for mesial temporal lobe epilepsy. *J. Neurol. Neurosurg. Psychiatry* **89**, 1138–1144 (2018).
7. Yaşargil, M. G., Teddy, P. J. & Roth, P. Selective amygdalo-hippocampectomy. Operative anatomy and surgical technique. *Adv. Tech. Stand. Neurosurg.* **12**, 93–123 (1985).
8. Hori, T. *et al.* Subtemporal amygdalohippocampectomy for treating medically intractable temporal lobe epilepsy. *Neurosurgery* **33**, 50–57 (1993).
9. Schmeiser, B. *et al.* Transylvian selective amygdalohippocampectomy for mesiotemporal epilepsy: Experience with 162 procedures. *Neurosurgery* **80**, 454–464 (2017).
10. Lutz, M. T., Clusmann, H., Elger, C. E., Schramm, J. & Helmstaedter, C. Neuropsychological outcome after selective amygdalohippocampectomy with transylvian versus transcortical approach: A randomized prospective clinical trial of surgery for temporal lobe epilepsy. *Epilepsia* **45**, 809–816 (2004).
11. Winston, G. P. Epilepsy surgery, vision, and driving: what has surgery taught us and could modern imaging reduce the risk of visual deficits?. *Epilepsia* **54**, 1877–1888 (2013).
12. Jeelani, N. U. O. *et al.* 'Hemispherical asymmetry in the Meyer's Loop': A prospective study of visual-field deficits in 105 cases undergoing anterior temporal lobe resection for epilepsy. *J. Neurol. Neurosurg. Psychiatry* **81**, 985–991 (2010).
13. Yeni, S. N. *et al.* Visual field defects in selective amygdalohippocampectomy for hippocampal sclerosis: the fate of Meyer's loop during the transylvian approach to the temporal horn. *Neurosurgery* **63**, 507–513 (2008) (**Discussion** 513–515).
14. van Lanen, R. H. G. J. *et al.* Visual field deficits after epilepsy surgery: A new quantitative scoring method. *Acta Neurochir.* **160**, 1325–1336 (2018).
15. Delev, D. *et al.* Vision after trans-sylvian or temporobasal selective amygdalohippocampectomy: A prospective randomised trial. *Acta Neurochir.* **158**, 1757–1765 (2016).
16. Jäncke, L. The plastic human brain. *Restor. Neurol. Neurosci.* **27**, 521–538 (2009).
17. Yogarajah, M. *et al.* Defining Meyer's loop—temporal lobe resections, visual field deficits and diffusion tensor tractography. *Brain* **132**, 1656–1668 (2009).
18. Winston, G. P. *et al.* Optic radiation tractography and vision in anterior temporal lobe resection. *Ann. Neurol.* **71**, 334–341 (2012).
19. Winston, G. P. *et al.* Diffusion tensor imaging tractography to visualize the relationship of the optic radiation to epileptogenic lesions prior to neurosurgery. *Epilepsia* **52**, 1430–1438 (2011).



20. de Souza, J. P. S. A. S. *et al.* Fractional anisotropy of the optic radiations correlates with the visual field after epilepsy surgery. *Neuroradiology* **61**, 1425–1436 (2019).
21. Chen, X., Weigel, D., Ganslandt, O., Buchfelder, M. & Nimsky, C. Prediction of visual field deficits by diffusion tensor imaging in temporal lobe epilepsy surgery. *NeuroImage* **45**, 286–297 (2009).
22. Powell, H. W. R. *et al.* MR tractography predicts visual field defects following temporal lobe resection. *Neurology* **65**, 596–599 (2005).
23. Yogarajah, M. *et al.* The structural plasticity of white matter networks following anterior temporal lobe resection. *Brain* **133**, 2348–2364 (2010).
24. Schoene-Bake, J.-C. *et al.* Widespread affections of large fiber tracts in postoperative temporal lobe epilepsy. *NeuroImage* **46**, 569–576 (2009).
25. Winston, G. P., Stretton, J., Sidhu, M. K., Symms, M. R. & Duncan, J. S. Progressive white matter changes following anterior temporal lobe resection for epilepsy. *NeuroImage* **4**, 190–200 (2014).
26. Yasuda, C. L. *et al.* Dynamic changes in white and gray matter volume are associated with outcome of surgical treatment in temporal lobe epilepsy. *NeuroImage* **49**, 71–79 (2010).
27. McDonald, C. R. *et al.* Changes in fiber tract integrity and visual fields after anterior temporal lobectomy. *Neurology* **75**, 1631 (2010).
28. Bueteftisch, C. M. Role of the contralesional hemisphere in post-stroke recovery of upper extremity motor function. *Front. Neurol.* **6**, 00214 (2015).
29. Dodd, K. C., Nair, V. A. & Prabhakaran, V. Role of the contralesional vs ipsilesional hemisphere in stroke recovery. *Front. Hum. Neurosci.* **11**, 1–10 (2017).
30. Seitz, R. J. *et al.* The role of diaschisis in stroke recovery. *Stroke* **30**, 1844–1850 (1999).
31. Mancuso, L. *et al.* The homotopic connectivity of the functional brain: A meta-analytic approach. *Sci. Rep.* <https://doi.org/10.1038/s41598-019-40188-3> (2019).
32. Ebeling, U. & Reulen, H.-J. Neurosurgical topography of the optic radiation in the temporal lobe. *Acta Neurochir.* **92**, 29–36 (1988).
33. Bzdok, D. Classical Statistics and Statistical Learning in Imaging Neuroscience. *Front. Neurosci.* **11**, 1 (2017).
34. Bzdok, D., Altman, N. & Krzywinski, M. Statistics versus machine learning. *Nat. Methods* **15**, 233–234 (2018).
35. Varoquaux, G. Cross-validation failure: Small sample sizes lead to large error bars. *NeuroImage* **180**, 68–77 (2018).
36. Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W. & Smith, S. M. FSL. *NeuroImage* **62**, 782–790 (2012).
37. Pierpaoli, C. *et al.* TORTOISE: An integrated software package for processing of diffusion MRI data. *Conference proceedings* (2010).
38. Yao, X.-F. & Song, Z.-J. Deformable Registration for Geometric Distortion Correction of Diffusion Tensor Imaging. in *Computer Analysis of Images and Patterns* (eds. Real, P., Diaz-Pernil, D., Molina-Abril, H., Berciano, A. & Kropatsch, W.) 545–553 (Springer, New York, 2011). [https://doi.org/10.1007/978-3-642-23672-3\\_66](https://doi.org/10.1007/978-3-642-23672-3_66).
39. Rorden, C., Karnath, H.-O. & Bonilha, L. Improving lesion-symptom mapping. *J. Cogn. Neurosci.* **19**, 1081–1088 (2007).
40. Medina, J., Kimberg, D. Y., Chatterjee, A. & Coslett, H. B. Inappropriate usage of the Brunner–Munzel test in recent voxel-based lesion-symptom mapping studies. *Neuropsychologia* **48**, 341–343 (2010).
41. Douaud, G. *et al.* Schizophrenia delays and alters maturation of the brain in adolescence. *Brain* **132**, 2437–2448 (2009).
42. Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M. & Nichols, T. E. Permutation inference for the general linear model. *NeuroImage* **92**, 381–397 (2014).
43. Smith, S. M. *et al.* Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *NeuroImage* **31**, 1487–1505 (2006).
44. Fischl, B. *et al.* Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron* **33**, 341–355 (2002).
45. Reuter, M., Rosas, H. D. & Fischl, B. Highly accurate inverse consistent registration: A robust approach. *NeuroImage* **53**, 1181 (2010).
46. Bhushan, C. *et al.* Co-registration and distortion correction of diffusion and anatomical images based on inverse contrast normalization. *NeuroImage* **115**, 269–280 (2015).
47. Greve, D. N. & Fischl, B. Accurate and robust brain image alignment using boundary-based registration. *NeuroImage* **48**, 63–72 (2009).
48. Behrens, T. E. J., Berg, H. J., Jbabdi, S., Rushworth, M. F. S. & Woolrich, M. W. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain?. *NeuroImage* **34**, 144–155 (2007).
49. Sotiropoulos, S. N. & Zalesky, A. Building connectomes using diffusion MRI: Why, how and but. *NMR Biomed* **32** (2019).
50. Tournier, J.-D., Mori, S. & Leemans, A. Diffusion tensor imaging and beyond. *Magn. Reson. Med.* **65**, 1532–1556 (2011).
51. Tournier, J.-D. *et al.* MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage* **202**, 116137 (2019).
52. Tournier, J.-D., Calamante, F. & Connelly, A. Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. *NeuroImage* **35**, 1459–1472 (2007).
53. Tournier, J.-D., Calamante, F. & Connelly, A. Determination of the appropriate b value and number of gradient directions for high-angular-resolution diffusion-weighted imaging. *NMR Biomed.* **26**, 1775–1786 (2013).
54. Smith, R. E., Tournier, J.-D., Calamante, F. & Connelly, A. Anatomically-constrained tractography: Improved diffusion MRI streamlines tractography through effective use of anatomical information. *NeuroImage* **62**, 1924–1938 (2012).
55. Smith, R. E., Tournier, J.-D., Calamante, F. & Connelly, A. SIFT2: Enabling dense quantitative assessment of brain white matter connectivity using streamlines tractography. *NeuroImage* **119**, 338–351 (2015).
56. Smith, R. E., Tournier, J.-D., Calamante, F. & Connelly, A. S. I. F. T. Spherical-deconvolution informed filtering of tractograms. *NeuroImage* **67**, 298–312 (2013).
57. Zalesky, A., Fornito, A. & Bullmore, E. T. Network-based statistic: Identifying differences in brain networks. *NeuroImage* **53**, 1197–1207 (2010).
58. Rubinov, M. & Sporns, O. Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage* **52**, 1059–1069 (2010).
59. Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
60. Schapire, R. E. The strength of weak learnability. *Mach. Learn.* **5**, 197–227 (1990).
61. Freund, Y. & Schapire, R. E. A decision-theoretic generalization of on-line learning and an application to boosting. in *Computational Learning Theory* (ed. Vitányi, P.) 23–37 (Springer, New York, 1995). [https://doi.org/10.1007/3-540-59119-2\\_166](https://doi.org/10.1007/3-540-59119-2_166).

## Author contributions

Study concept and design: B.D., J.E., D.D., J.G., J.W., B.Wa., T.R. Acquisition, analysis or interpretation of the data: B.D., J.E., D.D., J.G., C.C.P., A.R., J.W., J.S., B.Wa., T.R. Drafting of the manuscript: B.D., T.R. Critical revision of the manuscript for intellectual content: All authors. Statistical analysis: B.D., J.E., J.G., C.C.P., G.L. Administrative, technical or material support: J.W., E.H., J.S., B.Wa., J.S.B., A.R., B.W., R.S., C.E.E., T.R. Study supervision: D.D., A.R., R.S., C.E.E., T.R.

## Funding

Open Access funding enabled and organized by Projekt DEAL. JG received support from the BonnNi Promotionsskolleg Neuroimmunology of the University of Bonn and the *Else-Kröner-Fresenius Stiftung*. CP was funded by a promotion scholarship of the BONFOR research commission of the medical faculty of the University of Bonn. TR was supported by the BONFOR research commission of the medical faculty of the University of Bonn. This work was supported by the “Verein zur Förderung der Epilepsieforschung”.

## Competing interests

JW received fees as speaker or consultant from Eisai, Bial and UCB. AR lectures for Guerbet and Bayer. He receives institutional study support by Guerbet and Bayer. CE reports personal fees from UCB, Desitin, BIAL, and Eisai. CE was Medical Director for the Life & Brain institute until 2015 and also served as Platform Leader for Cognitive Neuroscience until 2017. RS has received fees as speaker or consultant from Bial, Cyberonics, Desitin, Eisai, LivaNova, Novartis, and UCB Pharma. The remaining authors have no conflict of interest.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-020-80751-x>.

**Correspondence** and requests for materials should be addressed to T.R.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021